The Effectiveness of an Academic Literacy Intervention to Help University Freshmen Recognize and Resolve Inconsistencies Across Multiple Texts

Patty Baldwin

University of San Francisco, pattiey@sfsu.edu

Follow this and additional works at: https://repository.usfca.edu/diss

Part of the Curriculum and Instruction Commons, Higher Education Commons, and the Reading and Language Commons

Recommended Citation
Baldwin, Patty, "The Effectiveness of an Academic Literacy Intervention to Help University Freshmen Recognize and Resolve Inconsistencies Across Multiple Texts" (2014). Doctoral Dissertations. 107.

https://repository.usfca.edu/diss/107

This Dissertation is brought to you for free and open access by the Theses, Dissertations, Capstones and Projects at USF Scholarship: a digital repository @ Gleeson Library | Geschke Center. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of USF Scholarship: a digital repository @ Gleeson Library | Geschke Center. For more information, please contact repository@usfca.edu.
THE EFFECTIVENESS OF AN ACADEMIC LITERACY INTERVENTION TO HELP UNIVERSITY FRESHMEN RECOGNIZE AND RESOLVE INCONSISTENCIES ACROSS MULTIPLE TEXTS

A Dissertation Presented

to

The Faculty of the School of Education
Learning and Instruction Department

In Partial Fulfillment
of the Requirement for the Degree
Doctor of Education

by
Patricia Baldwin
San Francisco
May 2014
The Effectiveness Of An Academic Literacy Intervention To Help University Freshmen Recognize And Resolve Inconsistencies Across Multiple Texts

Students must independently complete academic literacy tasks—including reading analytically to identify problems, resolving problems that arise, and using writing to demonstrate advanced knowledge acquisition—if they are to be successful in courses across their university careers. However, a significant portion of students arrives at the university underprepared to meet these expectations for academic literacy.

The purpose of this study was to examine the effectiveness of an instructional intervention to help developmental-level freshmen acquire the academic literacy skills that experienced academic readers demonstrate in order to promote independent learning. The four-week instructional intervention focused on two aspects of advanced academic literacy: 1) identifying inconsistencies across multiple texts and 2) flexibly employing evaluative heuristics (sourcing, corroboration, & contextualization) in order to resolve inconsistencies. The study, which took place at a large, urban, public university over the course of five weeks in two intact sections of a developmental-level academic literacy course taught by one instructor, used a pre-experimental one group pretest-posttest design. Participants (N = 31) were administered the Multiple Text Tasks as a pretest and a posttest in order to measure three dependent variables: 1) the number of inconsistencies identified, 2) the number of evaluative heuristics used in writing, and 3) the number of evaluative heuristics used in reading.
More participants were categorized as High Use in their ability to recognize inconsistencies across multiple texts postintervention. This result was statistically significant. Although participants did increase their use of evaluative heuristics in writing and in reading postintervention, these results did not reach statistical significance. One unique finding was that developmental-level freshmen in this study used the contextualization heuristic at higher rates than in previous studies.

The results suggest that the instructional intervention contributed to an increase in the number of inconsistencies identified. The increase in evaluative heuristic use suggests that the intervention may have contributed to increased use of evaluative heuristics. However, the failure to reach statistical significance suggests that the intervention was not of adequate intensity or duration.
This dissertation, written under the direction of the candidate’s dissertation committee and approved by members of the committee, has been presented to and accepted by the Faculty of the School of Education in partial fulfillment of the requirements for the degree of Doctor of Education. The content and research methodologies presented in this work represent the work of the candidate alone.

Patricia Baldwin
Candidate
May 14, 2014
Date

Dissertation Committee

Dr. Xornam Apedoe
Chairperson, Dr. Xornam Apedoe
May 14, 2014
Date

Dr. Yvonne Bui
Dr. Yvonne Bui
May 14, 2014
Date

Dr. Helen Maniates
Dr. Helen Maniates
May 14, 2014
Date
# TABLE OF CONTENTS

ABSTRACT.......................................................................................................................... ii

SIGNATURE PAGE............................................................................................................... iv

TABLE OF CONTENTS........................................................................................................ v

LIST OF TABLES.................................................................................................................. viii

LIST OF FIGURES.............................................................................................................. ix

I. INTRODUCTION TO THE STUDY ................................................................................. 1

  Statement of the Problem ............................................................................................... 1
    *Academic Literacy Expectations in Higher Education* ............................................... 1
    *The Role of Developmental-level Literacy Courses* ............................................... 2
    *How Inexperienced Freshmen Differ from Those Experienced with Academic
      Literacy* ...................................................................................................................... 3
    *Previous Investigations into Academic Literacy Behaviors* ..................................... 6
  Purpose of the Study ....................................................................................................... 9
  Theoretical Rationale .................................................................................................... 10
  Background and Need ................................................................................................... 16
    *Wineburg Makes Academic Literacy Explicit* ......................................................... 17
    *Research Need for the Study* ................................................................................... 21
    *Components of the Instructional Intervention* ......................................................... 26
  Significance .................................................................................................................... 28
  Research Questions ....................................................................................................... 30
  Definition of Terms ....................................................................................................... 31

II. REVIEW OF LITERATURE.............................................................................................. 38

  Developing Academic Literacy: Recognizing Inconsistencies .................................... 38
    *Students are Underprepared for University-level Academic Literacy Demands* ...... 39
    *Expert Readers Identify Inconsistencies* ................................................................. 41
    *Inexperienced Readers Do Not Recognize Inconsistencies* ..................................... 43
    *The Difficulty Paper* ............................................................................................... 45
    *Measuring Recognition of Inconsistencies* ............................................................. 51
    *Summary* .................................................................................................................. 52
  Developing Academic Literacy: Using Evaluative Heuristics .................................... 52
    *Academic Literacy is the Goal* ................................................................................ 54
    *Experienced Readers/Writers Use Evaluative Heuristics to Resolve Inconsistencies*
    ........................................................................................................................................ 56
    *Inexperienced Readers Rarely Evaluate Information Sources* ............................... 70
Previous Instructional Interventions for Evaluating Sources ............................................. 85
Methodological Considerations for Measuring Evaluative Heuristic Use ................ 102
Summary of the Literature on Evaluating Sources ......................................................... 106

III. METHODOLOGY ........................................................................................................ 109
Research Design ................................................................................................................. 110
Setting and Participants .................................................................................................... 112
  Setting ............................................................................................................................... 112
  Instructor .......................................................................................................................... 115
  Participants ......................................................................................................................... 115
Protection of Human Subjects .......................................................................................... 120
Instrumentation .................................................................................................................. 121
  Demographic Questionnaire ............................................................................................ 121
  Topic Familiarity ................................................................................................................ 124
  The Multiple Text Tasks ................................................................................................... 125
  Listing Inconsistencies ..................................................................................................... 126
  Decision Essay .................................................................................................................. 130
  Justify Trustworthiness ..................................................................................................... 139
Procedures .......................................................................................................................... 145
  Administration of Instruments ........................................................................................ 145
  Treatment Description ..................................................................................................... 147
  Instructional Intervention ................................................................................................ 156
Data Collection and Analysis ............................................................................................. 160
  Data Collection Procedures ........................................................................................... 160
  Data Analysis Procedures ............................................................................................... 161
Qualifications of the Researcher ......................................................................................... 164
Summary ............................................................................................................................... 165

IV. RESULTS ..................................................................................................................... 166
Results ................................................................................................................................. 170
  Research Question 1 ....................................................................................................... 170
  Research Question 2 ....................................................................................................... 172
  Research Question 3 ....................................................................................................... 178
  Summary of Results ....................................................................................................... 184
Scoring Anomalies .............................................................................................................. 185
  Difficulty to Categorize ................................................................................................. 186
  Challenges of the Demographic: Inexperienced Writers ............................................. 188
  Inexperienced Reader: How Wrong? ............................................................................ 191
  Difficulty Applying Consistent Scoring Standards ..................................................... 191
  Decision Rules ............................................................................................................... 192
  Summary .......................................................................................................................... 196
<table>
<thead>
<tr>
<th>Summary</th>
<th>197</th>
</tr>
</thead>
<tbody>
<tr>
<td>V. CONCLUSIONS</td>
<td>199</td>
</tr>
<tr>
<td>Summary of the Study</td>
<td>199</td>
</tr>
<tr>
<td>Summary of the Findings</td>
<td>203</td>
</tr>
<tr>
<td>Discussion of Results</td>
<td>204</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>217</td>
</tr>
<tr>
<td>Conclusions</td>
<td>219</td>
</tr>
<tr>
<td>Limitations</td>
<td>230</td>
</tr>
<tr>
<td>Implications for Future Practice</td>
<td>232</td>
</tr>
<tr>
<td>Recommendations for Future Research</td>
<td>235</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>241</td>
</tr>
</tbody>
</table>

**APPENDIXES**

<p>| Appendix A: Data Collection                                            | 250 |
| Appendix B: Multiple Text Tasks, based on Battle of Lexington          | 252 |
| Appendix C: Scoring Guide for Listing Inconsistencies                 | 266 |
| Appendix D: Evaluative Heuristics Scoring Guide &amp; Rubric              | 269 |
| Appendix E: PowerPoint Slides for Academic Literacy Expectations      |      |
| Appendix F: Anticipation Guide for Academic Literacy Expectations:    |      |
| Appendix G: Difficulty Paper Assignment and Feedback Sheet            | 286 |
| Appendix H: Introduction to the Evaluative Heuristics                 | 289 |
| Appendix I: Informed Consent for Research Participation               | 295 |
| Appendix J: Demographic Questionnaire                                  | 298 |
| Appendix K: Letter of Permission from Instructor to Participant in the Study | 300 |
| Appendix L: Education Readings for the Instructional Intervention      | 303 |
| Appendix M: Model Difficulty Paper                                    | 305 |
| Appendix N: Question Types Handout                                    | 320 |
| Appendix O: Contextualization Application Activity Handout            | 322 |
| Appendix P: Conditional Knowledge of Sourcing Activity                | 324 |</p>
<table>
<thead>
<tr>
<th>Table</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Evaluative Heuristics Used by Experienced Readers</td>
<td>5</td>
</tr>
<tr>
<td>2. Variables and Measurement Instruments</td>
<td>111</td>
</tr>
<tr>
<td>3. Data Collection</td>
<td>112</td>
</tr>
<tr>
<td>4. Mean and Standard Deviation for Topic Familiarity</td>
<td>118</td>
</tr>
<tr>
<td>5. Language Background Categories</td>
<td>123</td>
</tr>
<tr>
<td>6. Sample Inconsistencies from the Battle of Lexington Document Set</td>
<td>127</td>
</tr>
<tr>
<td>7. Scoring Ranges for Classification as Low Use and High Use</td>
<td>129</td>
</tr>
<tr>
<td>8. Aspects of Evaluative Heuristics</td>
<td>134</td>
</tr>
<tr>
<td>9. Overview of the Instructional Unit</td>
<td>148</td>
</tr>
<tr>
<td>10. Gradual Release of Responsibility in the Instructional Intervention</td>
<td>149</td>
</tr>
<tr>
<td>11. Example of Levels of Knowledge in the Instructional Intervention</td>
<td>150</td>
</tr>
<tr>
<td>12. Data Collection</td>
<td>161</td>
</tr>
<tr>
<td>13. Scoring Ranges for Classification as Low Use and High Use</td>
<td>168</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Zig Zag Pattern of Advanced Reader’s Intertextual Links</td>
<td>61</td>
</tr>
<tr>
<td>2. Frequency Counts for Listing Inconsistencies</td>
<td>171</td>
</tr>
<tr>
<td>3. Frequency Counts for the Decision Essay: Sourcing</td>
<td>173</td>
</tr>
<tr>
<td>4. Frequency Counts for the Decision Essay: Corroboration</td>
<td>175</td>
</tr>
<tr>
<td>5. Frequency Counts for the Decision Essay: Contextualization</td>
<td>176</td>
</tr>
<tr>
<td>6. Frequency Counts for the Decision Essay: Total</td>
<td>177</td>
</tr>
<tr>
<td>7. Frequency Counts for Justify Trustworthiness: Sourcing</td>
<td>179</td>
</tr>
<tr>
<td>8. Frequency Counts for Justify Trustworthiness: Corroboration</td>
<td>180</td>
</tr>
<tr>
<td>9. Frequency Counts for Justify Trustworthiness: Contextualization</td>
<td>182</td>
</tr>
<tr>
<td>10. Frequency Counts for Justify Trustworthiness: Total</td>
<td>183</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION TO THE STUDY

Statement of the Problem

The knowledge and abilities represented by earning a university degree allow individuals to keep pace in an increasingly complex, competitive world. Failure to earn a bachelor’s degree translates into failure to meet minimum qualifications for a significant number of job openings (Bureau of Labor Statistics, 2007), failure to position oneself for career advancement (National Commission on Writing, 2004), and loss of earning potential (Baum & Ma, 2007; Strong American Schools, 2008).

Academic Literacy Expectations in Higher Education

Attaining academic literacy is key to attaining a university education. Academic literacy is concerned with effective reading and writing at the university (Amos, 1999). Academic literacy includes the abilities to read analytically, negotiate multiple texts, apply information to novel situations, and demonstrate knowledge mastery with writing (Pugh, Pawan, & Antommarchi, 2000). Academic literacy is positively correlated with academic success at the university (Bosley, 2008; Holder, Jones, Robinson, & Krass, 1999; Pugh et al., 2000). In university contexts, students are expected to go beyond simple reading comprehension—to use reading to independently build knowledge, to apply what they learned from reading, often in the form of written work, and to solve novel problems.

The goal of higher education is to produce independent learners (Amos, 1999; Blau, 2006; Intersegmental Committee of the Academic Senates [ICAS], 2002; Paris, Lipson, & Wixson, 1983; Tetreault & Center, 2009). University students are expected to
use reading and writing to learn on their own. University expectations for academic literacy include building knowledge, reading analytically to identify problems, resolving problems that arise, and using writing to demonstrate advanced knowledge acquisition. These are tasks students must complete independently to be successful in their courses across their university careers.

However, a significant portion of students arrives at the university underprepared to meet these expectations for academic literacy. In the California State University (CSU) system, approximately 50% of incoming first-time freshmen are deemed underprepared for university-level coursework based on their scores from the English Placement Test (California State University [CSU], 2010). University faculty believe two-thirds of incoming freshmen are underprepared to meet university academic literacy expectations, particularly the expectations for analytic writing (ICAS, 2002).

The Role of Developmental-level Literacy Courses

Students deemed underprepared to meet the reading and writing expectations of the university are placed in developmental-level English courses. Postsecondary reading professionals prefer the term developmental to the label remedial because they recognize that learning to read is an ongoing, developmental process. Developmental-level courses are remedial courses whose curriculum has been designed to take individual learners’ developmental trajectory into account.

Students who are underprepared for the academic literacy demands of a university education are placed in developmental-level integrated reading and writing courses which recognize that learning to read and write are complementary, developmental processes (Goen & Gillotte-Tropp, 2003). The goal of these developmental-level academic literacy
courses is to help inexperienced readers/writers to become more like experienced readers/writers.

*How Inexperienced Freshmen Differ from Those Experienced with Academic Literacy*

Although reading research has focused primarily on the reading of single texts, a growing body of literature investigates how readers negotiate multiple texts, particularly historical documents (Britt & Aglinskas, 2002; Rouet, Favart, Britt, & Perfetti, 1997; Bråten, Strømsø, & Britt, 2009; Wineburg, 1991, 1998). Multiple text studies are more closely aligned with the complex academic literacy practices of the university, where students will be expected to build knowledge from multiple sources. Much research into readers working with multiple texts was conducted in the late 1990’s and early 2000’s. One line of inquiry pursued changes to students’ personal epistemology (e.g., Strømsø, Bråten, & Samuelstuen, 2008). A second branch explored the use of hypertext to make connections between texts (e.g., Strobel, Jonassen, & Ionas, 2008). The third branch, which holds the most promise for classroom instruction, focused on improving students’ skills (e.g., Wiley et al., 2009). These multiple text studies suggest that inexperienced, or novice, readers differ from experienced (expert and advanced) academic readers in their lack of awareness of the complexity of academic literacy. In particular, inexperienced readers fail to 1) detect inconsistencies across texts, and 2) employ conditional knowledge to strategically resolve inconsistencies.

Incoming freshman are rarely familiar with the complex intellectual work of the academy. First-year students are often inexperienced with academic literacy, not realizing that they should attend to both the content and the rhetorical features of a text (Scardamalia & Bereiter, 1991; Young & Leinhardt, 1998). Freshmen also do not yet
realize that questioning and grappling with difficulty are valued at the university, nor do they recognize that requirements for good reading are context-dependent, changing from text to text and task to task.

Inexperienced readers do not notice inconsistencies within a single text or across multiple texts and, therefore, do not utilize strategies for resolving inconsistencies. In a study of comprehension monitoring, when the last sentence of a paragraph explicitly contradicted the rest of the paragraph, inexperienced readers tended not to notice the textual inconsistency (Otero & Kintsch, 1992). Inexperienced readers also tend to gloss over contradictory evidence provided in different texts (Britt & Aglinskas, 2002; Wineburg, 1991). In contrast, experienced readers notice inconsistencies, ask specific questions, and formulate action plans to resolve these inconsistencies (Wineburg, 1991, 1998). This strategic behavior is a key difference between expert and novice readers. Experienced readers utilize conditional knowledge, knowing when and why to apply a strategy (Paris et al., 1983). Inexperienced readers tend not to notice comprehension issues and, therefore, may not realize they should mobilize a strategic approach (Garner, 1994).

Examinations of how readers negotiate texts containing contradictory information suggest that inexperienced readers, unlike experienced academic readers, either do not evaluate sources or only do so superficially and are, therefore, unlikely to use the evaluative heuristics that experienced readers rely on to resolve inconsistencies (Bråten et al., 2009; Britt & Aglinskas, 2002; Twait, 2005; Wiley et al., 2009; Wineburg, 1991). In a pair of landmark studies, Wineburg (1991, 1998) identified three evaluative heuristics that expert academic readers use to resolve inconsistencies: a sourcing heuristic, a
corroboration heuristic, and a contextualization heuristic. Sourcing is “the act of considering the source of the document when determining its evidentiary value” (Wineburg, 1998, p. 322); corroboration is a document comparison strategy for weighing evidence; and contextualization refers to attempts to reconstruct the spatial-temporal scene of events referred to in the document (Wineburg, 1991). Experienced academic readers use elements of these three heuristics flexibly to evaluate evidence and resolve inconsistencies (Rouet et al., 1997; Wasson, 2001; Wineburg, 1991, 1998). Although undergraduates possess declarative knowledge (they can tell you that they should evaluate sources) and procedural knowledge (they know how to evaluate them), they demonstrate a lack of conditional knowledge by not utilizing evaluative heuristics when necessary. Inexperienced, developmental-level students are not engaging in the very activities that could contribute to their academic success at the university.

Table 1

<table>
<thead>
<tr>
<th>Evaluative Heuristic</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sourcing</td>
<td>Evaluating the source of the information</td>
</tr>
<tr>
<td>Corroborating</td>
<td>Evaluating the information presented in light of other texts</td>
</tr>
<tr>
<td>Contextualizing</td>
<td>Evaluating information presented in light of the socio-temporal context</td>
</tr>
</tbody>
</table>

Previous Investigations into Academic Literacy Behaviors

Most multiple text studies have been descriptive, limited to describing the approaches readers use to decide a controversy presented in multiple texts (e.g., Bråten et al., 2009; Wineburg, 1998). These studies have increased understanding of the challenges students encounter with academic literacy. However, these descriptive studies have not provided insight into how to address these deficits.

There are few experimental or quasi-experimental studies addressing the efficacy of possible interventions and these have been limited to focusing on building procedural knowledge, not the conditional knowledge that developmental-level students need to be successful at the university. The researchers have provided an explicit inconsistency for participants to focus on, so students have not needed to recognize inconsistencies. The instructional interventions have focused on helping students identify information (procedural knowledge), not evaluate or apply information (conditional knowledge).

One intervention study in a high school history class found that instruction in evaluating sources improved students’ use of evaluative heuristics on the sourcing posttest (Britt & Aglinskas, 2002). However, the instructional intervention was a software program that provided a template for filling in source characteristics. A large portion of the posttest measure was the number of source characteristics identified, providing the experimental group \( n = 8 \) an advantage over the control group \( n = 7 \) on this particular measure. In effect, the researchers provided instruction to the experimental group, but not the control group, in a process that was then part of the scoring on the posttest. The intervention was somewhat successful in helping students identify the type
of information (procedural knowledge) used by experienced readers, but did not address students’ use of that information to resolve an inconsistency (conditional knowledge).

A recent study in high school history classes reported success with direct instruction of the three evaluative heuristics—sourcing, corroboration, and contextualization (Nokes, Dole, & Hacker, 2007). In this quasi-experimental study, intact classes were assigned for one of four instructional conditions: Textbook-Content, Textbook-Heuristics; Multiple Texts-Content, or Multiple Texts-Heuristics. Heuristic instruction resulted in significantly more instances of the sourcing heuristic in the essays, but both Multiple Texts conditions demonstrated more corroboration heuristics. Students rarely demonstrated use of the contextualization heuristic, which is commonly used by experienced readers. Exposure to multiple texts and not necessarily heuristic use accounted for improved scores on a test of history content. In addition, the change in instructional format from lecture before the intervention to small group activities and discussion during the intervention may have accounted for some of the increase in demonstrated content knowledge. Therefore, it is unclear which part of the intervention accounted for the noted improvement in evaluative heuristic use.

Most recently, a study of undergraduates’ sourcing behavior during an inquiry project in science reported improved sourcing skills after a sourcing evaluation intervention (Wiley et al., 2009). The experimental group received a 3-page handout discussing the SEEK strategy which was designed to help students remember four essential aspects of source evaluation: the Source of information, the nature of Evidence, how information fits with the Explanation, and how new information fits with prior Knowledge. The experimental group (the SEEK group) and the control group evaluated
documents from six Web sites related to the Atkins diet, but only the experimental group filled out a SEEK template for each source. In comparison to the control group, twice as many of the SEEK group demonstrated explicit use of the corroboration heuristic, but again use of the contextualization heuristic was rare. The essays written by the SEEK group were categorized as more conceptually integrated, while the control groups’ essays were deemed to have only local connections. This suggests that use of the evaluative heuristics can improve the depth of understanding students demonstrate in their written assessments, particularly essays. Although the intervention group demonstrated more instances of the sourcing and corroboration heuristics, the SEEK template approach is focused on building procedural knowledge—sloting information into a template—rather than building the conditional knowledge that students will need to learn independently from complex texts.

Although these three intervention studies offer support for the instructional value of direct instruction in using evaluative heuristics, all three limited themselves to improving declarative and procedural knowledge instead of attending to conditional knowledge (Britt & Aglinskas, 2002; Nokes et al., 2007; Wiley et al., 2009). At the university, students are expected to learn independently from texts (Pugh et al., 2000). To meet this expectation, students need to be able to identify inconsistencies across texts and to decide which evaluative heuristics will be most effective in resolving each particular inconsistency.

Therefore, this pre-experimental study investigated the instructional efficacy of an intervention designed to improve developmental-level students’ academic literacy skills through identification of inconsistencies across multiple texts and use of all three
evaluative heuristics (sourcing, corroborating, and contextualizing) to resolve inconsistencies.

Purpose of the Study

The purpose of the pre-experimental study was to examine the effectiveness of an instructional intervention to help developmental-level freshmen acquire the academic literacy skills that experienced academic readers demonstrate in order to promote independent learning and degree attainment. The study took place at a large, urban, public university over five weeks in two intact sections of a developmental-level integrated reading and writing course taught by one instructor. The developmental-level students in the two sections formed one group ($N = 31$).

The intervention focused on two aspects of advanced academic literacy: identifying inconsistencies across multiple texts and flexibly employing evaluative heuristics (sourcing, corroboration, and contextualization) in order to resolve inconsistencies. Students completed the Difficulty Paper assignment (Levinson, 2003; see also Salvatori, 1996)—a multiple step, integrated reading and writing strategy—in order to explicitly acknowledge confusion, ask questions, and specify gaps in knowledge which signal the need to mobilize conditional knowledge. Students defined and practiced using the three evaluative heuristics to evaluate information across multiple texts with the goal of reconciling inconsistencies and increasing understanding.

To investigate the effectiveness of this treatment, the study used a pre-experimental one group pretest-posttest design. The independent variable was the instructional intervention to improve developmental-level students’ academic literacy skills by identifying inconsistencies and using evaluative heuristics to resolve those
inconsistencies. There were three dependent variables in the study: 1) the number of inconsistencies identified, 2) the number of evaluative heuristics used in writing, and 3) the number of evaluative heuristics used in reading. Measures included Listing Inconsistencies, a researcher-designed measure of the number of inconsistencies identified; the Decision Essay to measure the number of evaluative heuristics used in writing, and the Justify Trustworthiness task, designed to measure the number of evaluative heuristics used in reading. Appendix A provides an overview of the data collection for this study.

Theoretical Rationale

Cognitive Flexibility Theory (CFT) provided the theoretic framework for this study (Spiro, Coulson, Feltovich, & Anderson, 2004). With its emphasis on problem solving, CFT reframes developmental-level students’ struggle with university-level reading as a problem to be solved through instruction, not an inherent failing on the part of the students. CFT illuminates the problem: inexperienced readers do not engage in the same behaviors that contribute to experienced readers’ success with academic literacy. CFT offers specific guidance for designing instruction that solves the problems that first-year, developmental-level university students encounter when reading multiple academic texts.

CFT (Spiro et al., 2004) explains the acquisition of advanced knowledge in complex domains, and, thus provides principles for effective interventions in university reading contexts. CFT frames problem solving as a natural part of advanced learning and defines the role that instruction can play in helping students acquire advanced knowledge (Spiro, 2001; Spiro et al., 2004; Spiro, Feltovich, & Coulson, 1996; Spiro & Jehng, 1990;
Spiro, Vispoel, Schmitz, Samarapungavan, & Boerger, 1987). Spiro et al. (2004) set three criteria for advanced learning: 1) learners must have been exposed to introductory learning, 2) learners must not yet be experts in the domain, and 3) learners are expected to apply knowledge. Learners are expected to have received introductory learning in a domain in order to form a foundation for future learning. This introductory learning may be characterized by factual learning, rote memorization, and multiple-choice assessments. The advanced learner has successfully acquired introductory level knowledge in the domain, but is not yet an expert in the domain. In order to advance their knowledge, the learner is expected to move beyond retelling facts, and apply conceptual understanding to diverse ill-structured problems.

Students enrolled in the developmental-level academic literacy courses satisfy the three conditions CFT sets for advanced learning: they have been exposed to introductory learning, they are not yet experts, and they are expected to apply knowledge (Spiro et al., 2004). These students have been exposed to introductory learning and come equipped with basic reading comprehension skills. They can extract literal meaning from textbooks, but they are not fully prepared for the rigors of academic literacy. They have limited experience reading multiple texts, learning complex concepts and engaging in analytic thinking. At the university, they are expected to go beyond memorizing facts in order to apply knowledge. Successful university students are expected to learn independently and developmental-level courses are designed to help students meet this expectation. Developmental-level literacy courses, like the one that served as the setting for this study, meet the criteria for applying CFT.
CFT assumes that problems are a natural part of learning. Part of learning in any content area is learning to solve problems within that particular content area. Learners must use previous experiences to interpret and comprehend new information. The goal of advanced learning is to be able to apply knowledge in novel problem-solving situations. CFT describes the way advanced readers and writers flexibly organize knowledge as it is acquired so that it is readily available for a variety of applications (Spiro & Jehng, 1990). CFT theorists originally conceptualized advanced learners as acquiring knowledge only in a specific domain, such as cardiology (Spiro et al., 2004). However, CFT describes the way any advanced learner, such as a university student (Spiro et al., 1996), engages in advanced learning in any academic domain (Spiro, 2001), such as the ones they encounter while completing their general education courses. In the current study, CFT predicts that the intervention will help developmental-level freshmen make use of knowledge and strategies as do more experienced academics.

CFT focuses on the acquisition of advanced knowledge in ill-structured domains (Spiro et al., 2004). An ill-structured domain is characterized by problems that do not have a single, explicit answer. Instead, the problem solver must draw on prior case knowledge and mobilize several resources to find a satisfactory solution. Conditional knowledge is a valuable asset for ill-structured problem solving because each problem has unique characteristics. Therefore, the problem solver needs to assess the problem and select a strategy that helps address the specific features of each new problem.

CFT represents prior knowledge (background knowledge) as a series of interrelated cases, a flexible system of organizing knowledge that can be rapidly mobilized in novel situations (Spiro et al., 2004). An individual’s prior knowledge is
represented as schema, an organized structure of all the experiences and information he or she possesses. Unlike traditional schema theory, which represents schema as a monolithic structure (Rumelhart, 1980), CFT conceptualizes prior knowledge in the form of cases (Spiro et al., 1987; Spiro et al., 2004; Spiro & Jehng, 1990; Spiro et al., 1996). Each experiential event is a case (Spiro et al., 2004). Each case is like a packet of information tied to a specific event within a specific context (Spiro et al., 1987).

Case-based structure allows for flexible organization of cases in memory, since each case can be mobilized independently. These cases are, in turn, composed of minicases, which are elements of the case (Spiro et al., 1987; Spiro et al., 2004). For example, a case might cover reading a history chapter for a university-level history course. The minicases would include specific elements of that experience: reading at a study carrel in a quiet library, taking notes, researching a primary source to answer a lingering question. All these aspects are connected, by being part of the original reading event, but can be decomposed into independent packets of information. One minicase would inform the reader’s sense of preferred study environment: a table at the quiet library.

Combined with other like minicases from other reading events, these minicases allow for multiple perspectives on each variable. For example, the reader might deduce that she prefers reading in quiet locations or at tables or near other potentially useful resources. Because these minicases allow case information to be deconstructed and combined with any other minicase, they allow for flexible adaptation to new situations, offering multiple perspectives on any new experience (Spiro et al., 2004; Spiro & Jehng, 1990; Spiro et al., 1996).
Although CFT acknowledges that students’ previous educational experiences may limit their ability to work productively with complexity, CFT suggests ways to overcome these limitations (Spiro et al., 2004; Spiro et al., 1996). Previous literacy experience and instruction has led students to believe that knowledge is simple and explicitly stated in a single source (Spiro et al., 2004). When reading multiple texts, a reader is more likely to encounter inconsistencies across texts. In their previous educational experience, difficulty was to be avoided as it did not help with memorization or answering multiple choice-type items on assessments. In the new context of the university where students are expected to appreciate complexity and navigate multiple perspectives (Spiro, 2001), recognizing difficulty in the form of textual inconsistency is considered productive. This represents a shift in their understanding of academic literacy for many first-year students.

CFT suggests that instruction matters in acquiring advanced knowledge and defines its role. In CFT, instruction is viewed as surrogate aptitude. Students could learn from various experiences in their lives, but we “do not want to have to wait that long for experience to accrue” (Spiro et al., 1987, p.191). The purpose of developmental-level coursework is to help first-year readers acquire the same effective academic literacy behaviors as experienced academic readers and writers, which include building conditional knowledge, increasing awareness of multiple perspectives, appreciating the complexity of tasks, and utilizing evaluative heuristics, in short, to acquire advanced academic literacy.

In order to do this, CFT envisions a content domain as a landscape in which the same features assume different patterns of significance when placed in different contexts (Spiro et al., 1987). Terminology provides a simplified example of the effect context has
on meaning. In an undergraduate English course, a thesis is a sentence that explicitly states the argument of an essay, while in graduate courses a thesis is a sixty-page culminating project. Likewise, an historian might see being a participant in an event as evidence of credibility because of the value of eyewitness testimony, but would also consider the increased likelihood of bias. Cognitive flexibility based instruction encourages seeing the same information from different perspectives: the thesis statement, the thesis project, and the relationship between them—a focused exploration of a topic. Instruction should allow for exploring a domain, reexamining cases from various perspectives, and connecting knowledge into an interconnected whole, albeit complex, landscape. This instructional approach is referred to as “criss-crossing” the landscape (Spiro & Jehng, 1990, p. 169; Spiro et al., 2004). Criss-crossing encourages students to analyze cases and explore ways to recombine minicases. Re-examining and recombining case elements will lead to flexible mobilization of knowledge and facilitate its transfer to new situations.

This study took place in two sections of a developmental-level integrated reading and writing course at a large, urban public university. The purpose of the course is to help first-year university freshman develop academic literacy strategies that they can flexibly deploy, so they may use reading across disciplines to acquire knowledge. First-year, developmental-level freshmen tend to have a narrow schema for literacy and exhibit unidimensional strategy deployment. In this course, they are asked to read multiple texts, yet they struggle to evaluate texts and recognize inconsistencies. The intervention included instruction targeted towards increasing students’ awareness of the expectations for academic literacy through recognition of inconsistencies within and across texts.
Because students also struggle to reconcile inconsistencies when they are noted, the intervention included instruction in how to flexibly use case knowledge (Spiro & Jehng, 1990; Spiro, 2001) to deploy prior knowledge and strategies in the form of evaluative heuristics to resolve inconsistencies across cases (Wineburg, 1991).

CFT predicted that the intervention would improve first-year, developmental-level students’ ability to recognize textual inconsistencies and draw on multiple strategies to reconcile the inconsistencies, including using the same evaluative heuristics experienced readers use—sourcing, corroborating, and contextualizing. If students’ previous educational experiences have convinced them that knowledge is unidimensional, explicit instruction would help them to see multiple perspectives. If students are not recognizing textual inconsistencies because they avoid difficulty, then reframing difficulty as a positive attribute of texts and explicitly structuring assignments to help students discover inconsistencies should address the problem. If students are unable to reconcile inconsistencies because of limited strategies, then instruction in evaluative heuristics (sourcing, corroboration, and contextualization) would enable them to employ the same strategies experienced academic readers use.

Background and Need

In California, more than 60% of the 40,000 freshmen admitted to the CSU require remediation (National Center for Public Policy and Higher Education, 2008; Strong American Schools, 2008). Because of the large number of high school graduates deemed underprepared for university-level coursework, many universities offer developmental-level literacy courses to help students gain effective academic literacy skills (National Center for Educational Statistics [NCES], 2001, 2003).
Prior research can help shed light on how to address the poor academic literacy skills of developmental-level university freshmen. Prior research suggests that experts in well-structured domains such as physics utilize problem-solving templates (Chi, Feltovich, & Glaser, 1981; Larkin, McDermott, Simon, & Simon, 1980). In a study centered on solving problems in geometry—a well-structured domain—problem-solving abilities appeared more important than domain knowledge (Schoenfeld, 1985). Problem-solving and meaning-making strategies are even more important in ill-structured domains, such as history and literacy (Scardamalia & Bereiter, 1991; Wineburg, 1991).

Wineburg, whose primary investigations are in the domain of history, acknowledged the similarities between historical knowing and academic literacy (1998).

*Wineburg Makes Academic Literacy Explicit*

In a pair of landmark studies, Wineburg made explicit the academic reading behaviors of experienced readers (1991, 1998). Wineburg’s work with experienced historians and novice high school students illuminated two aspects of academic literacy: recognizing inconsistencies and using evaluative heuristics to resolve those inconsistencies. The research base for those two ideas will be explored in greater detail in the Literature Review. However, in this section, Wineburg’s two studies will be discussed briefly along with studies that support his findings.

Recognizing that history is an ill-structured domain, Wineburg (1991) designed a study to examine how experts and novices constructed historical understanding from contradictory accounts of a historical event. Wineburg (1991) used think alouds to examine how experienced historians \( n = 8 \) and high school students \( n = 8 \) used sources to resolve a historical controversy across eight texts. After completing a measure of prior
knowledge, participants read eight written documents about the Battle of Lexington and examined three paintings of the event while engaged in a think aloud procedure.

In order to identify finer distinctions between experienced readers, Wineburg (1998) focused on the contextualization heuristic, whose use seemed unique to experienced academic readers. Wineburg examined how an expert reader (a content area specialist in Abraham Lincoln and the Civil War period) and an advanced reader (a discipline expert in the field of American history) differed in their approach to multiple contradictory texts. Spiro et al. (2004) define an advanced reader as one who has been exposed to introductory learning and, although not yet an expert, is expected to apply information. This describes most university students during their undergraduate years.

Both the advanced and expert reader engaged in think alouds while they read seven documents related to President Lincoln’s personal stance on race. Concurrent and retrospective protocols were coded for evidence of the contextualization heuristic, intertextual links, and specification of ignorance (moments of difficulty). The findings of this study (1998) along with Wineburg’s earlier study (1991) will be discussed below.

Experienced readers note inconsistencies across texts and between textual information and their understanding. An inconsistency is a contradiction between two or more interpretations or accounts within or across texts. To pinpoint these moments of confusion is a mark of expert behavior. While reading about President Lincoln, both historians engaged in specification of ignorance—identifying gaps in their understanding and detailing the knowledge they would need to make a judgment (Wineburg, 1998). The expert reader specified ignorance 7 times, while the advanced reader specified ignorance 21 times. In one instance, the advanced reader struggled to understand the
meaning behind a particular phrase in a one document: “capable of thinking like a white man” (p. 335). He was explicit in identifying his confusion, referring to it as “a baffling statement” (p. 335).

After identifying gaps in their knowledge, experienced readers create action plans to address their questions. For example, the expert reader encountered uncertainty over whether Lincoln’s use of “white men” should be read as “free men” (Wineburg, 1998, p. 335), so he suggested a possible plan to resolve that uncertainty, a search of the abolitionist literature for similar references. In the earlier study, an expert reader had puzzled over the anxiety level of the colonists as they waited on Lexington Green—a concept not addressed within the document set—and decided that he might be able to obtain personal letters in order to gain insight into the emotional state of the Minutemen (1991). When these experienced readers encountered inconsistencies, they created action plans to resolve the problem.

In contrast, novice readers rarely demonstrated awareness of uncertainty (Wineburg, 1991). Students provided more descriptive comments about the paintings, but infrequently qualified their selection. Expert readers use qualification an average of eight times each, while inexperienced readers used qualification an average of one time each. The novice readers’ explanations for picture selection were characterized by certainty. For example, one student chose the 1859 depiction of the Battle of Lexington because of the inclusion of a hill—a feature not mentioned in any of the written documents. Inexperienced readers, like the developmental-level students in the current study should benefit from explicit instruction in the problem-solving approaches that experienced academic readers use, including recognizing inconsistencies.
Wineburg’s findings indicate that the experienced academic readers use reading strategies, including the evaluative heuristics—sourcing, corroboration, and contextualization—in order to make sense of contradictory information. Experienced academic readers used the source information to preview the document (Wineburg, 1991). All eight historians attended to source information before reading the Battle of Lexington documents, using the *sourcing* heuristic 98% of the time. One historian noted that knowing the source permitted her to predict what she might find. Previewing the source information activated genre and author knowledge. When one historian previewed the excerpt from a high school textbook, a rich textbook schema was activated. Experienced academic readers used their prior knowledge to better understand the current texts.

Experienced academic readers used the *corroboration* heuristic—a document comparison strategy—significantly more often than inexperienced readers (1991). Experienced readers looked back to previous documents an average of six times apiece. The think aloud protocols provide more precise evidence of triangulating information. In particular, historians referred back to Document 2 in an attempt to evaluate the claims about Minutemen troop size in Document 4, creating scenarios, posing questions, and noting the lack of precise detail. Experienced academic readers compared and contrasted information across texts to better understand the historical event.

Experienced readers used the *contextualization* heuristic, attempting to reconstruct the spatial-temporal scene of events (1991). Participants needed to decide on the time of day and the order of events leading up to the assembling of troops on Lexington Green. For instance, one historian found support for the claim that the usually
disciplined British troops violated direct orders to hold their position, by noting that the British were tense, sleep deprived, and wearing wet, uncomfortable clothing.

Although experienced readers frequently use the evaluative heuristics to understand events and evaluate the credibility of evidence (1998), students rarely engaged with the heuristics, most notably corroboration and contextualization (1991). These inexperienced readers failed to use strategies to improve their understanding of the event or their evaluation of the information. Since the goal of developmental-level courses is to provide students with the same strategies that experienced readers would use, this study included explicit instruction in the existence and the use of the evaluative heuristics in order to help developmental-level students achieve academic literacy.

Therefore the intervention portion of the study was designed to help developmental-level freshmen increase their skills in academic literacy: to help students recognize inconsistencies and use the evaluative heuristics flexibly in order to resolve difficulties as both readers and writers.

Research Need for the Study

Because inexperienced readers tend not to use the evaluative heuristics that experienced readers use (Bråten et al., 2009; Britt & Aglinskas, 2002; Wineburg, 1991), researchers have designed interventions to help students learn to use these heuristics (Britt & Aglinskas, 2002; Nokes et al., 2007). However, research on instructional interventions have not been successful in helping students to use the contextualization heuristic, suggesting that students are not using the most helpful aspects of the heuristics in a given situation (Britt & Aglinskas, 2002; Nokes et al., 2007; Wiley et al., 2009). In order to be as successful as experienced academics, developmental-level freshmen need
to have access to all the strategic tools that the three evaluative heuristics represent. There is a need for research that examines interventions to help students learn to use all three heuristics, so they can flexibly apply the most helpful aspects of each heuristic to resolve any difficulty they encounter.

In addition, intervention research has overlooked the recognition of inconsistencies which is the key to successful reading at the university level, as noticing an inconsistency is the trigger for deploying conditional knowledge (Britt & Aglinskas, 2002; Nokes et al., 2007; Wiley et al., 2009). Once an inconsistency is identified, the reader can then specify what information he or she needs and devise a plan for resolving the inconsistency. At this point, the evaluative heuristics would become useful steps in a plan to resolve the difficulty. There is a need for research that examines interventions to help students learn to identify inconsistencies across multiple texts.

Therefore, the intervention portion of this study made the expectations for academic literacy explicit through instruction 1) in recognizing inconsistencies when reading multiple texts; and 2) in using the three evaluative heuristics to resolve those inconsistencies. This study examined the effectiveness of this intervention in promoting academic literacy, using three measures: Listing Inconsistencies, to measure the number of inconsistencies noted within or across texts; the Decision Essay, a short argument essay used to measure the number and type of evaluative heuristics used in writing; and Justify Trustworthiness, a measure of the number and types of evaluative heuristics used in reading. All three measures are part of the Multiple Text Tasks that participants completed after reading the Battle of Lexington document set (Appendix B).
Listing Inconsistencies

Listing Inconsistencies is a researcher-designed measure that draws on work by Rouet, Britt, Mason, and Perfetti (1996) and Salvatori (1996) to measure the number of inconsistencies noted within and across texts (Appendix B). In a study of undergraduates’ ability to reason with multiple documents, participants were asked to list any additional documents they would have liked to have access to in order to write a decision essay about a historical event (Rouet et al.). This represents one type of expert behavior, noticing what information is not available for making an informed decision.

Salvatori pioneered the difficulty paper assignment in which students were asked to catalog and discuss any difficulties they had while reading a single text. The Listing Inconsistencies assessment asks students to create a simple list, similar to the list used in Rouet et al., but it covers any type of inconsistency that participants might have noticed within or across texts in the document set, as Salvatori advocates. In this study, students read the Battle of Lexington document set and completed the Listing Inconsistencies assessment as part of the Multiple Text Tasks. The researcher-created Scoring Guide for Listing Inconsistencies (Appendix C) was used to train scorers and to score the assessment.

The Decision Essay

The Decision Essay (Appendix B) is a 1-2 page essay written to decide a controversy which is posed as a question (e.g., Who fired the first shot at the Battle of Lexington?). Britt and Aglinskas (2002), Rouet et al. (1996, 1997), and Nokes et al., (2007) utilized similar decision essay assignments to decide historical controversies. Wiley et al. (2009) assigned a decision essay to decide a scientific controversy (i.e., the
cause of volcanic eruptions). The controversy is a researcher-selected, explicit inconsistency that participants need to make a decision about. The decision essay assignment has two parts. First, participants read a document set that presents multiple perspectives on a controversy. Secondly, participants write a short essay of approximately 200 words to explain their decision on the controversy. In this study, the Decision Essay was coded for evidence of evaluative heuristics, using The Evaluative Heuristics Rubric and the Evaluative Heuristics Scoring Guide adapted from Nokes et al. (2007), in order to measure the number and type of evaluative heuristics used in writing.

*The Evaluative Heuristics Rubric and Study Guide*

The Evaluative Heuristics Rubric (Appendix D) has been adapted from Nokes et al.’s (2007) Heuristic Rubric by the researcher for use with non-history topics. Nokes et al.’s Heuristic Rubric was based on the coding scheme used by Britt and Aglinskas (2002) and Wineburg (1991, 1998). Nokes et al. modified it for use with high school students, creating a coding sheet as well as detailed coding instructions that operationally defined each heuristic (sourcing, corroboration, and contextualizing), listed descriptors of what should and should not be considered an instance of heuristic use, and gave examples of heuristic use. Britt and Aglinskas, Wineburg, and Nokes et al. used the coding scheme with a historical controversy. Since the goal of the current study was to help students see the applicability of the evaluative heuristics to problem solving in a variety of academic domains, the coding scheme was modified slightly so that it would work with history and non-history topics.
Justify Trustworthiness

The Justify Trustworthiness task, based on a ranking measure used in Wiley et al. (2009), Wasson (1991), and Rouet et al. (1996, 1997), was used to measure evaluative heuristic use in reading (Appendix B). In a study of evaluating information in an inquiry-project in high school, Wiley et al. (2009) asked students to rank the trustworthiness of the documents they read as part of the instructional intervention. The exercise was returned to them with a content area expert’s ranking. Students then answered questions designed to help them better understand the expert’s rankings. At the university level less emphasis is placed on reproducing expert knowledge. Students are expected to enact personal judgments based on their own knowledge. Therefore, expert rankings were not used in this study.

Wasson (1991) asked participants to rank the 13 documents—written, pictorial, and video—that they read or viewed about the Battle of the Plains of Abraham. Wasson compared the rankings of inexperienced readers (high school students) with experienced readers (university historians). Although not explicitly prompted to do so, many participants explained their rankings as part of the think aloud procedure. However, these comments were not provided by the researcher.

Rouet et al.’s use of ranking and justifying documents is closest to the measure for this study. Rouet et al. (1996, 1997) asked students to rank the documents they had read to decide on each of four historical events (e.g., Was U.S. intervention in Panama justified?) and justify their reasoning in a sentence or two. The participants were given the source information for each of the documents, as it had appeared on the document. Participant justifications were classified as content, author, document, and opinion, which
correspond roughly to the sourcing and corroboration heuristics. In this study, as in Rouet et al. (1996, 1997) the source information for each document was provided and students were asked to rank each document and provide one to two sentences to justify their ranking. The rankings were not scored, but the justifications were coded for evaluative heuristics, using The Evaluative Heuristics Rubric and Evaluative Heuristics Scoring Guide adapted from Nokes et al. (2007). In this study, the Justify Trustworthiness task was used to measure evaluative heuristic use by readers.

The three measures—Listing Inconsistencies, the Decision Essay, and the Justify Trustworthiness task—were selected to provide evidence of students’ ability to recognize inconsistencies and resolve them using evaluative heuristics. Therefore, these three measures were used as part of the Multiple Text Tasks to evaluate the effectiveness of the explicit instructional intervention to teach academic literacy skills to developmental-level freshmen in the current study.

Components of the Instructional Intervention

The instructional intervention used an anticipation guide for which students are given a series of statements about the topic of instruction in order to activate prior knowledge in preparation for acquiring new information (Readence, Bean, & Baldwin, 2004). A PowerPoint lecture (Appendix E) and class discussion based on the concepts presented in the Anticipation Guide for Academic Literacy Expectations (Appendix F) were used to explicitly provide additional information about expectations for academic literacy.

The Difficulty Paper, an elaborated reading strategy, was used to make the problem-solving process of experienced readers and writers explicit to developmental-
level readers/writers. As part of the Difficulty Paper assignment (Appendix G) students were asked to 1) identify any difficulties (inconsistencies) they noticed; 2) select one inconsistency and create a plan for resolving that inconsistency; 3) deploy strategies; and 4) reflect on learning outcome and choice of strategy. The goal of this portion of the intervention was to provide students with practice identifying inconsistencies, to help students to appreciate the value of difficulties (Salvatori, 1988; Miller, 1994) and to provide a strategy for resolving difficulties that developmental-level readers/writers are likely to encounter throughout their university education (Fisher, 2006; Levinson, 2003).

Direct instruction in the three evaluative heuristics—in the form of the Introduction to Evaluative Heuristics (Appendix H) and in-class activities—was provided to help students understand the heuristics and learn to use them flexibly as expert academics do. Expert utilization of evaluative heuristics has been well documented (e.g., Jacobson, 2001; Rouet et al., 1996; Smith et al., 1991; Wineburg, 1991, 1998; Wasson, 1991). Experts use the heuristics flexibly to solve novel problems. Rouet et al. (1997) demonstrated that advanced learners transfer strategic knowledge from prior experiences when faced with novel problems in different disciplines which mirrors the desired outcome for undergraduates who must achieve success in courses from a variety of disciplines. Other researchers have studied instruction in the evaluative heuristics in order to help inexperienced readers and writers become more like expert academic readers and writers (Britt & Aglinskas, 2002, Nokes et al., 2007, Wiley et al., 2009). Findings from Britt and Aglinskas suggest they were successful in helping students gain procedural knowledge by identifying the type of information that expert readers would evaluate. Findings from Wiley et al. suggest they were successful in helping students use
two evaluative heuristics (sourcing and corroboration) and that learning about the evaluative heuristics helped students to increase the depth of their understanding as demonstrated in an essay. Findings from Nokes et al. suggest that direct instruction was a successful method for teaching inexperienced readers about the evaluative heuristics.

However, none of the intervention studies presented findings to suggest that participants had reached the conditional level of strategy use. Developmental-level freshmen will need a firm grasp on how to use each of the three heuristics and when to use each to resolve a particular difficulty, if they are to attain advanced academic literacy status. The research base supports the importance of teaching the evaluative heuristics to inexperienced readers and suggests that direct instruction is an effective method of providing instruction. Therefore, the current study utilized direct instruction in all three evaluative heuristics, especially contextualization, in order to help developmental-level students gain experience with academic literacy.

Significance

This intervention study is significant for four reasons. First, this study provides additional data on an intervention to help inexperienced readers become more successful academic readers. Although descriptive research has identified the skills and knowledge that inexperienced readers lack (e.g., Bråten et al., 2009; Wineburg, 1991), few intervention studies have been conducted to find ways to meet these students’ instructional needs. This study contributes to the literature on possible instructional interventions.

Secondly, this study attempted a more comprehensive instructional intervention to help students become independent scholars. Previous intervention studies have been
limited in scope, primarily focused on acquiring procedural knowledge of evaluative heuristics (Britt & Aglinskas, 2002; Nokes et al., 2007; Wiley et al., 2009). Although this procedural knowledge of potential strategies is helpful, for students to succeed at the university-level they must possess conditional knowledge of when and how to use the strategies. Recognizing inconsistencies is the trigger for using evaluative heuristics. Students will need to master the contextualization heuristic—one of the evaluative heuristics—which several interventions have failed to demonstrate (Britt & Aglinskas, 2002; Nokes et al., 2007; Wiley et al., 2009) if they are to flexibly draw on the strategies. This study examined the effectiveness of an intervention focused on helping students to acquire conditional knowledge related to recognizing inconsistencies and using all three evaluative heuristics.

Thirdly, this study examined the benefits of an intervention to help developmental-level university freshmen. Previous studies have drawn samples from high school students for whom building procedural knowledge may be more appropriate (Britt & Aglinskas, 2002; Nokes et al., 2007; Wineburg, 1991) and undergraduates who could be upper-classmen already apprenticed to the advanced academic literacy expectations of the university (Bråten et al., 2009; Wiley et al., 2009). Thus, upperclassmen may have already acquired some strategies related to the evaluative heuristics. First-year students identified as developmental remain an understudied, yet needy, population.

Fourthly, this study is significant because it explored the value of Wineburg’s (1991, 1998) evaluative heuristics as academic literacy strategies to be used across domains. Most of the previous multiple text research has focused on the domains of
history (e.g., Nokes et al., 2007; Wineburg, 1998) and science (Wiley et al., 2009), so there is little empirical research addressing whether the benefits of evaluative heuristic usage would generalize to other domains.

Research Questions

The pre-experimental study investigated the effectiveness of an explicit academic literacy intervention on one group of developmental-level freshmen. This group \((N = 31)\) was composed of two intact sections of a developmental-level integrated reading and writing course. There are three dependent variables: 1) the number of identified inconsistencies; 2) the number of evaluative heuristics used in writing; and 3) the number of evaluative heuristics used in reading.

This study attempted to answer the following research questions:

1. What is the effect of an explicit academic literacy instructional unit on the number of inconsistencies identified by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Listing Inconsistencies measure?

2. What is the effect of an explicit academic literacy instructional unit on the number of evaluative heuristics used in writing by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Decision Essay measure?

3. What is the effect of an explicit academic literacy instructional unit on the number of evaluative heuristics used in reading by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Justify Trustworthiness measure?
Definition of Terms

*Academic Literacy* –

Academic literacy refers to students’ “ability to read and write effectively within the university context” (Amos, 1999, p. 178). Academic literacy includes using reading to acquire content knowledge and knowledge of rhetorical processes (Paxton, 1999; Paris & Jacobs, 1984; Scardamalia & Bereiter, 1991; Young & Leinhardt, 1998) as well as demonstrating the ability to apply that knowledge to solve problems or make sense of new information (Pugh et al., 2000).

*Advanced Learning* –

Advanced learning is acquisition of knowledge “which is intertwined and dependent, has significant context-dependent variations, and requires the ability to respond flexibly to ill-structured application structures” (Spiro et al., 2004, p. 641). Advanced learning would occur after the learner has been exposed to the subject area, but is not yet an expert. However, advanced learning includes an expectation to apply knowledge (Spiro et al., 2004).

*Advanced Reader* –

An advanced reader is an experienced reader who can decode and comprehend text (who has mastered introductory level reading), but has not reached the level of expert reader. An advanced reader is expected to apply knowledge acquired from reading to solve novel problems (e.g., decision essay). At the university, this often involves writing essays.
Anticipation Guide –

An anticipation guide is an instructional activity. Students are given a series of statements about the content they are about to read. Students are asked to respond to each statement, usually by noting True/False or Agree/Disagree in order to activate their prior knowledge about the topic (Readence et al., 2004).

Cases –

A case is a knowledge application situation (Spiro et al., 2004). Cases are “examples, occurrences, and events (occasions of use of conceptual knowledge)” (Spiro, et al.).

Case-based Structure –

This term refers to the structure of this knowledge in the form of cases. The structure is composed of cases decomposed into minicases and interrelationships between cases and minicases. This term refers to just the structure without any content and may be represented in the mind or in the real world.

Cognitive Flexibility –

Cognitive flexibility is “the human ability to adapt the cognitive processing strategies to new and unexpected conditions in the environment” (Cañas, Antolí, Fajardo, & Salmerón, 2005, p. 95).

Cognitive Flexibility Theory (CFT) –

CFT is a transdisciplinary theory that posits flexible problem solving is a means to advanced knowledge acquisition (Spiro et al., 2004).
**Conditional Knowledge** –

Conditional knowledge is knowledge about *when* to use a skill or a strategy and *why* it is effective (Paris et al., 1983).

**Contextualization** –

*Context* comes from the Latin root is *contextere* which means to weave together, “to connect strings in a pattern” (Wineburg, 1998, p. 339). In this study, contextualization is operationally defined as an effort to reconstruct the spatial-temporal scene—geographic, political, historical, and cultural context of an event—and to comprehend documents within that context (Nokes et al., 2007; Wineburg, 1991).

**Corroboration** –

Corroboration is operationally defined as a document comparison strategy for weighing evidence, which includes making connections between information found in different texts, and noting both contradictions and similarities (Britt & Aglinskas, 2002; Nokes et al., 2007; Wineburg, 1991).

**Criss-cross** –

Criss-crossing in both an instructional strategy for helping students build conceptual understanding and a description of the reading process (Spiro et al., 1987; Spiro et al., 2004). Instructionally, criss-crossing means to look at a feature of the knowledge landscape from various perspectives. Behaviorally, a reader is criss-crossing when he or she looks across sections of a single text or across texts to build conceptual understanding of an event.
Developmental-level –

Developmental-level courses are remedial courses whose curriculum has been designed to take individual learners’ developmental trajectory into account. Students who are underprepared for the academic literacy demands of university education are placed in developmental-level integrated reading and writing courses which recognize that learning to read and write are complementary, developmental processes (Goen & Gillotte-Tropp, 2003).

Difficulty Paper –

The original difficulty paper was an assignment that asked readers to catalog and discuss difficulties they encountered while reading a single text as a way to gain further insight into the text (Salvatori, 1996). Levinson’s Difficulty Paper is an elaborated reading strategy scaffolded into 4 parts that explicitly take the reader/writer through the process that experienced readers use to resolve difficulties (2003). In this study, the Difficulty Paper assignment is a modified version of Levinson’s Difficulty Paper for use with two texts.

Evaluative Heuristics –

Wineburg (1991) identified three heuristics (sourcing, corroboration, and contextualization) that experts use in navigating multiple texts, which he referred to as sourcing heuristics. However, one of the heuristics is commonly referred to as the “source heuristic.” To avoid confusion about terminology, the three heuristics will be referred to as evaluative heuristics in this study.
**Experienced Reader –**

For the purpose of this study, the term “experienced reader” will refer to both expert academic readers, like university scholars, and advanced readers, like graduate students.

**Expert Reader –**

It is only possible to be an expert reader in a specific domain. Expert readers have been reading for about 10 years in their specific domain (see Ericsson, Krampe, & Tesch-Romer, 1993). They have acquired the domain-specific content knowledge and the skills to be considered an expert in that domain. University professor are an example of expert readers.

**Heuristic –**

Heuristics are specific problem-solving strategies or “sense-making activities, [which] help their user resolve contradictions, see patterns, and make distinctions among different types of evidence” (Wineburg, 1991, p. 77).

**Ill-structured –**

Ill-structured means that many context-dependent concepts interact in a typical case of knowledge application (Spiro et al., 2004). Ill-structured problems do not have a single, explicit answer.

**Inconsistency –**

An inconsistency is a contradiction between two or more interpretations or accounts, within or across texts. An inconsistency arises when two or more facts or claims are not compatible with each other. In this study, the term *inconsistency* is synonymous with contradiction (Wineburg, 1991), controversy (Rouet et al.,
conflicting perspectives (Nokes et al, 2007), and difficulty (Salvatori, 1996).

**Integrated Reading and Writing** –
Studies of literate expertise indicate that reading and writing are reciprocal processes (see Goen & Gillotte-Tropp, 2003). As an individual becomes a better reader, acquiring content and rhetorical knowledge, he/she becomes a better writer because he/she is able to apply that new knowledge (Scardamalia & Bereiter, 1991).

**Introductory Learning** –
Introductory learning is the beginning stage of learning. Introductory learning is a required precursor to advanced learning and, eventually, expertise. In reading, introductory learning would include instruction and practice in decoding texts. In university contexts, introductory learning is characterized by lecture and survey classes, which introduce college students to various disciplines.

**Minicase** –
A minicase is a segment of a case (Spiro et al., 2004). Minicases can be recombined with other case elements (minicases) to form flexible knowledge structures.

**Sourcing** –
Sourcing is “the act of considering the source of the document when determining its evidentiary value” (Wineburg, 1998, p. 322). To source, means to weigh a text’s information in light of its source (Rouet et al., 1996, p. 478).
Specification of Ignorance –

Specification of ignorance occurs when the reader “explicitly acknowledge[s] confusion, express[es] puzzlement or wonder, ask[s] questions, or specify[s] gaps in knowledge” (Wineburg, 1998, p. 325). Specifying ignorance means that a reader has explicitly identified an inconsistency.
CHAPTER II
REVIEW OF LITERATURE

University students are expected to learn independently and developmental-level literacy courses are designed to help students meet this expectation. The purpose of developmental-level literacy courses is to help underprepared students—the target population for this study—gain advanced academic literacy (Goen & Gillotte-Tropp, 2003). The literature review is divided into two sections that examine the concept of academic literacy and its implications for students enrolled in developmental-level literacy courses—Developing Academic Literacy: Recognizing Inconsistencies and Developing Academic Literacy: Using Evaluative Heuristics.

Developing Academic Literacy: Recognizing Inconsistencies

A majority of first-year students are unaware of the academic literacy expectations at the university level, including the expectation that they become independent learners. University freshmen are unlikely to understand that academic literacy is primarily concerned with problem solving. High school students may be accustomed to an “assign and tell” format, in which the reading is assigned, but the teacher then tells students that same information in a lecture (Vacca & Vacca, 2005). However, at the university students are expected to read independently to build background knowledge to contextualize lectures on selected topics. Reading is expected to go beyond comprehension to include evaluation and depth of processing (Grabe & Stoller, 2002; Olson & Land, 2007; Scarcella, 2003; Sommers & Saltz, 2004). Students are expected to build knowledge of content and rhetorical processes, in order to acquire academic literate expertise (Paxton, 1999; Paris & Jacobs, 1984; Scardamalia & Bereiter,
1991; Young & Leinhardt, 1998). In order to do this students must read analytically—explicitly identifying multiple perspectives and inconsistencies within and across texts, as well as formulating their own questions. Then, they must take what they learned from the text (content knowledge) and the problem-solving process (metacognitive knowledge) and apply it to the next literacy task.

Students are Underprepared for University-level Academic Literacy Demands

First-year freshmen are rarely prepared for university expectations regarding academic literacy. Although 63% of high school students will go to college, only 43% take college preparation courses of study while in high school (Brenneman & Haarlow, 1998). Once first-year students arrive at the university they often find that it is more challenging than their high school work for a number of reasons.

At the university, their assigned reading will be composed primarily of complex, expository texts (Wiley, Griffin, & Thiede, 2005). Across four years, high school textbooks increase in difficulty by approximately 100 Lexiles, a measure that uses a common scale to assess reader ability and text readability, whereas the jump between high school and college texts is 260 Lexiles (Williamson, 2008). The average high school student will find him or herself reading many college texts at frustration level, with less than 50% comprehension (Williamson).

Furthermore, first-year university students are often unaccustomed to working with multiple texts, because throughout much of secondary education, the course textbook is regarded as the authority. High school students tend to have a textbook mentality (Wineburg, 1991), a belief that the textbook is factual and is above criticism (Luke, de Castell, & Luke, 1989; Paxton, 2002). Common evaluation procedures like
multiple-choice tests can reinforce the perceived authority of the textbook; there is only one correct answer, the one found in the course textbook. While textbooks present a tidy factual presentation of knowledge, multiple texts present students with ill-structured, messy problems. Multiple texts present multiple perspectives and inconsistent information, requiring the reader to actively identify inconsistencies, a requirement that may be foreign to incoming freshmen.

An inconsistency is a contradiction between two or more interpretations or accounts within or across texts. An inconsistency arises when two or more facts or claims are not compatible with each other. A common first-year composition text, Helena Viramontes’ short story “Snapshots,” provides an example of an inconsistency within a text. In this short story, the protagonist attempts to find meaning in her life by reviewing family snapshots. When discussing this text, students frequently report difficulty in following when events are happening. However, inconsistent chronology is actually a key feature of the text that reflects the character’s experience of being disoriented. An example of inconsistencies across multiple texts is provided by the Battle of Lexington document set from Wineburg (1991). In Document 2 of the set, the Minutemen attest to facing the British soldiers at 5 a.m. In Document 3 of the set, a novelist presents the action with sunlight glistening, suggesting a later start to the engagement. The inconsistency plays an important role in evaluating why the first shot was fired and evaluating the credibility of the sources. In both cases, recognizing the inconsistency is the reader’s first step towards analytic reading.

Recognizing inconsistencies presents several benefits. Noting inconsistencies helps readers monitor their comprehension, define their questions about a text, and trigger
problem-solving strategies. Once readers note an inconsistency, they can devise an
action plan in order to resolve the issue. However, inexperienced academic readers rarely
identify inconsistencies (Wineburg, 1991). Because the goal of a developmental-level
literacy course should be to help students become advanced academic readers, these
courses should offer instruction in recognizing inconsistencies.

University freshmen have encountered few reading experiences that would serve
to prepare them for the academic literacy expectations they will face at the university.
Assigned reading in university courses consists of multiple, challenging texts, which
students will be expected to read independently. Students will encounter a wide range of
perspectives that will often be in conflict with one another, and will, therefore, need to be
able to identify inconsistencies and develop plans for resolving those inconsistencies. A
significant portion of university freshmen, over 50% at one large, urban public university,
arrives at the university underprepared to meet rigorous literacy expectation and are, thus,
enrolled in developmental-level courses to help them build their academic literacy skills
(Goen & Gillotte-Tropp, 2003). Developmental-level integrated reading and writing
courses should provide explicit instruction in recognizing inconsistencies because that is
a key part of academic literacy and a signal to employ problem-solving strategies.
Therefore, this instructional intervention utilized direct instruction and practice with
identifying inconsistencies and creating a plan to resolve those inconsistencies.

*Expert Readers Identify Inconsistencies*

The multiple text research suggests that experienced academic readers, like
history professors, recognize inconsistencies within and across texts as part of evaluative
reading. However, because much of the research into reading multiple expository texts
has been conducted in the domain of history, many of the studies reported below involve historians and reading concerning history.

Expert academic readers identify inconsistencies across information sources. Wineburg (1991) designed a study to examine how experts and novices constructed historical understanding from inconsistent historical accounts. As noted in Chapter 1 above, Wineburg used a think aloud protocol to examine how experienced historians \((n = 8)\) and high school seniors \((n = 8)\) used sources to resolve a historical conflict across eight written documents and three paintings about the Battle of Lexington. When asked to decide which painting most accurately reproduced the events of the Battle of Lexington, experts noted more details and more inconsistencies than students.

Wineburg found that expert readers engaged in what he termed specification of ignorance which he defined in a later study as “instances when historians explicitly acknowledged confusion, expressed puzzlement or wonder, asked questions, or specified gaps in knowledge” (1998, p. 325). In particular, experts noted that the paintings did not show how the firing started and that the inclusion of a wall in one painting was inconsistent with the other accounts (Wineburg, 1991). In these instances, specification of ignorance—explicit identification of the gaps in their understanding—helped the historians analyze and interpret the paintings.

Other research has captured similar expert behavior. Wasson (2002) conducted a similar study focused on navigating multiple historical texts, but included a film clip as one text. The inexperienced students were least critical of the film clip even though they based their reasoning primarily on the information it yielded. In contrast, expert readers evaluated the utility of the film clip by noting what was not included in it. These findings
provide additional support to suggest that expert readers engage in specification of ignorance.

Investigating this phenomenon further, Wineburg (1998) found that specification of ignorance was often the trigger for enacting procedural and conditional knowledge in the form of an action plan. Wineburg recorded an advanced reader (a discipline expert in the field of American history) and an expert reader (a content area specialist in Abraham Lincoln and the Civil War period) engaging in think alouds while they read seven documents in order to determine President Lincoln’s personal stance on race. The think-aloud protocols were coded for three cognitive behaviors: use of the contextualization heuristic, intertextual linkages, and specification of ignorance. These experienced readers acknowledged inconsistencies between texts a total of 28 times. Both created “action plans” for addressing the gaps in their knowledge and reconciling the inconsistency (Wineburg, 1998, p. 335). In the earlier study, Wineburg (1991) also recorded a clear example of this phenomenon when an expert historian created an action plan to obtain information that would help resolve an inconsistency. The historian was puzzling over the anxiety level of the colonists (specification of ignorance) and pondered what other sources he might use (plan of action). He decided that he might be able to obtain personal letters in order to gain insight into the emotional state of the Minutemen. As was the case for this experienced reader, recognizing an inconsistency is often the trigger for creating a plan of action.

Inexperienced Readers Do Not Recognize Inconsistencies

However, inexperienced academic readers do not systematically recognize inconsistencies. Otero and Kintsch (1992) reported that high school students (\(N = 118\))
from four public schools in Madrid often do not notice textual inconsistencies within a short text. Tenth ($n = 116$) and twelfth ($n = 102$) graders read six short paragraphs (< 100 words), four of which contained explicit contradictions. After reading each passage, students noted any difficulties they encountered. Subjects failed to detect 40 percent of the inconsistencies in short paragraphs which contained a concluding sentence that explicitly contradicted the information provided in the paragraph. These findings support prior research indicating that inexperienced readers often do not notice inconsistencies (Scardamalia & Bereiter, 1991; Baker, 1985; Garner, 1981).

Even when asked explicitly to identify any additional information that might be helpful, inexperienced readers tend not to identify what they do not know. In another multiple texts study only 35% of undergraduates asked for additional texts to help interpret a controversy about the Gulf of Tonkin (Rouet et al., 1996). In a follow up study, Rouet et al. (1997) found that discipline novices were unlikely to ask for additional information. Only 2 of 11 participants asked for additional primary sources to help decide a controversy (Rouet et al., 1997). These findings suggest that inexperienced academic readers fail to engage in specification of ignorance.

Other researchers have found that inexperienced readers ignore contradictory evidence when reading multiple texts (Britt & Aglinskas, 2002; Scardamalia & Bereiter, 1991; Stahl et al., 1996; Wineburg, 1991). One instructor found that students enrolled in a senior sociology seminar avoided engaging with inconsistent perspectives by skipping response questions (Persell, 2004). Inexperienced readers may not notice an inconsistency or gloss over ones they have noted because inconsistencies hamper their
understanding (Stahl et al., 1996) which can make reading multiple texts a frustrating experience (Bråten & Strømsø, 2006).

Many researchers agree that inexperienced readers need instruction in recognizing inconsistencies (e.g., Bråten & Strømsø, 2006; Englert, Hiebert, & Stewart, 1988; Wineburg, 1991). Inexperienced academic readers may be aware that they are experiencing difficulty, but not be able to specifically identify inconsistencies in the text or across texts as the cause of the problem (Danner, 1976). If the reader fails to pinpoint the problem, he or she is unlikely to create a plan for resolving the issue. Because the goal of a developmental-level literacy course is to help inexperienced readers acquire academic literacy, an instructional intervention is necessary to help first-year students recognize inconsistencies, especially across multiple texts.

**The Difficulty Paper**

Salvatori (1988) advanced a Theory of Difficulty to explain the value of identifying difficulty. Difficulty has two causes: uncertainty caused by features of the text (textual inconsistency) and uncertainty caused by text-reader interaction (prior knowledge inconsistency). Salvatori explains that scholars engage difficulty as an opening to explore (1996), whereas inexperienced academic readers see moments of difficulty as an indictment of their abilities and seek to avoid getting caught up in difficulty (1988). Salvatori argues that students must be taught to view inconsistencies in texts as features “to be critically engaged rather than ignored” (1996, p. 448). In fact, acknowledging inconsistencies and attempting to resolve them lead to what Scardamalia and Bereiter (1991) term *knowledge transforming*. When readers add new knowledge gained from the text to their prior knowledge, they have engaged in knowledge
transformation. Inconsistencies open space to learn new things, to form opinions or revise perspectives, and as such are an essential part of the intellectual work of the university. In contrast, avoiding inconsistent information can lead to knowledge telling—regurgitation of facts with little or no change in understanding (Scardamalia & Bereiter, 1991). Thus, developmental-level learners must learn to engage difficulty as a way to both gain insight into a specific text and acquire academic literacy skills.

In order to help her undergraduates learn the value of difficulty, Salvatori (1988) designed a difficulty paper assignment. For this assignment, students read a course text, like T.S. Eliot’s *The Wasteland*, and write about one difficulty they encountered while reading. During the next class, students are invited to share their difficulties. Students are then asked to read supplementary materials and “chart any change in their understandings” (1988, p. 85) during an in-class writing session. In the third writing assignment, students are asked to reread the original text in order to gain further insight on the difficulty they had encountered. Through this process, Salvatori attempts to scaffold students’ interactions with the text so that they may reframe their experience of difficulty as an inconsistency in the text to be resolved instead of a failure of comprehension on their part.

Salvatori (1988) includes an example and extended analysis of a particular student’s (Jan’s) progress through this assignment sequence. Jan originally identified two instances where the text seems to refer to something that is not there (the third who walks beside the pair, but is not counted by the narrator and the water that is heard, but not seen). In writing about her difficulty, Jan explains that she experiences a “sense of disappointed expectations and an accompanying feeling of confusion and disorientation”
In her third written assignment, Jan notes that “the poem suggests…the unknown” (p. 90) and that the link between the two instances in the text is emotional, giving the reader the same feeling of frustrated expectation. Jan begins to see her response—her “difficulty”—not as simply her failure to “get” the poem but as a legitimate response to the writer’s rhetorical choices, images Eliot may have consciously used to disorient the reader in order to communicate an emotional message—frustration over what is unknown. In doing so, Jan’s engagement of a difficulty—recognition of specific inconsistencies in the poem—leads to a deeper understanding of the message of the poem and of the ways readers and writers communicate.

Miller (1994) has used a similar approach with expanded scaffolding in a composition classroom. Students are asked to note difficulties in Gloria Anzaldúa’s “Entering into the Serpent.” They are then asked to outline a plan for reconciling those inconsistencies and discuss their new re-reading of the text. A sample student assignment shows a student initially “bashing” Anzaldúa for including numerous words in Spanish, which he skipped (p. 406). In the assignment discussing his re-reading, that same student gains insight into the very features of the text that caused him difficulty. The student notes that he can “now see her strategy of using language and culture choice and placement” as a cultural commentary (p. 406). The student originally found Anzaldúa’s inconsistent language use to be an impediment to his understanding, but by focusing on that very difficulty he came to see it as a critical feature in the author’s argument.

Levinson (2003) has adapted the difficulty paper to help developmental-level learners critically engage with a single text. Levinson’s Difficulty Paper is an elaborated reading strategy composed of four parts. In Part 1 (Identifying Difficulty), students note
places that interested them or difficulties they encountered. In Part 2 (Creating a Plan of Action), students select a single, focal inconsistency and outline a plan for addressing their question. In Part 3 (Implementing Your Plan), students discuss the insights they have gained as a result of executing their plan. In Part 4 (Evaluating Your Plan), students reflect on the effectiveness of their plan. The Difficulty Paper assignment sheet which has been modified for use with multiple texts can be found in Appendix G. By breaking the process into explicit steps that mirror the problem-solving process that expert readers use, the assignment makes the process experienced readers use visible to inexperienced readers. The assignment forces students to engage with one inconsistency and to strategize about comprehension. Lastly, the assignment includes a reflective component to help students assess the effectiveness of their plan so they can fine-tune a strategic process that they can apply to future texts.

Reporting on her use of the Difficulty Paper at the California Reading Association Conference, Levinson (2003) presented a sample student difficulty paper focused on an expository piece, “Pat Cull: Carpenter” from Molly Martin’s *Hard-Hatted Women*. For Part 2 of the Difficulty Paper assignment (Creating a Plan of Action), the student selects why Pat Cull had given up a stable job as a social worker to face the hardships of becoming a carpenter as his focal difficulty and outlines a plan for answering his question. After following his plan, the student has a new perspective on Cull’s choice: she needed to find a job for which she could feel satisfaction. By attending to what the he perceived as an inconsistency (throwing away a master’s degree in social work for physically demanding work with no job security) and strategically planning a course of
action, the student came to understand the central message of the piece: satisfaction trumps benefits.

Research into the effectiveness of Levinson’s Difficulty Paper (the single text version) indicates that it is effective in helping students recognize inconsistencies. Fisher (2006) conducted a study that described developmental-level students’ interactions with texts as documented by the Difficulty Paper and how they used those readings in an essay assignment. The researcher presented a case study based on first-year students in his own developmental-level literacy course at a large, urban, public university. Students read five texts that they could incorporate in an end of the unit essay on the factors that shape identity. Students were assigned to complete the Difficulty Paper assignment for one of two short stories, either James Baldwin’s “A Stranger in the Village” (n = 6) or “Saint Marie” by Louise Erdrich (n = 8). The Difficulty Papers were collected and the first three parts were analyzed. Part 1 (Identifying Difficulty) and Part 2 (Creating a Plan of Action) were analyzed for difficulties, while Part 3 (Implementing Your Plan) was analyzed for insights gained from completing the plan. All three sections were coded for connections to the text, figurative language, the reader, or the world. The end of the unit essays were collected and examined for ways the Difficulty Paper focal texts were incorporated. The coding scheme focused on components of expository writing: thesis statement, point, illustration, and explanation.

Fisher (2006) found that completing Levinson’s Difficulty Paper assignment positively impacted students’ ability to recognize inconsistencies. The students (N = 14) generated 69 difficulties identifying specific places in the text where they encountered questions, confusion, or a break down in understanding. Analysis of the Levinson’s
Difficulty Paper assignments showed that when identifying the source of their difficulty, students elaborated on specific features of the text, including the use of metaphors and social/historical contexts. Six out of eight students used the assignment to better understand the use of metaphor in “Saint Marie,” a key rhetorical feature of the text, while five of six students connected to larger social/historical contexts, a necessary element for understanding “A Stranger in the Village.”

The difficulty paper sequence helped move students from a general sense of confusion towards identifying specific inconsistencies. For instance, one student identified the last lines of Erdrich’s short story (“Rise up, Rise up and walk. There is no limit to this dust.”) as a moment of difficulty because the inspirational message of rising up seemed inconsistent with her understanding of the insignificance of dust. In fact, the student is rightly identifying contradictory metaphorical meanings, for Erdrich is representing humans as simultaneously insignificant and transcendent. Fisher concluded that students’ engagement with specific features of the text indicate that Levinson’s Difficulty Paper is an effective scaffold for helping students identify the source of their difficulties as specific inconsistencies within the text. Therefore, in this study, direct instruction was provided to help students recognize textual inconsistencies that cause difficulties. Levinson’s Difficulty Paper, which has successfully been used to help this population of developmental-level students recognize where they are having difficulties including identifying inconsistencies (Fisher, 2006), was used as part of the instructional intervention.

In his analysis of end of the unit essays which he used as a measure of transfer, Fisher found that only half of the class (7 students) included the text for which they had
completed the difficulty paper. Although those seven students included multiple texts in their essays—the difficulty paper reading as well as other course texts—they did not demonstrate an ability to synthesize across texts. In fact, they dealt with each text in a separate paragraph. Their essays were structured in a simplistic way, starting with the thesis statement which included a list of multiple texts, instead of producing a more sophisticated, integrated essay whose thesis included a single central idea identifying the student’s message across texts. Although Levinson’s Difficulty Paper, an assignment designed for use with a single text, helped students to recognize inconsistencies within a single text, those benefits did not transfer to recognizing inconsistencies across multiple texts. Therefore, this study used a modified version of Levinson’s Difficulty Paper for use with multiple texts, referred to as the Difficulty Paper assignment (see Appendix G) in order to make recognizing inconsistencies across multiple texts explicit for students.

**Measuring Recognition of Inconsistencies**

In multiple text studies, especially those focused on history, participants are rarely asked to recognize inconsistencies. The researchers usually provide a specific controversy for the participants to resolve, such as asking participants to decide who fired first at the Battle of Lexington. Participants are asked to respond to an identified inconsistency instead of being asked to identify inconsistencies on their own (e.g., Bråten & Strømsø, 2006; Britt & Aglinskas, 2002; Stahl et al., 1996). Aside from Wineburg (1991, 1998) who identified the phenomenon of specification of ignorance, only Rouet et al. (1996, 1997) asked students to engage in specification of ignorance. Rouet et al. used a very broad measure, prompting students to ask for additional information to address a lack of information about the provided inconsistency. In this study, a specific researcher-
designed measure of recognizing inconsistencies, called Listing Inconsistencies, was utilized as part of the Multiple Text Tasks (Appendix B). Participants were asked to list any inconsistencies they noted after reading multiple documents about the Battle of Lexington.

**Summary**

Experts believe it is acceptable to not know something. In fact, to pinpoint that lack of knowledge is expert behavior. When reading multiple texts, expert readers are able to tolerate difficulty long enough to devise a plan that will enable them to arrive at a potential interpretation (Wasson, 2002; Wineburg, 1991; Wineburg, 1998). Experienced academic readers understand that they are forming a perspective, not arriving at a fact. Because inexperienced academic readers often fail to engage in this useful behavior, direct instruction is needed to help students successfully navigate multiple texts (Bråten & Strømsø, 2006; Englert et al., 1988; Fisher, 2006; Wineburg, 1991). One successful intervention is the difficulty paper (Fisher, 2006; Miller, 1994; Salvatori, 1988; Salvatori, 1996; Salvatori, 2000). In this study, a modified version of Levinson’s Difficulty Paper (2003) was used to provide practice in noticing inconsistencies across multiple texts. The effectiveness of this instructional intervention was measured using a pretest-posttest within-subjects design by asking students to list inconsistencies across seven documents regarding the Battle of Lexington.

**Developing Academic Literacy: Using Evaluative Heuristics**

The research base for evaluating sources of information, a key aspect of academic literacy, indicates a need to offer instruction in evaluative heuristics (Davis, 2003; Fister, 1992; Fitzgerald & Galloway, 2001; Kolstø, 2001; Twait, 2005). Wineburg (1991,
1998) identified three evaluative heuristics—specific problem-solving strategies—that experienced readers use to evaluate information. Research by Wasson (2001) and Rouet et al. (1997) also provide evidence of experienced academic readers using these three evaluative heuristics—sourcing, corroboration, and contextualization. These evaluative heuristics are employed by experts in a variety of disciplines who flexibly draw on all three heuristics—sourcing, corroboration, and contextualization—in order to resolve difficulties (Jacobson, 2001; Smith et al., 1991; Strømsø, Bråten, & Samuelstuen, 2003; Wyatt et al., 1993).

However, novice learners rarely employ these evaluative heuristics (Wineburg, 1991). In fact, students often approach problems in the exact opposite way that experts would. While experienced academic readers attempt to utilize as much textual information as possible to improve comprehension, inexperienced readers tend to narrow their focus when they encounter difficulties (Bråten & Strømsø, 2011; Garner, 1981; Kletzien, 1991; Stahl et al., 1996). This approach leaves inexperienced readers with fewer resources to help them resolve difficulties.

Providing students with explicit instruction in the evaluative heuristics helps students become more experienced with academic literacy, but skills for evaluating sources are often not taught at all (e.g., Devet, 2007). Even when evaluating sources is included in instruction, the process and purpose are oversimplified; most instruction in evaluating information focuses on procedural knowledge, leaving student to slot information into templates (e.g., Britt & Aglinskas, 2002; Gardner, Benham, & Newell, 1999; Wiley et al., 2009). Despite this shortcoming in instruction, educators do value source evaluation, which at the university-level requires conditional knowledge, and
educators want to help students master the evaluative heuristics (Calkins & Kelley, 2007).

One group of researchers has encountered some success teaching Wineburg’s evaluative heuristics, yet their findings are inconclusive (Nokes et al., 2007). Therefore, this intervention was built on past research and included more focus on building developmental-level freshmen’s conditional knowledge about when to employ the evaluative heuristics.

*Academic Literacy is the Goal*

Developmental-level courses in academic literacy should help students to use reading to build their background knowledge and use background knowledge to enhance their understanding of texts with the ultimate goal of helping students become independent learners (Goen & Gillotte-Tropp, 2003). Students are placed in these courses to acquire the academic literacy skills necessary to be successful at the university. In a seminal exploration of reading and writing expertise, Scardamalia and Bereiter (1991) found that literate experts engage in a dialectical process when reading: incorporating information from the text into background knowledge and integrating background knowledge with information derived from the text to increase understanding of the text. Scardamalia and Bereiter refer to this phenomenon as “knowledge transformation” because the reader’s knowledge has been transformed by adding new information (p. 179). In order to transform knowledge, literate experts tend to monitor their comprehension, identify difficulties, and engage in problem solving to resolve those difficulties. In contrast, a novice’s domain knowledge is little influenced by reading; inexperienced readers tend to engage in “knowledge telling,” which emphasizes retelling
the information contained in the text (p. 179). Because inexperienced readers are not engaging with the text, the information derived from texts generally resides in memory as isolated bits of information instead of becoming a cohesive knowledge structure. Therefore, students are not building their knowledge when engaged in “knowledge telling” activities.

Developmental-level freshmen need to acquire literate expertise in order to be successful at the university-level. Even introductory coursework at the university, such as survey courses, are designed to provide a foundation for advanced learning in later coursework. Because students are expected to read multiple texts in order to build knowledge, students must understand the complexity of the learning task, which includes using strategies, like the evaluative heuristics, to transform knowledge and make it their own. University students are asked to demonstrate knowledge transformation by applying information to novel situations, often in the form of written assessments. Therefore, students need explicit instruction in the academic literacy behaviors of more experienced academic readers/writers.

Developmental-level freshmen who are inexperienced with academic literacy tend towards a reliance on certainty, consistent with knowledge telling, while more advanced readers demonstrate knowledge transformation by focusing on essential characteristics of texts and flexibly applying strategies to improve understanding (Jacobson, 2001; Scardamalia & Bereiter, 1991; Smith et al., 1991; Wineburg, 1991; Wineburg, 1998). Experienced academic readers and writers also demonstrate flexibility through their ability to criss cross the knowledge landscape, drawing flexibly on the evidence available to them (Scardamalia & Bereiter, 1991; Smith et al., 1991; Spiro & Jehng, 1990;
Wineburg, 1998). The goal of a developmental integrated reading and writing course is to help students gain the procedural and conditional knowledge to engage in academic literacy like experienced academic readers (Goen & Gillette-Tropp, 2003). Therefore, developmental-level freshmen need instruction in source evaluation, including the evaluative heuristics used by expert readers.

**Experienced Readers/Writers Use Evaluative Heuristics to Resolve Inconsistencies**

As noted in section one of the literature review, experts recognize and attempt to reconcile inconsistencies. In contrast, students often do not recognize inconsistencies either within a single text (Otero & Kintsch, 1992) or across texts (Wineburg, 1991). However, even when inconsistencies are noted, recognition does not lead automatically to reconciling those inconsistencies. Expert readers approach inconsistencies strategically, using what Wineburg (1991) called sourcing heuristics to resolve inconsistencies within or across texts. Because one of the three heuristics is called the sourcing heuristic, the term *evaluative heuristics* has been substituted in this document to refer to all three heuristics (sourcing, corroboration, & contextualization). These three evaluative heuristics help experienced academic readers and writers resolve inconsistencies. In this section, Wineburg’s findings on evaluative heuristic use from two studies will be discussed, followed by a discussion of studies that support his findings across disciplines.

In the earlier study described briefly in section one of the literature review, Wineburg (1991) examined how experts and novices constructed historical understanding from contradictory accounts. Historians \( n = 8 \) and high school students \( n = 8 \) engaged in think alouds while they resolved a historical conflict across texts. Participants read
eight written documents about the Battle of Lexington and examined three paintings of the event while engaged in a think aloud procedure. Participants were asked to select the painting that most accurately depicted what happened at the Battle of Lexington and rank the trustworthiness of the texts. Picture evaluation was coded for description, analysis, corroboration, and qualification (including identifying limitations of information). Document evaluation was coded for sourcing, corroboration, contextualization, and qualification. The findings of this study will be discussed in greater detail below, along with the findings of the second study, which will be introduced in the next paragraph.

In order to identify finer distinctions between expert and advanced readers, Wineburg (1998) focused on the contextualization heuristic, whose use seemed unique to experienced academic readers. Wineburg examined how an expert reader (a content area specialist in Abraham Lincoln and the Civil War period) and an advanced reader (a discipline expert in the field of American history) differed in their approach to multiple contradictory texts. Both experienced readers engaged in think alouds while they read seven documents related to President Lincoln’s personal stance on race. Concurrent and retrospective protocols were coded for evidence of the contextualization heuristic, intertextual links, and specification of ignorance (moments of difficulty). The findings of both studies (1991, 1998) that relate to evaluative heuristic use will be discussed in greater detail below.

Explicit Problem-Solving Strategies: The Evaluative Heuristics

Wineburg’s research makes explicit the problem-solving strategies of expert readers who apply strategic knowledge in the form of evaluative heuristics to build new
knowledge (1991). Wineburg identified three evaluative heuristics (sourcing, corroboration, & contextualization) that experts use in navigating multiple texts.

*The sourcing heuristic.* Expert readers use source information to preview the document (Wineburg, 1991). Wineburg defines sourcing as “the act of considering the source of the document when determining its evidentiary value” (1998, p. 322). All eight expert readers attended to source information before reading the documents about the Battle of Lexington (Wineburg, 1991). Historians used the *sourcing* heuristic 98% of the time. Experienced readers use the sourcing heuristic to evaluate the value of evidence provided by different sources. Previewing source information activates genre and author knowledge. When one historian previewed the excerpt from high school textbook, a rich textbook schema was activated. This reader was more skeptical of the information from this source because of his previous experience with the limited perspective of high school textbooks. As noted by Scardamalia and Bereiter (1991), expert readers use previewing to activate relevant background knowledge to aid in understanding. For example, another experienced reader noted that source information permitted her to predict what she might encounter in the document (Wineburg, 1991). Experienced readers preview source information in order to gain a fuller understanding of texts.

*The corroboration heuristic.* Expert readers use the corroboration heuristic—a document comparison strategy—to evaluate evidentiary value (Wineburg, 1991). The think aloud protocols provide evidence of triangulating information across documents. For example, all eight expert readers evaluated the claim about Minutemen troop size in Document 4 by comparing it to other documents, especially Document 2. In addition to evaluating claims, corroboration was also used to set the scene. One historian who
scored poorly on the measure of prior knowledge used the corroboration heuristic to decide which painting most accurately depicted the Battle of Lexington. She used a couple of pieces of information from Document 4 and one detail from Document 5. She also noted details in the painting that were uncorroborated (omitted) by the document set, like the presence of women and a hill. Similarly, seven of the eight expert readers rejected the claim that the minutemen “stood their ground” (Document 7) based on discorroborating evidence from other documents.

The contextualization heuristic. Contextualization is part of constructing a problem space when faced with an ill-structured problem. The contextualization heuristic refers to attempts to reconstruct the spatial-temporal scene of events (1991). For this historical event, students and historians needed to decide on the time of day and the order of events leading up to the assembling of troops on Lexington Green. For instance, one historian found support for the claim that the usually disciplined British troops violated direct orders to hold their position, by noting that the British were tense, sleep deprived, and wearing wet, uncomfortable clothing. In order to better understand the event, another expert reader reconstructed the anxiety the minutemen must have felt after (and while) waiting three hours for the British to arrive.

Locating events in place and time helps experienced readers decide on the credibility of evidence. One historian used information from Document 2 to situate the event in time. If it was around five o’clock in the morning, it was dark. This expert reader questioned how credible eyewitness testimony was since the event occurred in the dark.
Wineburg (1998) focused on experienced readers’ use of the contextualization heuristic because the earlier study indicated a marked difference in its use between inexperienced and experienced readers. In order to gain a finer grain view, the coding for contextualization was divided into six subcategories: spatio-temporal, social-rhetorical, biographic, historiographic, linguistic, and analogical comments. The expert reader utilized the contextualization heuristic more than the advanced learner (61% of the time vs. 31% of the time). The expert reader drew frequently on his background knowledge to contextualize events. He had a higher percentage of historiographic and analogical comments (11% vs. 8%, and 11% vs. 4%, respectively). Historiographic comments reference other historical texts, while analogical comments are explicit comparisons to other historical periods. In contrast, the advanced reader drew on documentary evidence to build a context.

As the figure below indicates, the advanced reader used intertextual links made among the documents to build an understanding of what Lincoln meant when he invoked God to discuss natural rights (Wineburg, 1998). As he read and reread the documents, the advanced reader created a historical context for understanding the documents by making a series of eight intertextual links. Throughout the think aloud the advanced reader made a total of 20 intertextual links. This recursive process echoes Scardamalia and Bereiter’s (1991) claim that expert readers use background knowledge to comprehend a text, while simultaneously assimilating knowledge from the text to build their background knowledge. This zig-zagging pattern of intertextual links corresponds to Spiro et al.’s description of criss-crossing the knowledge landscape (Spiro et al., 2004).
Despite the differences in their background knowledge, both experienced readers relied heavily on the contextualization heuristic to resolve the historical inconsistency. Wineburg’s research indicates that experienced readers use all three heuristics flexibly, selecting the most strategic approach to resolving inconsistencies as they read (1991, 1998).
Experienced Readers Differ from Inexperienced Readers

Wineburg (1991) suggests that students need strategy instruction to help them think more like experienced readers. In examining the Battle of Lexington, expert readers used the sourcing heuristic 98% of the time, while students used this heuristic 31% of the time (Wineburg, 1991). Wineburg presents several examples of problems that students could have avoided by attending to source information. In one instance, an inexperienced reader struggled to understand the perspective of a document, including establishing pronoun references. Another student reached the end of a document where the attribution was located and exclaimed, “Oh, my God it is British” (p. 79). In both instances, inexperienced readers expended time and energy to construct meaning from documents at a local level, when quickly previewing the source information would have provided a framework for global understanding.

Students were rarely explicit about making connections across documents. Expert readers used the corroboration heuristic referring back to documents on average six times, while students averaged just two look backs. In response to the controversial comment that the minutemen “stood their ground” (Document 7), only one student mentioned other accounts of the event. In contrast, the experienced reader with the least background knowledge used information from several documents to resolve inconsistencies. Similarly, Wineburg (1998) documented the advanced reader making eight intertextual links in an attempt to understand Lincoln’s personal stance on racism. Although experienced academic readers use corroboration to evaluate the credibility of claims and to increase their understanding, students rarely engage in corroboration.
Inexperienced readers rarely used contextualization and appeared reluctant to question the information provided. Experienced readers were able to flexibly balance their linguistic awareness, an aspect of the contextualization heuristic, with their use of corroboration to more fully evaluate claims. For example, one historian pointed out that “Stood their ground is a very different connotation from being just a bunch of rebels who won’t disperse” (Wineburg, 1991, p.81). Similarly, another expert reader commented about that same phrase: “Here we have a sense of purpose” (p. 81). Both are linguistic comments. In contrast, a third expert utilized the corroboration heuristic to examine that claim: “What, this is the seventh document? Not one of those six documents said they ‘stood their ground’” (p. 81). However, student comments suggest that they failed to evaluate the claim and were, instead, persuaded by how direct the statement was: “It seems in a way just reporting the facts...just concise, journalistic in a way, just saying what happened there” (p. 81). This inexperienced reader sees information (from the textbook excerpt) as fact and does not interrogate this version of events, preferring its clarity. Similarly, another student noted, “That’s really straightforward” (p. 81). In both cases, the student relies on the clarity of the writing, without questioning the information.

Although experienced readers frequently use the evaluative heuristics to understand events and evaluate the credibility of evidence (1998), students rarely engaged with the heuristics, most notably corroboration and contextualization (1991). These inexperienced readers fail to use strategies to improve their understanding of the event or their evaluation of the information (1991).
Experienced Readers Use Evaluative Heuristics across Disciplines

Wineburg’s two studies indicate that experienced readers use evaluative heuristics when attempting to resolve historical controversies. Rouet et al.’s (1997) findings suggest that experienced academic readers transfer strategies across reading situations, including those in different disciplines. When graduate students in psychology were asked to complete a history task with multiple texts, they used document knowledge from their field. This suggests that students do not need to know how to read differently in each domain, but rather how to be more effective as readers and to transfer effective strategies from one domain to another.

Although Rouet et al. (1997) intended to contrast discipline specialists’ (graduate students in history) and discipline novices’ (graduate students in psychology) use of study strategies, evaluation of documents, and use of documents in an essay, their study provides further insight into the comparable behaviors of advanced readers. As graduate students, both groups were familiar with and successful at satisfying expectations for academic literacy, and as such both groups could be considered advanced academic readers and writers.

The participants—graduate students in history ($n = 11$) and graduate students in psychology ($n = 8$) at a French university—read documents sets comprised of conflicting accounts of two controversies regarding the Panama Canal. Participants studied the seven documents (2 historian essays, 2 participant accounts, 2 official documents, and 1 textbook excerpt) and background facts for up to 15 minutes in a hypercard environment. Participant were then given 10 minutes to draft an essay expressing their opinion on the controversy and were explicitly told that they were not expected to quote from
documents. Participants then ranked the usefulness and trustworthiness of the documents and provided one-sentence justifications of their ranking of each document.

Despite differences in discipline background, both groups of advanced learners used similar types of justifications of trustworthiness. Justification types were coded as content, source, task, and other. The source justifications most closely match Wineburg’s sourcing heuristic. Both groups ranked documents that did not take an explicit stance on the controversy (the textbook excerpt and the official documents) to be more trustworthy than opinionated documents (participant accounts and historian essays). Official documents (military correspondence and 1846 Treaty) were considered by both groups to be the most trustworthy. Graduate students in history were only slightly more likely than the graduate students in psychology to use source justifications (49 vs. 41 times). Overall, the groups did not differ significantly in the types of justifications they used, $\chi^2 (3, N = 311) = 7.07, ns$.

Use of the sourcing and corroborating heuristics in the decision essays was comparable. Sourcing was operationalized as referencing a document, so each mention of an author’s name or a document was scored one point. Both groups attended to source information and 92% of the essays contained at least one reference. Both groups focused their referencing on the official documents (62% of the references).

Both groups demonstrated use of the corroborating heuristic with equal frequency. Three types of corroborations were identified: positive connections between sources showing that documents agreed, negative connections between sources showing an inconsistency across documents, and general references to groups of sources. Discipline novices (psychology graduate students) tended towards general references.
Only one expert utilized negative corroboration. Two discipline novices and two discipline experts pointed out evidence that was missing or might be helpful, a component of specification of ignorance.

The two groups differed only in their use of the contextualization heuristic. Three categories of contextual statements were identified: specific *problem context statements*, general *historical context statements*, and *general context statements*, which were not specific to historical thinking. Graduate students in history used significantly more contextual statements in their essays, $F(1, 17) = 5.94, p < .05$. Context statements were defined as “statements presenting relevant information that could not be traced back to the documents” (p. 98). Discipline specialists included more historical context statements than novices (55% vs. 20%). Graduate students in history tended to begin their essays with a statement contextualizing the controversy. Qualitatively, the context statements made by discipline specialists were more elaborate.

Participants in this study tended to demonstrate use of evaluative heuristics. Both groups used sourcing and corroboration. Differences in contextualization differed along discipline knowledge lines. However, the inclusion of the *general context statements* category in the coding scheme for contextualization suggests that participants could demonstrate contextualization even without specific historical knowledge. Graduate students in psychology did not incorporate general knowledge, suggesting they would benefit from instruction to make the task demands for contextualization more explicit. In general, participants were able to draw on their discipline-general problem-solving skills, even when they did not have discipline-specific content knowledge. These graduate students join Wineburg’s expert general historian (1998) as advanced learners using
evaluative heuristics, even when faced with deficits in background knowledge about a
topic. Instruction should promote heuristic use at the undergraduate level in order to give
students, who are inexperienced with academic literacy, strategies to use across the
disciplines, especially those in which they are discipline novices (Rouet et al., 1997).

Strømsø et al. (2003) found that successful first-year law students use evaluative
heuristics to improve comprehension of independent reading material. Although first-
year law students are not yet discipline-experts, they have expertise in academic literacy
having successfully completed four years of undergraduate coursework. Strømsø et al.
present three case studies. Two cases suggest a sophisticated approach to texts (Cases 1
& 3), while Case 2 indicated a naïve approach to multiple texts. Although Case 2
possessed strong decoding and comprehension skills, she demonstrated fewer intertextual
links, fewer instances of the evaluative heuristics, and a low end of the year examination
score. Cases 1 and 3 possessed low to average decoding and comprehension skills, but
made intertextual links, used all three evaluative heuristics, and, therefore, earned higher
scores on the end of the year examination, which required applying legal knowledge.
Case 1 utilized strategies such as previewing, a component of the sourcing heuristic, and
demonstrated the zig zag pattern of reading similar to the advanced learner in Wineburg’s
study (1998), a component of the corroboration heuristic. Case 3 used the evaluative
heuristics to resolve “perceived inconsistencies” as successful academic readers do
(Strømsø et al., 2003, p. 137). These three case studies indicate that advanced readers,
even those with weak decoding skills who use evaluative heuristics benefit both in
attaining more sophisticated reading behavior and improved course performance
(Strømsø et al., 2003).
Findings from Wyatt et al. (1993) suggest that skilled academic readers use evaluative heuristics as part of academic literacy tasks. Wyatt et al. investigated the reading behaviors of social scientists \((N = 15)\) who were asked to engage in think alouds while reading research articles from professional journals. Expert readers monitored their comprehension (e.g., noting difficulty), employed a range of strategies (e.g., surveying the reading), and evaluated the texts (e.g., engaging in source evaluation). Expert academic readers were flexible in their strategy use. In addition to using a wide range of reading strategies, these expert readers were continually using evaluative heuristics throughout the reading process.

In an expert/novice study of complex problem solving, Jacobson (2001) captured evidence of professionals in science and engineering—disciplines that vary substantially from history and English—employing evaluative heuristics. Undergraduates in the humanities and social sciences \((n = 7)\) served as the novices, while credentialed, advanced graduate students and scientists who belonged to a scientific professional organization \((n = 9)\) represented the experts. Both groups responded to questions about complex problems (e.g., How would you design a city so that there will be goods and services but minimal shortages or surpluses?). The protocols were coded for eight ontological component beliefs and categorized as either consistent with a Clockwork Mental Model (which corresponds to a basic, mechanistic understanding of complex phenomena) or a Complex Systems Mental Model (which corresponds to an understanding of the interrelationships between aspects of a complex system). Significant differences were found between the two groups, including for the revised component belief scales. The scientists possessed more background knowledge,
increased scientific literacy, and a discipline-specific epistemology. But, they also used evaluative heuristics. For example, one expert response offers an example of contextualization through the use of an analogy, explaining that a component in a complex, non-linear system cannot be labeled as causal by saying that “[a butterfly] no more ‘causes’ a storm than a baseball pitcher ‘causes’ a homerun” (p. 45). The coded responses of both experts that were provided in the article showed evidence of the sourcing and corroboration heuristics (naming a source and triangulating across sources). Experts in a complex scientific discipline employ evaluative heuristics in order to solve ill-structured problems in their domain (Jacobson, 2001).

Scientists in the medical field have also been observed using evaluative heuristics. Smith et al. (1991) conducted two studies to investigate how immunohematologists resolve multiple solution problems. In Study 1, an expert immunohematologist with ten years of experience engaged in a think aloud protocol while reading blood panel data sheets to identify antibodies for fifteen test cases. This is a complex problem-solving (ill-structured) domain because multiple antibodies could be present at the same time and a given antibody may produce varying patterns of test results. The expert tended to analyze the problem, by breaking it down into a series of simpler questions. He engaged in evaluative behavior, looking at multiple data sets for each case (sourcing) and attempting to collect convergent evidence (corroboration). In one case (Case 4), he overturned his previous decision to rule out a particular antibody \( P_1 \), suggesting that he understands the uncertainty built into the process.

In a second study, sourcing and corroboration were frequently used heuristics. In Study 2, subjects at varying levels of experience \( N = 4 \) each attempted to identify
antibodies in one panel while engaging in a think aloud procedure (Smith et al., 1991).

Even the inexperienced intern engaged in evaluative heuristics: contextualizing the case, evaluating multiple data sources (sourcing), and looking for converging evidence (corroborating). The least experienced subject, a student, looked for certainty, leading to an incorrect solution. In contrast, the expert from Study 1 carefully considered candidate solutions, suggesting that he trusts the process. All four subjects used a variety of strategies, including the evaluative heuristics to reach a conclusion (Smith et al.).

From both a theoretical perspective and a research perspective, there is evidence that flexible use of strategic knowledge extends across disciplines. Experienced academic readers are flexible in their strategy selection, evaluation approach, and application of information because they are accustomed to working in ill-structured problem spaces. Experts in varied disciplines utilize evaluative heuristics—sourcing, corroborating, and contextualization—to evaluate texts and to resolve inconsistencies (Jacobson, 2001; Scardamalia & Bereiter, 1991; Strømsø et al., 2003; Smith et al., 1991; Wineburg, 1991; Wineburg, 1998; Wyatt et al., 1993). The evaluative heuristics are useful problem-solving tools that transcend discipline specific boundaries. Because of their evident utility, evaluative heuristics should be taught to developmental-level university students as a means of helping them progress from introductory to advanced knowledge acquisition in the variety of disciplines they encounter while completing their general education requirements.

Inexperienced Readers Rarely Evaluate Information Sources

The previous subsection explored the research on the academic literacy behavior of experienced readers. Experienced academic readers tend to understand that they are
working in complex, ill-structured domains (e.g., they notice inconsistencies) and are aware that they need to flexibly apply knowledge. Experts in many disciplines rely on the evaluative heuristics first identified by Wineburg (1991) to help them apply knowledge and solve complex problems. In contrast, inexperienced readers rarely use evaluative heuristics when faced with inconsistencies. In fact, although inexperienced readers know they should evaluate sources, they often do not evaluate sources (Eysenbach & Köhler, 2002). This section will look at how novice behavior differs from that of the experienced academic reader.

Students rarely verify information from sources (Metzger et al., 2003; Flanagin & Metzger, 2007), relying instead on superficial criteria: convenience (Fister, 1992; Fitzgerald & Galloway, 2001; Metzger et al., 2003; Twain, 2005), relevance (Kolstø, 2001; Twain, 2005), design elements (Calkins & Kelley, 2007; Flanagin & Metzger, 2007), and currency (Metzger et al., 2003). Because students often use superficial evaluation criteria, instruction in evaluative heuristics is essential to help them become more expert academic readers. Use of inappropriate sources “merits attention because it both devalues and places at risk a central assumption of academic writing: that a writer will support claims with appropriate, valid, and authoritative evidence” (Burton & Chadwick, 2000).

Instead of attending to source information as expert readers do, undergraduates indicate an over-reliance on content for selecting sources. In a qualitative study of undergraduates’ source selection behavior, thirteen undergraduates who were working on research assignments were interviewed and engaged in a think aloud while conducting research for class assignments at the library (Twain, 2005). Students indicated that
content was the major criteria for source selection, as has also been found in previous research (e.g., Fister, 1992; Fitzgerald & Galloway, 2001). Perceived relevance to their topic, which was mentioned 51 times (44% of total responses), was the most often utilized source selection criteria for students in Twait’s study. Familiarity (15%) was the second most common criteria for selecting sources (Twait, 2005). Relevance and familiarity of the content were two factors that influenced undergraduates’ selection of information.

Similarly, Bråten et al. (2009) found that undergraduates rely primarily on content to judge trustworthiness. Norwegian undergraduates ($N = 122$) preparing to enter a teacher education program completed a measure of prior knowledge and read seven texts on climate change. Participants ranked the extent to which they trusted each source and the trustworthiness of six source categories (author, publisher, type of text, content, own opinion, and date of publication). Content was the highest ranked reason for trusting a source (Bråten et al., 2009).

Convenience is also a top reason for source selection decisions (Fister, 1992; Fitzgerald & Galloway, 2001; Kolstø, 2001; Metzger et al., 2003; Twait, 2005). In a qualitative study of Norwegian high school students ($N = 22$) reconciling a controversy about the role power transmission lines might play in incidence of childhood leukemia, Kolstø (2001) reported that students only superficially evaluated sources. Convenience and not quality dictated which sources were used to complete the task (Kolstø, 2001). This trend towards superficial examination of sources may be exacerbated by the growing prevalence of Internet research (Metzger et al., 2003). Dependence on convenience might also explain the shift from using peer reviewed journals to using more popular yet
less credible media, such as magazines and websites, which Davis and Cohen (2001) and Davis (2003) documented.

Students, including undergraduates, rarely evaluate information sources. Instead they rely on convenience, relevance, and other superficial source characteristics when deciding to use information. Consistent with their failure to evaluate information, research into their use of evaluative heuristics suggests that inexperienced readers use evaluative heuristics less frequently than more experienced readers.

Inexperienced Readers Do Not Use Evaluative Heuristics

Wasson (2001) corroborated Wineburg’s findings that inexperienced readers are less likely to use the evaluative heuristics employed by experienced academic readers. Wasson (2001) replicated Wineburg’s (1991) study, describing the heuristic use of experienced and novice readers in Canada. Instead of using the document set about the Battle of Lexington, Wasson used 14 documents—written, pictorial, and film—about the Battle of the Plains of Abraham—an equivalent historical event in Canada. University historians ($n = 6$), three of whom represented expert readers and three of whom represented advanced readers, and high school students ($n = 6$) viewed the documents while engaged in a think aloud procedure. Then, all participants were asked to decide why the British had won the battle. Finally, all participants rated the trustworthiness of the documents.

Sourcing heuristic. In support of Wineburg, Wasson (2001) found that inexperienced readers used the sourcing heuristic less frequently than experienced readers. High school students demonstrated 29 instances of sourcing compared to 140 instances by the more experienced historians. Document 6 (Memoir on Canada) was a
sourcing focal point for experienced academic readers. These experienced readers commented extensively on the credibility of the source because it did not have an author or a date. In their discussion of the information, experienced readers tended to qualify their inferences because of the questionable pedigree of the information provided. Expert readers—historians whose specialty covered the Battle of the Plains of Abraham—were twice as likely to use the sourcing heuristic as advanced readers (91 vs. 49 instances).

Wasson used less rigorous criteria for sourcing than Wineburg (1991). Novices were credited with “superficial sourcing” (25 of the 29 instances of sourcing) for identifying source information even when they did not evaluate how it might affect the information (p. 7). The high school students demonstrated 4 instances of “deep sourcing” in contrast to the 93 instances exhibited by the experienced readers (p. 7). Despite this generous coding scheme, experienced readers used the sourcing heuristic significantly more often than novice readers.

The ranking of credibility provides further evidence that novice readers do not evaluate information in light of source information. The historians ranked the primary documents highest in trustworthiness. In contrast, the novices ranked the secondary sources, like the film clip, as more credible.

*Corroboration heuristic.* Corroboration was the least used heuristic, yet experienced readers used the corroboration heuristic seven times as often as novice readers (Wasson, 2001). All the historians used the corroboration heuristic for a total of 28 instances. In contrast, only three of the students used it for at total of 4 instances. In particular, the experienced readers criss crossed between Document 2 (Letter from the
British general to his troops) and Document 3 (Letter from British officers to General Wolfe) as they tried to establish how well coordinated the British strategy was.

The pictorial documents provide a window through which to view the differing behavior of experienced and inexperienced readers. Experienced academic readers tended to delineate the uncorroborated details present in the pictures. Historians’ use of the corroboration heuristic took the form of wanting additional information. However, students used the pictures to inform their understanding of the battle without corroboration from other documents. Additionally, novices relied on information that could not be corroborated within the paintings. For instance, one student discussed the emotional state of the Quebecois as inferred from their facial expressions in Document 7 (Painting of the Quebec Ruins), despite the inability to see facial expressions in the painting. Students tended to draw uncorroborated inferences from the pictorial documents.

Source 1—the video clip—serves to differentiate the corroborating behavior of participants at each skill level. Experts were immediately able to critique what was not included in the video clip. Advanced learners were able to use the corroboration heuristic to critique the film clip after reading other documents. In contrast, students used the film as the main source for their response to the decision question of why the British were successful, yet devoted the least amount of analysis to the clip during the think aloud as measured by number of transcribed lines.

*Contextualization heuristic.* Wasson found that experienced readers were seven times more likely to demonstrate the contextualization heuristic than novice readers. The instances of contextualization demonstrated by students were all classified as “intra-event
oriented,” which means they were focused on creating a spatial and temporal context within the event (p. 59). For example, all students who contextualized addressed the time between the British arrival in Quebec and the start of the battle (a month). Experienced readers also provided temporal contextualization. All historians created timelines, and they incorporated “extra-event oriented” context which included outside knowledge to contextualize the event (p. 59). Historians demonstrated 33 instances of intra-event contextualization and 29 instances of extra-event contextualization for a total of 62 instances of contextualization. In contrast, students demonstrated a total of 9 instances of contextualization (6 intra-textual and 3 extra-textual). Experienced readers used the contextualization heuristic significantly more often than inexperienced readers and were able to bring in a broader range of contextual knowledge.

In Wasson (2001), as in Wineburg (1991) experienced academic readers utilized the evaluative heuristics. However, use of the evaluative heuristics by inexperienced academic readers was rare. Wasson provided a more generous definition of the heuristics, for example including superficial sourcing when Wineburg would only have credited deep sourcing. Even with this more generous coding scheme, inexperienced readers demonstrated significantly fewer instances of heuristic usage than experienced readers.

Rouet et al. (1996) also used a generous coding scheme, but found few instances of evaluative heuristic use among undergraduates. The participants (N = 24), who were on average 20 years old, read a constructed textbook excerpt comprised of agreed upon facts about the Panama Canal from other documents. The experimental group read two primary source documents, which were replaced by two history essays for the control
Participants wrote a one-page decision essay for each of four explicitly presented controversies (e.g., justification for U.S. intervention in the Panamanian revolution) and ranked the usefulness and trustworthiness of each document.

Superficial sourcing dominated participants’ justifications of document credibility. Content justifications were given for historian essays and the constructed textbook excerpts. Obvious factors such as being an eyewitness or participant were common for participant accounts, which were coded as author justifications.

Despite a more relaxed criteria for counting instances of evaluate heuristic use, students exhibited minimal use of evaluative heuristics. In the decision essay, two types of references were identified: general (e.g., “according to the texts I have read”) and specific. Wineburg’s (1991) coding scheme would not have credited general textual references as heuristic usage. Although all students explicitly referred to a document in at least one of the four controversies, only one student did so for all four essays. Roughly half of the essays (52%) contained at least one specific reference. In other words, nearly half of the essays had no explicit reference to the document set and, therefore, could not be considered sourcing even with a generous coding scheme that only required referencing that documents had been used. Findings from Rouet et al. (1996) suggest that students rarely use the evaluative heuristics.

Britt and Aglinskas (2002) reanalyzed data from Rouet et al. (1996), finding that participants rarely used the evaluative heuristics, basing the majority of their evaluations on content or personal opinion. Britt and Aglinskas categorized the approximately 1500 student justifications for trustworthiness and usefulness. Rouet et al. had used a four category system (content, author, document, and opinion). Britt and Aglinskas provided
a finer-grain categorization scheme with seven categories: author position (e.g., credentials), author motivation, author participation (e.g., how the author came to know the events), author evaluation, date, document type, and document evaluation. Author justifications accounted for 24% of the justifications, while document characteristics accounted for 17%. More than half of the justifications were based on the content or opinion instead of assessing the credibility of the source.

Britt and Aglinskas (2002) found that students do not spontaneously use evaluative heuristics to evaluate source information. High school (n = 60) and college students (n = 49) read six documents related to the Panama Canal and took notes that could be used to respond to test items because they would not have access to the documents after the reading phase. Participants completed a 14-item sourcing skills test which included two essay questions. Sourcing scores composed of the number of correct answers on the sourcing skills test (23 points possible) and the number of pieces of correct source information that students recorded on the notes sheet (42 points possible) were calculated. High school students earned on averaged 15% of available points on the sourcing skills test and undergraduates averaged 23%. Of the students who answered at least one of the listing facts questions 32% of high school students and 41% of undergraduates used information that was unique to the excerpt from a historical novel, suggesting that they did not utilize the sourcing or corroboration heuristics (Britt & Aglinskas, 2002).

Wiley et al. (2009) also found that students do not spontaneously employ the evaluative heuristics as experienced academic readers do. Wiley et al. investigated the effect writing task manipulation (argument prompt vs. description prompt) had on depth
of understanding with regards to an Internet-based science inquiry activity. Although students read the same document set, one group received a writing prompt that asked them to compose an argumentative essay (argument prompt), while the other group was prompted to describe the situation (description prompt). Participants ($N = 110$), whose average age was 19, read the seven edited Internet texts about volcanoes. Four of the sources were credible and three were not. Participants were asked to rank and justify the trustworthiness of the texts. The sourcing and corroboration heuristics were infrequently used. Relevance was the most frequent type of justification (25.1% with the description prompt vs. 43.6% with the argument prompt). The closest measure of Wineburg’s sourcing heuristic was source justifications that accounted for 10.1% and 4.3% of the justification, respectively. Agreement in justifications, the closest match to Wineburg’s corroboration heuristic, accounted for 7.0% and 3.6% of the justifications, respectively.

One participant cited repetition across sources (corroboration) as a justification for a low credibility ranking. Although the participants assigned to the argument prompt condition used slightly more heuristics, the undergraduates in this study rarely used the evaluative heuristics used by experienced academic readers (Wiley et al., 2009).

Students rarely evaluate source information in a manner consistent with university expectations for academic literacy. Students rarely use general strategies for source evaluation. Often they do not evaluate sources at all. When they do attempt to evaluate sources, they use surface features instead of rigorous evaluation criteria. Students do not use the specific strategies—evaluative heuristics—as frequently as experienced academic readers (Britt & Aglinskas, 2002; Rouet et al, 1996; Wasson, 2001; Wiley et al., 2009).
The rarity of identifiable evaluative heuristic use by students in most multiple text studies suggests that student use of evaluative heuristics is not deliberate or strategic.

*Inexperienced Readers Narrow Focus in the Face of Difficulty*

Inexperienced academic readers do not spontaneously evaluate information sources. Failure to use evaluative heuristics can compromise comprehension. Inexperienced readers tend to be text-bound (Keck, 2007) and to take a piecemeal approach to reading texts (Bråten & Strømsø, 2011; Strømsø et al., 2003). When inexperienced readers encounter difficulty (i.e., textual inconsistencies) they narrow their focus. This hyperfocus prevents them from making use of resources like source information or contextual cues.

Based on his findings, Wasson (2001) suggested that students approach texts as “snapshots” reading each documents in isolation as if it could provide the whole picture. Their use of the video clip as the primary basis for answering the decision question, despite the presence of 13 other documents, supports this interpretation. Wasson noted that students asked questions and then let them go. Wasson characterizes the questions that students posed as “roadblocks” (p.16). In contrast, experienced readers used questioning as a “vehicle for gaining deeper understanding” (p.16). The students tended to view the information in each document separately from the other documents. In contrast, the experienced readers used sourcing, corroboration, and contextualization to bring in additional information in order to create a broader understanding of the event.

Prior research has suggested a tendency among inexperienced readers to narrow their focus when they encounter difficulty. Although this may allow for a closer reading of the difficult passage, it deprives the reader of rich sources of additional information,
including source information, context provided by the document, corroborating information in other documents, and applicable prior knowledge. Kletzien (1991) found that as text difficulty increased, inexperienced readers used fewer strategies and fewer types of strategies. Forty-eight average-ability high school students were divided into good comprehenders and poor comprehenders. All participants read three expository cloze passages adapted from high school social studies textbooks. The passages were modified to represent independent, instructional, and frustration level passages for each group. Poor comprehenders used fewer strategies as text difficulty increased. Although good comprehenders used more and different strategies on the instructional and frustration passages, both groups tended to narrow their focus as the difficulty increased. Participants reading at the independent level used broad context (e.g., visualization and prior knowledge). Participants reading at the instructional level used more passage-level strategies (e.g., organizational strategies). When reading at the frustration level, participants focused only on the chunk they were struggling with, attending to microlevel information, such as new vocabulary. As they encountered increased difficulty, participants focused in on the difficulty to the exclusion of potentially helpful additional information. These inexperienced readers failed to use the evaluative heuristics that experienced academic readers use to resolve inconsistencies.

Garner (1981) noted a “piecemeal-processing” approach in a study comparing the strategic monitoring of middle school students who were categorized as poor comprehenders and good comprehenders (p. 159). Poor comprehenders attended to intra-sentence comprehensibility more than inter-sentence comprehensibility. By extension, this suggests that struggling readers would attend more within a text than between texts.
Any reader, whether inexperienced, advanced, or expert, becomes a poor comprehender when they face difficulties, such as inconsistencies. Therefore, any struggling reader faced with an inconsistency is prone to circling the strategic wagons, when experts would open up the options for finding helpful information. All three evaluative heuristics ask inexperienced readers to be extra-textual and attend to the source, other texts, and other knowledge and information in addition to the text being evaluated. The evaluative heuristics could provide struggling readers with alternative avenues for incorporating additional information to resolve inconsistencies.

*Failure to Employ Evaluative Heuristics Impacts Comprehension*

Evaluating sources is not just an extraneous activity in which academic readers are expected to engage. Evaluating sources predicts comprehension (Bråten et al., 2009; Stahl et al., 1996; Strømsø, Bråten, & Britt, 2010; Wiley, 2009). When students do not attend to source information, it cannot be used to aid in text comprehension. Readers who do not evaluate source information are depriving themselves of resources that might aid in comprehension.

A study of how high school students negotiate multiple texts in history demonstrates that failure to use the evaluative heuristics goes hand in hand with an inability to engage in knowledge transformation as a literate expert (Stahl et al., 1996). Although 44 sophomores in Advanced Placement (AP) US history participated in Stahl et al.’s study, only twenty worked independently and provided complete data sets. These twenty high school sophomores completed measures of prior knowledge, read a set of six documents about the Gulf of Tonkin Incidents and five documents about the Gulf of Tonkin Resolution, took notes, completed a task-evaluation questionnaire, completed a
free recall task, and as a final writing task were assigned to either write their opinion or a description on the events.

A test of relational knowledge suggests that students did not integrate knowledge from the texts (Stahl et al.). One of the prior knowledge measures was the Gulf of Tonkin Relationship Task in which students rated the strength of the relationship between all possible pairs of 10 key words or phrases (e.g., North Vietnam, Defense, Aggression). Students also completed this task after reading each document. Sophistication of students’ mental models was measured by comparing their scores on the relationship assessment with the scores of advanced readers (e.g., a high school history teacher). The correlation at pretest between students and the experienced reader average was .26, indicating that students had little prior knowledge. The gain to .42 after reading one text indicates that students learned about some relationships. However, there was little gain from any additional readings, indicating that inexperienced readers did not integrate additional information from additional sources of information (Stahl et al.).

Coding of student work products suggests that students did not integrate knowledge from texts (Stahl et al.). The notes, free recall task, and final writing task were divided into idea units and coded for level of integration as copying, paraphrasing, reducing, making a gist, evaluating, or distorting/misleading. Reducing was described as summarizing across two or more sentences in the same text. Making the gist was defined as replacing nouns with superordinates. Evaluating was defined as “stating an opinion about the ideas in the text that were not merely the copied opinions of the authors or the opinions of the people the authors described” (e.g., “Johnson was an idiot”). The first three categories (copying, paraphrasing, & reducing) demonstrate surface level
understanding of texts and accounted for 47% of idea units. *Evaluating* included forming opinions without a textual-basis, accounting for 4% of the idea units. Only *making a gist* suggests that students were integrating information into knowledge, accounting for 15% of the idea units (Stahl et al.).

Analysis of the final writing tasks indicates that inexperienced readers were either text-bound or ignored the texts. Students, when asked to write an opinion, used broad generalizations without including factual evidence. Stahl et al. describe one example as “the task of giving his opinion was viewed as being disassociated from obtaining evidence from the text to support that opinion” (p. 444). In short, when asked to write opinions students wrote broad generalizations that could not be traced back to the evidence provided in the document set. Those students writing descriptions tended to copy and paraphrase from a single text. In addition, most students produced ideas in their final writing task in the same order in which they were presented in the reading (Stahl et al.), suggesting that they did not integrate information into a cohesive understanding, but rather engaged in knowledge telling which isolated each piece of information.

Despite broadly defined heuristic categories that required identification instead of evaluation, students rarely used sourcing, corroboration, or contextualization (Stahl et al., 1996). Students were credited with sourcing for mentioning the author or document, without making any evaluation of the evidence based on that information. Students were credited with corroboration for mentioning two documents. Students were credited with contextualization for mentioning the time of the events. In all three cases, students were credited with heuristic use for identifying information that they should then have evaluated, but were not required to engage in critical evaluation. Despite the liberal
redefinition of Wineburg’s heuristics, instances of their use were rare. Sourcing was by far the most frequently credited category with 30 instances. However, the authors noted that “the comments classified as sourcing did not use the source to understand the text…but merely noted it” (p. 446). Five instances each of corroboration and contextualization were recorded (Stahl et al.). This is a very small number, especially in light of the broad operationalized definitions.

Overall, students demonstrated superficial engagement with texts, tending to copy and paraphrase information in the order they encountered it or ignore the text altogether in favor of broad generalizations (Stahl et al.). There was little evidence of evaluation: instances of evaluation included unsupported opinion, Wineburg’s heuristics were rarely used, and consistent with the definitions of Wineburg’s heuristics in this study, evaluation was not required. Therefore, direct instruction in using evaluative heuristics might be a way to help inexperienced readers, such as these, interact critically with multiple texts.

*Previous Instructional Interventions for Evaluating Sources*

Evaluating sources is valued at the university (Intersegmental Committee, 2002). However, students do not spontaneously evaluate sources. For example, university freshmen enrolled in a world history course reported never before having been asked to evaluate a Web site (Calkins & Kelley, 2007). Despite faculty agreement that students need these skills to be successful in university courses (Intersegmental Committee, 2002), skills for evaluating sources are often not taught at the university (e.g., Devet, 2007). In one survey, 79% of the respondents reported not having any previous instruction in assessing the credibility of an Internet source (Wiley et al., 2009). When evaluating sources is included as a topic for instruction, the process and purpose are often
oversimplified (e.g., Gardner et al., 1999). The resultant instruction is procedural in nature and does not help students gain the conditional knowledge about evaluating sources that they need to become successful, independent learners.

Explicit Instruction in Source Evaluation is Rarely Provided

Devet (2007) reports on “instruction” to help first-year, university students verify source information, however, no explicit instruction in evaluating the credibility of sources was provided. In Step 1, students answer several questions about the Battle of Little Big Horn by engaging in Internet research, to gather what the author refers to as “verifiable facts” (p. 280). In Step 2, students read Longfellow's poem, “Revenge of Rain-in-the-Face,” about the Battle of Little Big Horn, in order to identify inconsistencies between the description of the battle presented in the poem and in the accounts presented on the Internet.

Devet (2007) accounts for the conflicting accounts of the battle by claiming that history constantly revises its conclusions. Despite explicitly encouraging students to recognize inconsistencies, this instructional approach does not include any explicit instruction in resolving contradictions across texts. While the assignment sheet does encourage students to verify information on more than one website, students do not receive any guidance in how to evaluate the credibility of Internet sources. Although the class looks for historical inaccuracies in the poem based on their understanding of the event garnered from the Internet, no instructional support was provided for evaluating internet sources so inconsistencies across potentially unreliable Internet sources remain unexamined and therefore unverified. Devet presents reconciling contradictions across texts in terms of personal opinion—encouraging the poor evaluation behavior that
undergraduates already engage in—rather than offering explicit instruction in the ways that experts reconcile conflict across texts—with evaluative heuristics. In short, students are assigned to evaluate sources, but students are provided no explicit instruction in how to evaluate sources like experienced academic readers.

*Instructional Interventions Are Often Procedural*

Because resolving inconsistencies and evaluating source information are important components of academic literacy, researchers have investigated several instructional interventions designed to help students acquire these skills. Although these interventions have registered modest success in raising students’ awareness of evaluating sources, most instructional interventions have been focused on improving students’ procedural knowledge (Britt & Aglinskas, 2002; Gardner et al., 1999; Calkins & Kelley, 2007; Wiley et al., 2009). At the university level, students need conditional knowledge of source evaluation in order to flexibly apply strategies to novel problems. Therefore, most instructional interventions have not been successful in helping students acquire the flexible academic literacy skills they will need to be successful at the university level.

One instructional intervention met with success by significantly improving high school students’ use of evaluative heuristics (Nokes et al., 2007). However, the findings are inconclusive as the change in instructional format to include multiple texts may have been the cause of improved sourcing rather than the instruction in evaluative heuristics.

*Procedural instructional intervention: Sourcer’s Apprentice.* Sourcer’s Apprentice (SA) is a computer application that prompts students to identify source features and fill in a template. This intervention is procedural in nature. Students are asked to identify pieces of information, not evaluate them. However, research indicates
that using SA contributes to identifying sources, using documents in writing, and increased causal connections when writing (Britt & Aglinskas, 2002).

Britt and Aglinskas describe SA which prompts students to identify source features. In SA, documents are presented as books on a bookshelf. Each “book” contains a title page, author page, document page, and content page. The most general document, a textbook excerpt on the historical controversy, is always presented first. Two of the books are historian accounts and four of the books are primary documents (two of which were mentioned in other documents). Students fill in “note cards” for each book by dragging and dropping or typing. The source features are categorized as author, document, and content (documents mentioned, main point, and comments). Then the students answer questions about the documents and sources, before moving on to write the decision essay. Only the note cards are available while writing the essay.

In a study conducted by Britt and Aglinskas (2002), high school students exposed to SA identified more information than a comparison group. One intact 11th grade history class ($n = 8$) received instruction through SA, while another intact 11th grade history class ($n = 7$) served as the comparison group, engaging in regular classroom activities as they studied the Vietnam conflict (Tonkin Resolution). The assessment procedure was the same as for Study 1 (which was discussed earlier in this section of the Literature Review): participants read and took notes on six documents, then completed a 14-item sourcing skills test. Sourcing scores are composed of the number to correct answers on the sourcing skills test (23 points possible) and number of pieces of correct source information that students recorded on the notes sheet (42 points possible). The pretest focused on the Battle of Lexington, while the posttest focused on the Panama Canal. A 2
(Condition) x 2 (Test Occasion) ANOVA was conducted. The SA group averaged 10.3 more items on the posttest over the pretest, while the control group identified 2.7 fewer items on the posttest than they had on the pretest. The experimental group earned about 20% more points than the control group at posttest (Britt & Aglinskas). Findings suggest that the SA prompts students to identify more source information, but not to evaluate it.

In another study, Britt and Aglinskas (2002) found that using SA increases use of sources and essay quality. The participants were two American history classes at a rural high school. The experimental group ($n = 9$) worked with eight documents concerning the Homestead Steel Strike in SA. The control group ($n = 14$) received the same eight documents which had been integrated into a single textbook-like document. The pretest was on Panama, while the post test focused on Tonkin Bay. Participants read the texts and took notes which they used to answer source and comprehension questions and draft an essay.

Essay analysis indicated that SA encouraged document use in writing (Britt & Aglinskas, 2002). A 2 condition (SA group vs. control group) x 2 source of information (narrative vs. documents) mixed design ANOVA revealed a significant main effect for source of information [$F(1, 21) = 33.65, p < .01$] and a significant interaction effect [$F(1, 21) = 8.36, p < .01$]. Post hoc tests indicate that both groups included an equal amount of narrative information in their essays [$t(21) = -1.15, p = .26$], but the SA group included significantly more document-based information [$t(21) = 4.53, p < .001$]. The SA group’s essays contained an average of three citations, whereas only one-third of the control group’s essays contained a citation. This suggests that SA prompts students to use more document information in their writing (Britt & Aglinskas).
Working with SA can also lead to improved essay quality (Britt & Aglinskas). Essay quality was assessed by a teacher grading on a five-point scale. The average grade on SA group essays \((M = 3.75)\) was significantly better than the average essay grade \((M = 2.67)\) of the control group, \(t(21) = 3.44, p < .01\). In addition, students in the SA groups used more causal connectors \([t(21) = 2.58, p < .05]\). Causal connectors are interpreted as a measure of transfer, suggesting that students who use SA are more deeply integrating what they read (Britt & Aglinskas). Findings indicate that SA contributes to including more citations and causal connectors in writing. However, no evidence of improved source evaluation is presented.

*Procedural instructional intervention: Evaluation guides.* Even when researchers report on interventions that provide explicit instruction in evaluating sources, the approach is often simplistic, reaching the level of procedural knowledge, but not conditional knowledge. Gardner et al. (1999) recommend five categories for evaluation: *authorship, accuracy, objectivity, currency, and coverage.* Gardner et al. provide an evaluation guide—the Internet Evaluation List—that students could use to evaluate an internet source. Several items ask students to identify information without evaluating it. For example, one item on the Internet evaluation guide asks students to “list the date of the last revision of the document” (Gardner et al., 1999, p. 44). Since students are assigned to list information without making any judgment about the currency of the document, they may be misled into believing that if the document is not current, it is not credible. However, there are instances in which currency is not the most salient feature. For example, a site containing primary source documents about the Battle of Lexington which has not been updated in ten years would be more credible than a site containing the
lyrics to a song referencing the Battle of Lexington that was posted yesterday. Students may follow the procedure item by item without seeing the interactions between factors, for example, between genre and currency. Because each piece of information is isolated, students may not consider the best approach to evaluating the text, which requires conditional knowledge. In fact, the evaluation guide explicitly states that all twelve items “must be present” if a site is to be deemed credible (Gardner et al.), suggesting that the absence of one piece of information would render the entire source not credible. In contrast, expert academic readers would employ conditional knowledge, selecting a few relevant pieces of information to evaluate in order to make a decision about credibility. Although students may benefit from the explicitness of the list, the itemization of information reinforces what Spiro et al. (1996) consider the inexperienced reader’s predisposition to oversimplify and treat knowledge as discrete facts. In short, learning this procedure may prevent students from building the conditional knowledge that they need to succeed in academic literacy.

However, when procedural instruction is provided, students are able to work with sources in academically appropriate ways (e.g., by contextualizing sources). Calkins and Kelley (2007) report on two case studies regarding undergraduates evaluating Internet sources. In Case Study 1, sophomores in a psychology course \(N = 20\) were offered instruction utilizing an evaluation guide—Robert Harris’ CARS checklist (Credibility, Accuracy, Reasonableness, and Support) to evaluate Internet research sources. The class performed a Google search on the topic: How music affects a child’s brain. Students are paired up, assigned an article to read, and engage in a Think-Pair-Share task as they respond to discussion questions. Students are then introduced to the CARS acronym and
brainstorm the criteria that would indicate Credibility, Accuracy, Reasonableness, and Support. The instructor introduces reading tips for reading research reports and reviews how to find articles in library databases. Students are assigned to read and critically comment on two articles provided by the instructor (Kelley) and to research the Mozart Effect and create an annotated bibliography for the six best sources. During the second session, the class debriefed about the bibliographies and critically evaluated the assigned readings based on the reading tips for research reports. Informal evaluation suggested that students who received instruction in evaluating sources tended to include more sources in their research project, use more library resources, select more credible sources, such as peer reviewed journal articles, and provide more context for the sources they include in their papers, than students who did not receive instruction (Calkins & Kelley, 2007).

Explicit instruction and instructor feedback can increase students’ use of sourcing and corroboration. In Calkins and Kelley’s second case study (2007), freshmen in a history course were asked to critique a historical Web site by comparing it with three scholarly journal articles. Calkins, the course instructor, scaffolded instruction in evaluating sources through the use of an evaluation guide, in this case detailed worksheets. Students were assigned to write an eight-to-ten-page research paper about a historical topic. After students selected a topic, they were asked to submit one credible website and three scholarly sources on that topic. After the instructor (Calkins) provides feedback on their source selection, students completed a detailed Historical Source Evaluation Worksheet. The worksheet contains items, such as What is the author’s evidence? (Archeology? Poetry? Weapons? Tapestries?) Is this a different kind of
evidence from the other sources?; Briefly describe the content and structure of the website; How does the website compare in substance and tone with your scholarly texts?; and How would you evaluate the overall strengths and weaknesses of the Web site (e.g., navigability, useful links, graphics, visual interest)? (Calkins & Kelley, p. 155). Students receive feedback after they complete the detailed evaluation worksheet, which they can use to complete the next assignment: writing a critical essay analyzing the content of the website and comparing it to the three scholarly sources. The instructor also provides feedback on a draft of each student’s critical essay. As a result of explicit instruction and instructor feedback, students were able to enact sourcing and corroboration heuristics, just as expert readers do (Calkins & Kelley, 2007). However, it should be noted that the researchers did not present data to support these conclusions. Instead, they rely on informal evaluation and instructor observation.

Procedural instructional intervention: SEEK. Instruction that focused on the SEEK template for source evaluation lead to increased use of corroboration heuristics and more sophisticated essays (Wiley et al., 2009). SEEK is an acronym for four essential aspects of source evaluation: Source of information; the nature of Evidence; how information fits with the Explanation; and how new information fits with prior Knowledge. Sixty participants were assigned to either the experimental (SEEK) group ($n = 30$) or a comparison group ($n = 30$). All participants evaluated documents from six Web sites related to the Atkins diet. In addition, the experimental group received a 3-page handout explaining SEEK (SEEK Declarative Materials) and filled out a SEEK Evaluation Template for each source related to the Atkins diet. The experimental group ranked and justified the credibility of each text. They received their ranking sheets back
with the rankings that “hypothetical” experts would have made, and then responded to questions about how their responses differed from those of experts. The comparison group was asked to rank and justify the credibility of each text, but did not get feedback on their selection.

The volcano inquiry task was used as a transfer measure for both groups. Participants read the seven edited Internet texts about volcanoes. Participants completed two measures of content learning: 1) an essay (“What caused the eruption of Mt. St. Helens?”), and 2) a 30-item volcano concept recognition test, a sentence verification instrument. Participants then completed two measures of evaluation: 1) ranking and justifying the credibility of each text and 2) evaluating an engineered “student” essay.

The experimental group who had received SEEK instruction preformed better than the comparison group on the content learning and evaluation measures. Twelve participants from the SEEK group and six from the control group explicitly mentioned corroboration as a justification for their evaluation. Essays were coded based on four categories assessing the presence of three causes of eruptions: Type 0 (Superficial models), Type 1 (Local models), Type 2 (Mixed models), and Type 3 (Integrated models). Participants in the SEEK group produced more Type 3 essays, while the majority of control group essays were categorized as Type 1. The SEEK group produced more sophisticated essays (Wiley et al.).

*Explicit instructional intervention: Using evaluative heuristics.* The previous three types of instructional approaches—SA, evaluation guides, and SEEK—have helped students identify more source information. However, they produced, at best, modest improvements in students’ use of evaluative heuristics. Nokes et al. (2007) present the
only instructional intervention to focus specifically on evaluative heuristics. After the instructional intervention, high school students demonstrated greater use of evaluative heuristics (the sourcing and corroboration heuristic, in particular) in their writing.

Nokes et al. (2007) conducted a quasi-experimental study to evaluate an instructional intervention designed to improve high school students’ use of Wineburg’s (1991) heuristics. Eight high school history classes from two schools were randomly assigned to one of four interventions: 1) traditional Textbook-Content instruction; 2) traditional Textbook-Heuristic instruction; 3) Multiple Texts-Content instruction, and 4) Multiple Texts-Heuristic instruction. The study was conducted with 11th grade students (N = 246) enrolled in mainstream history courses as part of a 15-day unit on United States history in the 1920s and 1930s. Classroom observations were conducted to ensure fidelity of implementation. Heuristic use (sourcing, corroborating, contextualizing, & using documents as evidence) was measured by a 3-stage assessment. Students read multiple documents focusing on a single historical event and critiqued a picture based on the document information. Secondly, students wrote a 200-word essay in which they took on the role of historian and explained whether or not the picture presented an accurate portrayal. Finally, they responded to four open-ended questions about the trustworthiness and usefulness of the documents. Based on the work of Britt and Aglinskas (2002), the researchers constructed a coding scheme to identify instances of heuristic use. Students in all conditions scored well on using documents as evidence (the fourth heuristic) during the pretest.

Using multiple texts enhanced content learning. Because a mixed-model ANCOVA revealed no teacher-within-intervention effect, students became the unit of
analysis. An ANCOVA conducted on the content posttest showed a significant intervention effect, $F(3, 213) = 21.93, p < .001$. A Tukey’s HSD showed students in the Multiple Texts-Content condition scored significantly higher than students in all other groups ($p < .01$). Students in the Multiple Texts-Heuristics conditions performed better than those in the Textbook-Content condition and significantly better than those in the Textbook-Heuristics condition ($p < .02$). In short, analysis of scores from both multiple texts conditions suggest that using multiple texts enhances content learning.

Instruction in evaluative heuristics appears to have helped students use sourcing and corroboration heuristics. Nokes et al. (2007) only reported heuristic use in the essays. Sourcing was the most commonly used heuristic. Seventy percent of the students used it for an average of three times each. There was a significant effect for sourcing, $F(3, 206) = 16.35, p < .001$. A Tukey’s HSD showed students in the Multiple Texts-Heuristics conditions scored significantly higher than all other groups ($p < .001$). There was also a significant effect for corroboration, $F(3, 205) = 10.02, p < .001$. Tukey’s HSD showed students in the Multiple Texts-Heuristics scored significantly higher than the two Textbook conditions ($p < .01$). The Multiple Texts-Content group scored significantly better than the Textbook-Content group on a measure of corroboration use in writing. No analyses were possible for contextualization because only seven percent of students used it. Because many students cited documents in the pretest and the posttest regardless of condition, there was no significant difference between the groups.

Students in the Multiple Text-Heuristic condition and the Multiple Texts-Content condition increased sourcing and corroboration heuristic use from pretest to posttest,
suggesting that both conditions should be present in future studies, as they are in the study.

Summary of Instructional Intervention Research Findings

All six instructional interventions transmit the expectation for evaluating source information to students (Britt & Aglinskas; Calkins & Kelley; Gardner et al.; Nokes et al.; Wiley et al.). Each instructional intervention utilizes different instructional methods to make explicit the type of source information the student should identify. Britt and Aglinskas advocate SA with predetermined slots into which students paste source information. Gardner et al. utilize an evaluation guide, the Internet Evaluation List. Calkins and Kelley present two evaluation guides: the CARS evaluation guide and the Historical Source Evaluation Worksheet. Wiley et al. present the SEEK template. Nokes et al. employed explicit instruction in evaluative heuristics. Therefore, this study included direct instruction to alert students to the expectation that readers evaluate source information in academic contexts at the university.

All six instructional interventions offer students opportunities to gain procedural knowledge about identifying source information. Students are asked to complete sourcing templates in SA (Britt & Aglinskas), the Internet Evaluation List (Gardner et al.), the CARS evaluation guide (Calkins & Kelley), the Historical Source Evaluation Worksheet (Calkins & Kelley), and the SEEK template (Wiley et al.) which force them to identify source information that should be evaluated. Some approaches incorporate evaluation opportunities: the Internet Evaluation List (Gardner et al.), the CARS evaluation guide (Calkins & Kelley), the Historical Source Evaluation Worksheet (Calkins & Kelley), and the SEEK template (Wiley et al.). However, only Nokes et al.
provide explicit instruction in the procedures for using evaluative heuristics. Since this explicit evaluative heuristic instruction led to increased use of the sourcing and corroboration heuristics that experienced academic readers use, this study included explicit evaluative heuristic instruction.

Three instructional approaches provide information on why readers should evaluate source information (Calkins & Kelley; Nokes et al.; Wiley et al.). Explicit instruction that explains the components of the CARS strategy and how to read research literature provide students with information about why evaluating sources is important academic behavior contributed to increases in contextualizing. Similarly, the SEEK declarative materials provide some information on why evaluating sources is important, which lead to increased corroboration (Wiley et al.). The reason to evaluate source information is best conveyed by offering instruction in the evaluative heuristics, as evidenced by increases in the use of both sourcing and corroboration (Nokes et al.). Therefore, this study included direct instruction in why to use each of the evaluative heuristics.

Although all six instructional interventions offer procedural knowledge of sourcing, only four scaffold students’ acquisition of this knowledge: Calkins and Kelley provide explicit evaluation guides and instructor feedback at strategic points in the learning process during two interventions; Wiley et al. provide expert feedback on students’ text credibility rankings; and Nokes et al. provide step-by-step explicit instruction of each evaluative heuristic. Each of these approaches increased students’ use of evaluative heuristics. In addition, students improved in their content knowledge (Nokes et al.; Wiley et al.) and in the sophistication of their writing (Nokes et al.; Wiley et al.).
et al.). Therefore, the intervention included careful scaffolding of the procedural knowledge related to evaluating sources, as well as instructor feedback during instruction.

Research suggests that explicit instruction in identifying source information leads to the inclusion of more source information by students. In two studies, high school students exposed to SA increased the number of references to sources (Britt & Aglinskas). Calkins and Kelley found that university sophomores enrolled in a psychology course who received explicit instruction in using the CARS strategy to evaluate sources tended to include more sources in their papers and to use more library resources, increasing the likelihood of finding credible sources. The most significant gains were achieved by Nokes et al. as the students receiving explicit instruction in using evaluative heuristics demonstrated the most gains in using evaluative heuristics. In order to encourage students to integrate information from multiple sources in their writing, this study included explicit instruction in identifying sources.

In two studies, explicit instruction in identifying source information led to improved content learning. University students who experienced the SEEK strategy for evaluating sources recognized more concepts related to the reading topic than the comparison group (Wiley et al.). The SEEK group also tended to write more sophisticated essays that integrated causal elements presented in multiple texts (Wiley et al.). High school students who received direct instruction with multiple texts (the Multiple Texts-Content and Multiple Texts-Heuristics groups) demonstrated enhanced content learning (Nokes et al.). However, the findings indicate that it might be multiple text instruction and not necessarily instruction in evaluative heuristics that accounted for
the increase. In order to help students gain academic literacy skills that will help them learn content across their undergraduate education, the current study included explicit instruction in identifying sources.

Two studies indicate that explicit instruction in identifying source information contributed to more sophisticated writing. Britt and Aglinskas found that students who experienced SA included more causal connectors and were judged to be of better quality by an independent grader. University students who experienced the SEEK strategy for evaluating sources wrote more sophisticated essays that accounted for multiple causal factors (Wiley et al.).

Two studies suggest that exposure to explicit instruction in identifying source information contributes to increased use of the sourcing heuristic. Although the authors only offered informal impressions of the effect of scaffolding explicit sourcing instruction involving worksheets and instructor feedback, they note that students increase their use of the sourcing heuristic (Calkins & Kelley). The findings for Nokes et al. are more concrete. Students who experienced explicit instruction in using the evaluative heuristics with multiple texts employed the sourcing heuristic more frequently in their essays.

Research suggests that explicit instruction in sourcing produced an even larger effect on students’ use of the corroboration heuristic. University freshmen were explicitly assigned to use the corroboration heuristic to evaluate a website (Calkins & Kelley). Although Calkins and Kelley only offered informal impressions of the effect of scaffolding explicit corroboration instruction involving worksheets and instructor feedback, they note that students increase their use of the corroboration heuristic.
Undergraduate who had experienced SEEK template instruction were twice as likely to utilize the corroboration heuristic to justify the credibility of a source than students who had just read the documents on the Atkins diet (Wiley et al., 2009). High school students who experienced explicit instruction in using the evaluative heuristics with multiple texts were significantly more likely to employ the corroboration heuristic in their essays than students who had experienced the other conditions (Nokes et al.).

Although Calkins and Kelley only offered informal impressions of the effect of scaffolding explicit sourcing instruction involving the CARS evaluation guide and instruction in reading research articles, they note that students exhibited increased contextualizing of sources. This is only one aspect of the contextualization heuristic. However, in comparison to the absence or minimal use of contextualization reported in other research (Britt & Aglinskas; Nokes et al.; Wiley et al.), this suggests that instruction in evaluating sources combined with evaluative reading instruction can improve undergraduates’ use of the contextualization heuristic.

Prior research indicates that explicit instruction in using evaluative heuristics leads to increased use of evaluative heuristics—the strategies that experienced academic readers use to resolve inconsistencies. Calkins and Kelley found that direct instruction in evaluating sources supported by explicit instruction in analytic reading leads to gains in evaluative heuristic use. Therefore, this study provided explicit instruction in using evaluative heuristics in order to help students acquire the academic literacy skills that experienced readers use to be successful at the university.
Previous Instructional Interventions Failed to Incorporate Conditional Knowledge

Researchers agree that inexperienced readers—like developmental-level freshmen—need instruction to build both procedural and conditional knowledge (e.g., Garner & Reis, 1981; Paris & Jacobs, 1984; Scardamalia & Bereiter, 1991). However, none of the previous instructional interventions for evaluating sources has incorporated building students’ conditional knowledge about when to use each evaluative heuristic. Therefore, this study attempted to fill the gap left by previous instructional interventions by offering instruction in evaluative heuristics that encompasses declarative, procedural, and conditional knowledge.

Methodological Considerations for Measuring Evaluative Heuristic Use

Evaluative Heuristic Use in Writing

Wineburg’s (1991) landmark study into novices and experts behavior while reading about a controversy in history used several short documents, which presented multiple perspectives on the Battle of Lexington. These eight documents proved to be rich material for investigating the approaches subjects used during think aloud procedures because there is a clear historical controversy over who started the battle, as well as numerous inconsistencies across the documents. Three of the documents are told from the colonists’ perspective, four from the British perspective, and the last document is an American textbook excerpt. Because the Battle of Lexington document set provides a clear controversy on a topic that most readers of American history textbooks present as a decided issue of fact, it helps students see that knowledge is not in the facts, but in navigating multiple perspectives. In this study, one text was removed from the document set in order to save time during administration. Seven of the original eight documents
about the Battle of Lexington were used because they present a clear central controversy, numerous inconsistencies across documents, and a case in which knowledge previously presented as fact is revealed as only a well-supported interpretation of available information.

An essay-writing task has been used in several of the studies investigating multiple text usage. Argument writing tasks have been shown to increase students’ use of evaluative heuristics and topic understanding (Wiley & Voss, 1999; Wiley et al., 2009). In fact, some researchers have asked subjects just to “imagine” that they were going to write an essay (Bråten et al., 2009), while others have asked students to write an essay, but not scored it (Britt & Aglinskas, 2002). Even when the essay is not a measure, researchers have used writing an essay as a way to help participants conceptualize the type of evaluative, purposeful reading that comprises academic literacy. The purpose of the Decision Essay is to assess how many times students use the three evaluative heuristics in academic writing. Therefore this study used a Decision Essay in which subjects must decide what they think happened at the Battle of Lexington (Appendix B).

Multiple text studies have used writing an essay to decide a controversy to measure several outcomes, including types of claims. Undergraduates and graduate students wrote essays deciding a controversy presented in a document set about the Panama Canal (Rouet et al., 1996). These essays were coded for types of claims (full, restricted, or no claim), references to documents, and type of document cited to support type of statements. Essays in this study were not coded for type of claims, because all students are taught to make strong claims as part of the course curriculum in developmental-level academic literacy courses.
Researchers have also measured the number of connectors contained in each essay (Britt & Aglinskas, 2002). These connectors included causal (“because”) and temporal (“afterwards”) connectors. Subjects’ use of connectors will not be measured in this study because developmental-level freshmen learn to use grammatical connectors as part of the writing instruction in developmental-level courses.

Rouet et al. (1997) conducted a similar analysis of essays, but included instances of heuristics used. However, the categories for the three heuristics were somewhat general. For example the three categories for contextualization were positive connections, negative connections, and general references. Britt and Aglinskas (2002) recoded the data collected by Rouet et al. (1996) to present more specific subcategories for each heuristic. Seven subcategories were used for identifying instances of the sourcing heuristic in previously collected data. These seven categories are author position, author motivation, author participation, author evaluations, publication date, document type, and document evaluation.

However, the one study to provide explicit instruction in evaluative heuristics used an even more specific coding scheme (Nokes et al.). Subjects completed an essay task based on the Battle of Lexington and those essays were scored using a very concrete rubric that identified the component aspects of each heuristic (Nokes et al.). Based Wineburg (1991) and Britt and Aglinskas, the Heuristic Rubric and accompanying Heuristic Scoring Guide identify seven sourcing aspects: author’s position, author’s motivation, author’s participation, other evaluation of author, date of production, document type, and other evaluation of document. Five aspects of corroboration are identified: direct comparison, direct contrast, claim of uniqueness, claim of omission, and
other. The seven aspects of contextualization are based on Wineburg’s work (1991) including time or location awareness, culture or setting awareness, biographic awareness, historiographic awareness, linguistic awareness, analogy, and other. This scoring rubric and the accompanying scoring guide clearly delineate what constitutes evidence of evaluative heuristic use. Therefore, this study used the Evaluative Heuristics Scoring Rubric and the Evaluative Heuristics Scoring Guide adapted from the rubric and scoring guide designed by Nokes et al. to measure heuristic use (Appendix D).

**Evaluative Heuristic Use in Reading**

A justification of trustworthiness task has been used by several multiple text researchers (Rouet et al., 1996; Rouet et al., 1997; Wiley et al., 2009). Rouet et al. used similar measures for both usefulness and trustworthiness. These researchers coded the justifications and scored the accuracy of the ranking as compared to expert historians. A more current multiple texts study used the ranking of sources as a measure and an intervention strategy (Wiley et al.). After completing the ranking sheet for a set of readings, subjects received their instrument back along with expert rankings. Students were asked to compare their ratings with the experts and respond to four questions about why the experts made those decisions about credibility. However, university education is not about replicating expert knowledge, but helping students to think for themselves. Therefore, this study did not compare student rankings with expert rankings.

A ranking and justification instrument can provide information on how participants evaluate sources (see also, Britt & Aglinskas, 2002; Rouet et al., 1996; Rouet et al., 1997; Wiley et al., 2009). This task explicitly asks participants to explain their decisions about credibility. Participants may not need to include evaluative heuristic
justifications in an essay-writing task, so a justification task provides a more complete picture of evaluative heuristic use. Therefore, the Justify Trustworthiness task was included as a measure of evaluative heuristic use for reading in this study (Appendix B).

Summary of the Literature on Evaluating Sources

Developmental-level literacy courses should help students acquire the academic literacy skills that experienced academic reader and writers use to succeed at the university. Wineburg (1991) identified three evaluative heuristics that experienced readers used. Experienced academic readers use evaluative heuristics to resolve inconsistencies across texts. Experts in a variety of disciplines, including history (Wasson, 2001), psychology (Rouet et al., 1996), law (Strømsø et al., 2003), engineering (Jacobson, 2001), immunohematology (Smith et al., 1991) have been observed utilizing the evaluative heuristics to solve complex problems.

However, students use evaluative heuristics significantly less often than experienced academic readers (Wasson, 2001). Students rarely verify source information (Britt & Aglinskas, 2002; Metzger et al., 2003; Flanagin & Metzger, 2007; Wiley et al., 2009). When students do attempt to verify information from sources, they rely on superficial criteria (Bråten et al., 2009; Calkins & Kelley, 2007; Flanagin & Metzger, 2007; Twait, 2005). Evaluating sources is an important component of academic literacy at the post secondary level (Burton & Chadwick, 2000). In addition to using credible sources in papers, evaluating source information predicts comprehension (Bråten et al., 2009; Strømsø et al., 2010; Wiley et al., 2009). Undergraduates must be able to evaluate source information if they are to be successful, independent learners.
Because students often use superficial evaluation criteria, instruction in evaluative heuristics is essential to help them become more expert academic readers (Davis, 2003; Fister, 1992; Fitzgerald & Galloway, 2001; Grimes & Boening, 2001; Kolstø, 2001; Twait, 2005). However, explicit instruction in source evaluation is rarely provided (e.g., Devet, 2007; Wiley et al., 2009). When instruction is provided, it is often procedural rather than conditional.

When procedural instruction in source evaluation has been provided, students have demonstrated improved source evaluation (Britt & Aglinskas, 2002; Calkins & Kelly, 2007; Gardner et al., 2007; Wiley et al., 2009). Instruction that builds procedural knowledge of source evaluation helps students understand why they need to attend to source information, contributes to more sophisticated writing, and leads to improved content learning.

However, explicit instruction in Wineburg’s evaluative heuristics (1991, 1998), which form the basis for this study, ensure that students are analyzing, not just identifying, aspects of credibility. Explicit instruction in the evaluative heuristics leads to increased use of the strategies expert readers use to successfully resolve inconsistencies (Nokes et al., 2007).

Prior research provides potential measures for assessing use of evaluative heuristics. Prior researchers utilized historical document sets, decision essays, heuristic scoring guides and heuristic scoring rubrics to measure heuristic use in writing (e.g., Nokes et al., 2007). Prior researchers have utilized rank and justify tasks to measure evaluative heuristic use in reading (e.g., Wiley et al., 2009). Therefore, this study
adapted measures used by previous researcher to assess the effectiveness of the instructional intervention.

From both a theoretical perspective and a research perspective, there is evidence that flexible use of strategic knowledge extends across disciplines. Experts in varied disciplines utilize similar strategies (Ericsson et al., 1993; Ericsson & Kintsch, 1995; Jacobson, 2001; Scardamalia & Bereiter, 1991, Smith et al., 1991; Wineburg, 1991; Wineburg, 1998). The evaluative heuristics transcend discipline specific boundaries. Because of their evident utility, evaluative heuristics should be taught to developmental-level university students as a means of helping them progress from introductory to advanced knowledge acquisition and build the academic literacy skills that will ensure their success at the university. Therefore this study utilized direct instruction in evaluative heuristics to help developmental-level students develop critical academic literacy skills.
CHAPTER III
METHODOLOGY

This chapter presents the methodology of this study, including descriptions of the research design, setting and participants, protection of human subjects, instrumentation, procedures, data collection, and data analysis. This chapter ends with a summary.

The primary purpose of this pre-experimental study was to investigate the effectiveness of an explicit academic literacy intervention with one group of developmental-level freshmen from two intact sections of a developmental-level integrated reading and writing course (N = 31). Specifically, this study investigated the influence of providing explicit instruction in recognizing and resolving inconsistencies across multiple texts on participants’ ability to 1) recognize inconsistencies, 2) use evaluative heuristics to reconcile inconsistencies when writing, and 3) use evaluative heuristics to reconcile inconsistencies when reading.

This study attempted to answer the following research questions:

1. What is the effect of an explicit academic literacy instructional unit on the number of inconsistencies identified by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Listing Inconsistencies measure?

2. What is the effect of an explicit academic literacy instructional unit on developmental freshmen’s ability to use evaluative heuristics as writers as measured by differences between their pretest and posttest scores on the Decision Essay measure?
3. What is the effect of an explicit academic literacy instructional unit on developmental freshmen’s ability to use evaluative heuristics as readers as measured by differences between their pretest and posttest scores on the Justify Trustworthiness measure?

Research Design

This study used a pre-experimental one group pretest/posttest design to investigate the effectiveness of an explicit academic literacy instructional intervention in the context of a developmental-level integrated reading and writing classroom. Participants received explicit instruction in recognizing inconsistencies within and across multiple texts and using evaluative heuristics, like those that experienced academic readers use to resolve these inconsistencies. A pretest was administered, followed by the four-week instructional intervention. Participants then completed the posttest. Table 2 presents the variables and instruments for the pre-experimental study. Measures of five student background variables—age, gender, ethnicity, language background, and familiarity with the topic of the document set for the Multiple Text Tasks—were administered prior to the start of the intervention.

Three dependent variables—1) the number of inconsistencies identified, 2) the number of evaluative heuristics used in writing, and 3) the number of evaluative heuristics used in reading—were measured at the beginning and the end of the study. For an overview of data collection, see Table 3. Participants completed the Multiple Text Tasks which includes reading a seven document set concerning the Battle of Lexington, listing inconsistencies identified within the document set, writing a decision essay about which side fired the first shot at the Battle of Lexington, and ranking and justifying the
credibility of each document in the set. The Multiple Text Tasks was administered prior to the four-week instructional intervention in recognizing and resolving inconsistencies and re-administered postintervention.

Table 2

Variables and Measurement Instruments

<table>
<thead>
<tr>
<th>Background Variables, Pretest Only</th>
<th>Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic Information</td>
<td>Demographic Questionnaire</td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>Language Background</td>
<td></td>
</tr>
<tr>
<td>Prior Topic Knowledge</td>
<td>Topic Familiarity</td>
</tr>
<tr>
<td>Dependent Variables, Pretest &amp; Posttest</td>
<td>Instruments a</td>
</tr>
<tr>
<td>The Number of Inconsistencies Identified</td>
<td>Listing Inconsistencies</td>
</tr>
<tr>
<td>The Number of Evaluative Heuristics Used in Writing</td>
<td>Decision Essay</td>
</tr>
<tr>
<td>The Number of Evaluative Heuristics Used in Reading</td>
<td>Justify Trustworthiness</td>
</tr>
</tbody>
</table>

*a All three instruments are part of the Multiple Text Tasks that use the Battle of Lexington Document Set.

This study took place in two sections of a year-long, developmental-level integrated reading and writing course during the spring semester. Approximately 16 students were enrolled in each section of the course. Both sections were taught by the same instructor and formed one group (N = 31) for the purpose of data analysis. The instructor who has experience teaching this developmental-level integrated reading and writing course volunteered to participate in this research project. This instructor
administered the pretest, posttest, and instructional intervention in both sections of the course.

Table 3

**Data Collection**

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Data Source</th>
<th>Data Collection</th>
</tr>
</thead>
</table>
| What is the effect of an explicit academic literacy instructional unit on the number of inconsistencies identified by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Listing Inconsistencies measure? | Listing Inconsistencies • Total number of inconsistencies identified | **Pretest:** Number of inconsistencies listed for the Battle of Lexington Document Set  
**Posttest:** Number of inconsistencies listed for the Battle of Lexington document set |
| What is the effect of an explicit academic literacy instructional unit on the number of evaluative heuristics used in writing by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Decision Essay measure? | Decision Essay • Subscores for number of times each evaluative heuristic is used in writing  
○ Sourcing  
○ Corroboration  
○ Contextualization  
• Total number of evaluative heuristics used in writing | **Pretest:** Number of evaluative heuristics used in the Decision Essay for the Battle of Lexington document set  
**Posttest:** Number of evaluative heuristics used in the Decision Essay for the Battle of Lexington document set |
| What is the effect of an explicit academic literacy instructional unit on the number of evaluative heuristics used in reading by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Justify Trustworthiness measure? | Justify Trustworthiness • Subscores for number of times each evaluative heuristic is used in reading  
○ Sourcing  
○ Corroboration  
○ Contextualization  
• Total number of evaluative heuristics used in reading | **Pretest:** Number of evaluative heuristics used in the Justify Trustworthiness task for the Battle of Lexington document set  
**Posttest:** Number of evaluative heuristics used in the Justify Trustworthiness task for the Battle of Lexington document set |

**Setting and Participants**

**Setting**

This study took place on the main campus of a large, urban, public, four-year university during a five-week period. Entering freshmen took an English placement test, which assessed reading and writing skills prior to enrolling in classes. Their scores on that English placement test were used to place them in either of two versions of Freshman
Composition (a one-semester written composition course or a year-long developmental-level integrated reading and writing course) or Composition for Multilingual Students. Students who scored at or above 147 on the English placement test were enrolled in a one-semester composition course. Those students who scored between 120 and 146 were required to enroll in a year-long developmental integrated reading and writing course. Students who scored below 120 usually elect to take Composition for Multilingual Speakers. However, students who score in the high 120s may enroll in the year-long developmental integrated reading and writing course despite needing the additional support of a multilingual composition course.

This study took place during the spring semester in two sections of a year-long developmental integrated reading and writing course. The year-long version of developmental integrated reading and writing is a six-unit (3 units in the fall and 3 units in the spring) course designed to provide additional support to students as they develop university-level, analytical reading and writing skills. This course focuses on integrating reading and writing strategies and includes instruction on the writing process, reading strategies, grammar, vocabulary, and study strategies. Students usually write six to seven essays on various topics during the year-long course.

The university offers approximately sixty sections of this year-long course each year. Each section meets twice a week with each class meeting lasting 75 minutes. Enrollment in each section is limited to 18 students. This study took place in two sections ($N = 31$) offered during the spring semester. The same instructor taught both sections. These sections met in the afternoon on Mondays and Wednesdays in standard classrooms on the main campus. Standard classrooms are equipped with either
chalkboards or white boards. Although some classrooms are smart classrooms equipped
with LCD projectors, all instructors can check out computer carts with an LCD projector,
laptop, and overhead projector from a technology resource room.

Each section was comprised of the same students and instructor from the fall
semester. Because this is a year-long course, the students and the instructor remained
together for the full academic year. The goal of the first semester of this integrated
reading and writing course is introducing college level reading, writing, and critical
thinking with an emphasis on exploring different genres. In the fall, the course focused
on the role of literacy in higher education and included a variety of readings focused on
the topic of education. Students completed weekly summary and response papers on
readings related to the course topic of education. The instructor assigned three major
papers: a five-paragraph essay about the five-paragraph essay (an essay organization
form common to high school English instruction) and a critique of the form, a difficulty
paper on an excerpt from Freire’s Pedagogy of the Oppressed, and an essay on
standardized testing. Unlike the Difficulty Paper utilized in this study, the Freire
difficulty paper assignment was reading focused, only incorporated a single text, and
allowed students to select several questions to pursue in Parts 2, 3, and 4.

The instructor incorporated iLearn courseware, a course management system, into
her course. Each section had an iLearn website where out-of-class assignments were
posted, readings could be downloaded, assignments including essays could be submitted,
and asynchronous discussions could take place. For this study, out-of-class assignments
were posted to iLearn, readings were available to download, and the Difficulty Papers
were submitted electronically.


**Instructor**

This study took place in two sections of the course taught by an instructor who had volunteered to participate in the study. The researcher approached this instructor about participating in the study because she had expressed interest in the use of evaluative heuristics even though she had no prior experience with using evaluative heuristics instructionally.

This instructor’s preparation and experience teaching reading and writing at the university level were consistent with the majority of the faculty in this university’s composition department. The instructor had earned a Master of Arts degree in Teaching Composition and had completed some coursework for the certificate in teaching postsecondary reading from a graduate program at this institution. This instructor has taught developmental-level courses for four years at this institution and was classified as a part-time lecturer.

She has experience incorporating difficulty paper assignments into her curriculum and used a single-text version of the difficulty paper in the fall with this group of students. Her curriculum focused on working through multiple challenging university-level texts, including research articles, making it a good fit for this study. The researcher scheduled one orientation meeting and two training sessions with the instructor. A discussion of the training she received is provided later in this chapter. The instructor’s letter of permission can be found in Appendix K.

**Participants**

Students eligible to enroll at this institution usually represent the top third of high school graduates in the state with a mean high school grade-point average of 3.16 (CSU,
Sixty-one percent of the freshman class is female (CSU, 2013). According to CSU (2013), the student population is racially and ethnically diverse, composed of American Indian (less than 1%), African American (5.4%), Asian American (18.9%), Filipino (9.3%), Mexican American (22.4%), Other Latino (8.9%), White/Non-Latino (22.6%), Pacific Islander (less than 1%), Two or more races (6.4%), Unknown (1.8%), Non-resident (3.7%).

At this campus, approximately sixteen hundred first-time freshmen (45.8% of incoming students) are required to take a developmental reading and writing course (CSU, 2010). These students are eligible to enroll in general education courses, but many are also enrolled in remedial mathematics courses. These students are able to decode texts and comprehend texts at the literal level, but struggle with inferential and evaluative reading skills. As readers, students enrolled in this course tend to struggle with textual analysis, tracing detailed arguments, and evaluating the credibility of information. Students enrolled in this course tend to struggle as writers with focus, organization, use of textual evidence, depth of analysis, and sentence variety. All students are taught to compose expository essays that range from three to ten pages in length and to read analytically.

The population in the year-long developmental courses is representative of the campus population with respect to many demographic characteristics, including gender distribution. However, almost all students enrolled in this course are first-time freshmen, and therefore are between 17 and 19 years of age.

Data on background variables were collected via the Demographic Questionnaire to demonstrate that the study participants from the two sections making up this sample
were representative of the population enrolled across all sections of the year-long course. Although the Demographic Questionnaire will be discussed in more detail in the Instrumentation section, the information it yielded about the background characteristics of this sample will be discussed here.

Thirty-three of the 34 students enrolled in these two sections completed the Demographic Questionnaire (97% response rate). In these two sections, 58% of the respondents were 18 years old at the start of the study \((n = 19)\), 39% were 19 years old \((n = 13)\), and 3% were twenty years olds \((n = 1)\). Sixty-four percent of the respondents were female \((n = 21)\), while 36% were male \((n = 12)\). Respondents identified as African American (12%), Asian American (9%), Filipino (33%), Mexican American (21%), Other Latino (3%), White/Non-Latino (12%), and Other (9%), suggesting that students identifying as African American and Filipino were slightly overrepresented in this sample.

Respondents were equally distributed across three categories of language background: Native Monolingual Students (39%), L1 Bilingual Students (27%), and Generation 1.5 Students (30%). Native Monolingual students speak only English. L1 Bilingual means those students are bilingual, but that English is their primary language. Generation 1.5 refers to students whose school experiences were conducted primarily in English, but whose home language is a language other than English. One student (3%) was categorized as L2 Bilingual, which means that the student is bilingual, but that the student’s primary language is other than English. These categories are discussed in more detail under the Demographic Questionnaire in the Instrumentation section.
Data on the final background variable, prior knowledge of the Multiple Text Tasks topic (Battle of Lexington), was collected via the Topic Familiarity measure during the pretest. The scoring range, mean, and standard deviation for the Topic Familiarity measure are presented in Table 4. The scale included 1- Not at all Familiar; 2 – Somewhat Familiar; 3 – Familiar, and 4 – Very Familiar. The mean score for the first item for which participants were asked to rate their familiarity with the topic was 1.71 (SD = 0.53). As expected, most participants were somewhat familiar with the Battle of Lexington. Only one participant indicated familiarity with the topic by selecting 3 – Familiar.

Table 4

Mean and Standard Deviation for Topic Familiarity

<table>
<thead>
<tr>
<th>Item</th>
<th>Scoring Range</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate Familiarity</td>
<td>1 - 4</td>
<td>1.71</td>
<td>0.53</td>
</tr>
<tr>
<td>Indicate Number of Times Studied</td>
<td>1 - 5</td>
<td>2.52</td>
<td>1.34</td>
</tr>
<tr>
<td>for a Class</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1 - 9</td>
<td>4.23</td>
<td>1.73</td>
</tr>
</tbody>
</table>

The second item asked participants to indicate the number of times the Battle of Lexington has been studied as part of a class (1 – Never to 5 – More than 3 times). Most participants had studied it between zero and two times, including eight responses for zero exposures. However, four participants selected 5 – More than 3 times. The mean number of times participants indicated they had studied the Battle of Lexington as part of a class was 2.52 (SD = 1.34). The mean total for the two-item Topic Familiarity survey
was 4.22 ($SD = 1.73$). Although some participants had been exposed to instruction about the Battle of Lexington on multiple occasions, overall, participants can best be described as somewhat familiar with this historical event.

The instructional intervention provided as part of this study was part of the teacher’s curricular plan for the spring semester. All students received course credit for their participation in the instructional intervention. Participation in the data collection phase of the study was voluntary. Three criteria would have excluded a participant from inclusion in the data analysis for this study. Students who were identified as English Language Learners and recommended for transfer to the Composition for Multilingual Students (CMS) program by the instructor would have been excluded because of weak English language skills. Students who had violated the attendance policy missing substantive amounts of instruction which would lead to withdrawal from the course would have been excluded from the data analysis. Lastly, as will be discussed in the next section (Protection of Human Subjects), students could opt out of data collection by not giving informed consent. Although these students engaged in the same tasks and received the same instruction, their data were not collected for inclusion in this study. Three of the 34 enrolled students declined to give consent and were excluded from the data analysis. Participant exclusion decisions were made prior to the administration of the pretest, except in the case of excessive absence during the intervention. Originally, if a student missed two or more class periods during the intervention, his or her data would have been excluded from the data analysis as he or she would not have had the benefit of the full instructional treatment. With such a small sample size, the researcher decided to
retain students who missed up to 3 of the 10 class periods allotted for the study. Therefore, no students were removed from the study for absenteeism.

**Protection of Human Subjects**

This study satisfies the standards for protection of human subjects. The researcher did not foresee any major risks to the participants. The researcher ensured that the fundamental rights of all subjects were preserved and adhered to the ethical standards of the American Psychological Association (2010). Approval from the Institutional Review Board at the University of San Francisco along with appropriate permission from the research site, including approval from that institution’s Institutional Review Board was obtained. Participation in this study was entirely voluntary; students could decline to participate in data collection procedures with no penalty. Informed consent to participate in this five-week study was obtained from each student (Appendix I). Receiving consent allowed the researcher to collect the following information from participants: demographic information, Topic Familiarity scores, Listing Inconsistencies scores, Decision Essay scores, and Justify Trustworthiness scores.

To maintain confidentiality, each participant was assigned a random number by a research assistant. The research assistant placed the list of participants and their assigned numbers in a sealed envelope. The researcher stored that list, sealed in the envelope, in a secure location. The researcher is the only person with access to that master list of participants and their assigned number. All testing materials, including completed assessments, notes pages, demographic questionnaires, and student work were coded with the assigned number. All data gathered has remained locked in a secure location that only the researcher has access to and participant confidentiality has been maintained.
Instrumentation

Five instruments were used in this study. Two instruments were used to collect information on student background variables: 1) a Demographic Questionnaire, and 2) a Topic Familiarity survey. The three dependent variables were 1) the number of identified inconsistencies, 2) the number of evaluative heuristics used in writing, and 3) the number of evaluative heuristics used in reading. In order to measure the dependent variables, three written instruments were administered as part of the Multiple Text Tasks using the Battle of Lexington document set (Appendix B) at pretest and posttest: 1) Listing Inconsistencies; 2) the Decision Essay; and 3) the Justify Trustworthiness task. Each instrument is described below.

Demographic Questionnaire

The Demographic Questionnaire (Appendix J) is an eight-item researcher-designed measure. The instrument was created to collect information on four background variables (age, gender, ethnicity, and language background) in order to describe the composition of the sample. This measure includes one item about age, one item about gender, one item about ethnicity, and five items about language background. There are five categories for age: 17 years, 18 years, 19 years, 20 years, and older than 20 years. Gender was presented as three categories (female, male, and other) to accommodate the diversity at a large, urban, public university.

Ethnicity was measured with the same categories that the institution that served as the research site uses to collect information on its student body. This large, urban public university uses nine categories for ethnicity: 1) African American, 2) American Indian or
Native Alaskan, 3) Asian American, 4) Filipino, 5) Mexican American/Mexican, 6) Other Latinos, 7) Pacific Islander, 8) White/Caucasian, and 9) Other.

Language background is a categorical variable measured with five items: 1) primary language for speaking; 2) primary language for reading and writing; 3) other language(s) spoken; 4) other languages for reading and writing; and 5) length of residency in the United States. The Language Background variable is designed to give information about a student’s language history that would influence his or her academic literacy (Table 5). There will be five categories: Native Speaker, Bilingual Student, Generation 1.5 Student, English Language Learning (ELL) Student, and International Student. These categories were designed in conjunction with a colleague who works in a program focused on multilingual students (M. Roberge, personal communication, February 12, 2010).

Native Speakers speak, read, and write in English as the primary or sole language of instruction. Therefore, they have usually developed an academic register for English. Generation 1.5 refers to a growing segment of college students who use English as their primary academic language, but who learned a different language as a child. They are fluent in American culture and have been educated within the American educational systems. Although they may speak another language, they probably have not acquired an academic register in that language. Bilingual, for the purposes of this study, refers only to individuals who speak and read in two languages, one of which must be English. L1 Bilingual means that those students are bilingual, but their primary language is English. L2 Bilingual means that those students are bilingual, but their primary language is a language other than English.
### Table 5

**Language Background Categories**

<table>
<thead>
<tr>
<th></th>
<th>Native Monolingual Students</th>
<th>L1 Bilingual Students</th>
<th>L2 Bilingual Students</th>
<th>Generation 1.5 Students</th>
<th>Generation 1.5 Students</th>
<th>ELL (^1) (CMS)</th>
<th>ELL (^2)</th>
<th>International Students</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary language for speaking</strong></td>
<td>E(^3)</td>
<td>E</td>
<td>O(^4)</td>
<td>E</td>
<td>O</td>
<td>E</td>
<td>E</td>
<td>O</td>
</tr>
<tr>
<td><strong>Primary language for reading &amp; writing</strong></td>
<td>E</td>
<td>E</td>
<td>O</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>O</td>
</tr>
<tr>
<td><strong>Other language for speaking</strong></td>
<td>N(^5)</td>
<td>O</td>
<td>E</td>
<td>O</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td><strong>Other language for reading &amp; writing</strong></td>
<td>N</td>
<td>O</td>
<td>E</td>
<td>N</td>
<td>N</td>
<td>O</td>
<td>N</td>
<td>E</td>
</tr>
<tr>
<td><strong>Length of residency</strong></td>
<td>11 or more years</td>
<td>6 - 10 years or longer</td>
<td>6 - 10 years or longer</td>
<td>6 - 10 years or longer</td>
<td>6 - 10 years or longer</td>
<td>2 - 10 years</td>
<td>5 years or less</td>
<td>2 years or less</td>
</tr>
</tbody>
</table>

\(^1\) More appropriately places in Composition for Multilingual Students (CSM) courses

\(^2\) Not enrolled at this institution

\(^3\) E – English

\(^4\) O – Other language

\(^5\) N – No second language

*English Language Learners* would have been excluded from the data analysis, because they should have been enrolled in an English as a Second Language or Composition for Multilingual Students course. Their low level of language skills would have significantly impacted their performance on the timed pretest and posttest. Participant exclusion decisions based on language skills would have been made prior to the administration of the pretest. However, no students were excluded based on English Language Learner status. *International students* differ from bilingual students only in that they only reside in the United States for a couple of years to study and thus are unfamiliar with the American educational system. They have most likely acquired an academic register in their first language, but may not have reached a similar level in
English. These categories illuminate not only the language skills of potential participants, but also the academic literacy resources they would be able to access.

Information on these variables was collected to describe the sample and was discussed in the Participants section above.

**Topic Familiarity**

The purpose of the Topic Familiarity instrument (Appendix B) was to assess participants’ familiarity with the topic of the document set regarding the Battle of Lexington that was administered at pretest and at posttest. Topic familiarity was measured in order to describe the sample.

Measures of prior topic knowledge used in previous multiple text research have indicated that participants do not have much prior topic knowledge on which to draw (Bråten et al., 2009; Rouet et al., 1997; Wineburg, 1991; Wolfe & Goldman, 2005). Although some researchers have decided not to use a prior knowledge measure (e.g., Britt & Aglinskas, 2002; Nokes et al., 2007), a measure of topic familiarity provides further evidence that the participants are homogenous.

This instrument assessed participants’ familiarity with the topic through two items. The first item was a Likert-like item: *Rate your familiarity with the topic. Not at all Familiar* was assigned 1 point; *Somewhat Familiar* was assigned 2 points; *Familiar* was assigned 3 points; and *Very Familiar* was assigned 4 points. The second item asked the participant to *indicate the number of times you have studied this topic for a class.* *Never* was assigned 1 point; *1 time* was assigned 2 points; *2 times* was assigned 3 points; *3 times* was assigned 4 points; *More than 3 times* was assigned 5 points. Participants circled their response.
The total points for both items were used to create a composite score of topic familiarity. Potential scores could range from 2 to 9 points. Based on previous research reported in the literature (e.g., Wineburg, 1991), it was expected that most participants would be only somewhat familiar with the Battle of Lexington, scoring from 2 to 4 total points on the Topic Familiarity Measure. The results from the Topic Familiarity measure were reported in the description of participants above.

**The Multiple Text Tasks**

Wineburg’s (1991) research on novices and experts reading about a controversy in history used several short documents that presented multiple perspectives on the Battle of Lexington. In this study an adapted set of seven documents focusing on the Battle of Lexington was used at pretest and posttest to measure participants’ ability to identify inconsistencies and use evaluative heuristics to reconcile those inconsistencies. Three of the documents present events from the colonists’ perspective, three from the British perspective, and the last document is a textbook excerpt. This adapted document set was used because it presented a clear central controversy, numerous inconsistencies across texts, and a case in which knowledge previously presented as fact is revealed as only a well-supported interpretation of available information. As part of the Multiple Text Tasks (Appendix B), students read the Battle of Lexington document set and then completed the three measures of the dependent variables: 1) Listing Inconsistencies, 2) the Decision Essay, and 3) the Justify Trustworthiness task.

All three measures of the dependent variables—Listing Inconsistencies, the Decision Essay, and the Justify Trustworthiness task—were co-scored by the researcher
and an instructor familiar with the developmental population in this study. The researcher then categorized each score as Low Use or High Use.

The researcher used scores reported in the literature and the scores of two expert readers to help calibrate the scoring categories. However, the scores reported in prior studies were often based on less rigorous definitions of the evaluative heuristics. In many studies, points were awarded for identifying information instead of evaluating information. Additionally, the expert readers who tested out the measures used in this study were university professors in the English department and, therefore, significantly more sophisticated in their approach to academic literacy than university freshman could be expected to be. Neither a review of the literature, nor the expert readers provide clear guidelines for the categorization. This study broke new ground, which is exciting, but does highlight the need for more research in order to normalize the scoring and categorization. Each instrument is described below.

*Listing Inconsistencies*

The purpose of the Listing Inconsistencies measure was to assess the number of inconsistencies across multiple texts that students identify (Appendix B). Research suggests that inexperienced readers do not notice inconsistencies when they read (Britt & Aglinskas, 2002; Otero & Kintsch, 1992; Wineburg, 1991), whereas experienced readers focus on inconsistencies (Wineburg, 1998).

Listing Inconsistencies was developed by the researcher to measure the number of inconsistencies developmental-level freshman identify after reading a topical document set (e.g., the Battle of Lexington). Samples of the types of inconsistencies present in the Battle of Lexington document set appear on Table 6. Procedures for constructing an
instrument were followed (Creswell, 2008). The purpose, audience, and format of the instrument were decided on after a thorough review of the literature and feedback from experts in university-level reading research. The listing format was selected because it required subjects to generate responses. However, subjects could use the notes they took while reading and studying the documents. Therefore, students were able to list the inconsistencies they noted while reading, rather than using a post facto reading strategy to ferret out information from the documents in response to the questions. Because subjects could use their notes, the measure captured information on what they noticed, not just what they remembered from their reading.

Table 6

*Sample Inconsistencies from the Battle of Lexington Document Set*

---

What was the British purpose for marching through Lexington?
- Document 5: secure two bridges
- Document 7: take possession of arms at Concord

Reference to the colonists
- Document 1: Lexington Company
- Document 3: peasants
- Document 5: rebels
- Document 6: colonists & peasants
- Document 7: embattled farmers, rebels, & patriots

Were drums sounded during the battle?
- Document 2: drumbeat alerted colonists to assemble on the Green
- Document 3: drumbeat during the battle
- Document 5: no drumbeat, but firing of guns & ringing bells alerted British

Only Document 7 brings up Paul Revere

---

Two content experts in postsecondary reading and one expert in research design reviewed the draft instrument. The first content expert held a certificate in teaching
postsecondary reading and a certificate in teaching composition along with a Master of 
Arts degree in English. She had taught developmental-level courses at the research site 
for three years. The second content expert held a Master of Arts degree in Teaching 
Composition and a certificate in teaching postsecondary reading. She had taught 
developmental-level integrated reading and writing courses at the research site and at 
other post-secondary institutions for ten years. Both content experts completed the 
Listing Inconsistencies task on Multiple Text Tasks with the Battle of Lexington 
document set and offered feedback on the content, format, and presentation of the 
instrument. In addition to completing the Listing Inconsistencies measure, the second 
content expert completed the Decision Essay and the Justify Trustworthiness task. 
Expert evaluation of representativeness and coverage of the task provided evidence of 
content validity.

The Listing Inconsistencies task was co-rated by two scorers. The researcher was 
one scorer and trained an instructor experienced with teaching integrated reading and 
writing at the postsecondary level to be the second scorer. The Scoring Guide for Listing 
Inconsistencies (Appendix C) and packet of six Listing Inconsistencies assessments was 
used for training and norming.

During the first scoring meeting, the researcher reviewed the purpose of the study, 
the Listing Inconsistencies measure, and The Scoring Guide for Listing Inconsistencies. 
The researcher presented and discussed three completed Listing Inconsistencies 
protocols. Then, both scorers scored three completed Listing Inconsistencies and 
compared their assessment. The scorers discussed their evaluation. Since the scorers 
were in agreement, they proceeded to score the rest of the protocols independently in sets
of six. They then compared and discussed the scores they assigned for each set of six protocols.

Scores for this measure are a single number value representing the number of inconsistencies listed on the assessment. Based on frequency counts, each participant was categorized as either High Use or Low Use. Table 7 presents the scoring ranges for all three measures of dependent variables.

Table 7

Scoring Ranges for Classification as Low Use and High Use

<table>
<thead>
<tr>
<th>Measure</th>
<th>Low Use</th>
<th>High Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listing Inconsistencies</td>
<td>0 - 3</td>
<td>4 or more</td>
</tr>
<tr>
<td>Decision Essay</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sourcing</td>
<td>0 - 1</td>
<td>2 or more</td>
</tr>
<tr>
<td>Corroboration</td>
<td>0 - 1</td>
<td>2 or more</td>
</tr>
<tr>
<td>Contextualization</td>
<td>0</td>
<td>1 or more</td>
</tr>
<tr>
<td>Total Score</td>
<td>0 - 2</td>
<td>3 or more</td>
</tr>
<tr>
<td>Justify Trustworthiness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sourcing</td>
<td>0 - 4</td>
<td>5 or more</td>
</tr>
<tr>
<td>Corroboration</td>
<td>0 - 2</td>
<td>3 or more</td>
</tr>
<tr>
<td>Contextualization</td>
<td>0</td>
<td>1 or more</td>
</tr>
<tr>
<td>Total Score</td>
<td>0 - 6</td>
<td>7 or more</td>
</tr>
</tbody>
</table>

The researcher identified 26 inconsistencies in the document set which are listed on the Scoring Guide for Listing Inconsistencies. Therefore, the ceiling score for this measure was set at 26. Both content experts hold master’s degrees and teach university-level courses, and therefore, can be considered expert academic readers/writers. Both content experts completed the Listing Inconsistencies assessment and found the time
limit to be a constraining factor in listing inconsistencies. Therefore, it was unlikely that any participant could actually list all 26 inconsistencies within the five minutes of allotted time.

The first content expert identified 5 inconsistencies on the Listing Inconsistencies assessment, while the second content expert noted 7. Because both expert readers have completed master’s degrees, the researcher deemed them to be much more skilled with academic literacy than someone successfully earning a bachelor’s degree, let alone university freshmen taking their first classes toward a bachelor’s degree. The researcher took the average expert score—six inconsistences identified—and set that as the threshold for expert reading. Findings from Otero and Kintsch (1992) were used to calibrate the expectations for inexperienced readers at three inconsistencies, slightly higher than the 40 percent recognized by high school students. Advanced learners, such as university students should be able to identify inconsistencies and can be expected to outperform high school students, but perform more modestly than expert readers. The researcher decided to use 4 as the cut score between Low Use and High Use for developmental-level freshmen. Participants were expected to identify at least 4 inconsistencies in order to be categorized as High Use. Participants identifying between 0 and 3 inconsistencies were categorized as Low Use.

**Decision Essay**

The purpose of the Decision Essay (Appendix B) is to assess how many times students use the evaluative heuristics in writing. The Decision Essay was administered as a pretest and a posttest as part of the Multiple Text Tasks using the Battle of Lexington document set. A raw score was calculated based on how many identifiable aspects of the
evaluative heuristics a student used in his or her Decision Essay. This section discusses the Decision Essay that students wrote explaining their decision about who fired first at the Battle of Lexington, as well as information on how that assessment was scored.

An essay-writing task has been used in several of the studies investigating multiple text usage. Argument writing tasks have been shown to increase students’ use of evaluative heuristics and topic understanding (Wiley & Voss, 1999; Wiley et al., 2009). An essay writing task, such as the Decision Essay, provides an authentic academic literacy task. The Decision Essay prompt was modified from the one used by Nokes et al. (2007). Students were asked to decide who fired first at the Battle of Lexington. Consistent with other multiple text studies, students had 24 minutes to read and take notes on the Battle of Lexington document set. Although the time limits varied in previous studies, participants in this study were allotted 30 minutes to write a Decision Essay, which is consistent with previous course instruction presenting writing as a process. Binder paper on which to write the Decision Essay was provided.

Both content experts reviewed the Decision Essay prompt. Both felt the use of the Battle of Lexington, an event that seemed settled, but for which a question remains, was an appropriate controversy for university-level reading and writing tasks. Both experts felt that the prompt was clear. The first content expert suggested that the researcher add the phrase “you have a chance to convince me of your decision” to give students a clearer sense of their audience. Both experts felt that this phrase at the end of the first paragraph of directions and the use of “Your essay should explain your decision” in paragraph two, made clear to participants that they needed to offer their perspective on the historical controversy. In addition, the second content expert appreciated the
inclusion of “Be sure to write about the documents” to clarify that this was not strictly an opinion piece, but an academic task that called for synthesizing information across sources.

Scoring: Evaluative Heuristics Rubric and Scoring Guide

A specific coding system previously used by Nokes et al. (2007) in a high school history intervention was adapted for use in this study to measure evaluative heuristic use. The Evaluative Heuristics Rubric and Evaluative Heuristics Scoring Guide are included in Appendix D. In the Nokes et al. multiple texts study, subjects completed an essay task based on the Battle of Lexington and those essays were scored using a concrete rubric that identified the component aspects of each of the three evaluative heuristics. The rubric and accompanying scoring guide were based on work by Wineburg (1991) and Britt and Aglinskas (2002) and identified eight aspects of sourcing: 1) author’s position, 2) author’s motivation, 3) author’s participation, 4) evaluation of author, 5) date of production, 6) document type, 7) evaluation of document, and 8) other. Five aspects of corroboration were identified: 1) direct comparison, 2) direct contrast, 3) claim of uniqueness, 4) claim of omission, and 5) other. The seven aspects of contextualization were based on Wineburg’s work (1991) including 1) time or location awareness, 2) culture or setting awareness, 3) biographic awareness, 4) historiographic awareness, 5) linguistic awareness, 6) analogy, and 7) other. Because this scoring rubric clearly delineated what constitutes evidence of the use of evaluative heuristics, it was adopted with some minor modifications, as the measure of evaluative heuristic use in the study (Appendix D).
Two content experts with experience teaching developmental-level college students and one expert in conducting educational research reviewed the instrument. Modifications were made to the aspect list for sourcing and contextualization. Table 8 lists the aspects for each of the evaluative heuristics and details the modifications to sourcing and contextualization. The name of one aspect of sourcing was changed from *author’s position* to *author’s credentials*, which expands the information that might be evaluated from career to include influences like educational background. The labels of the aspects were modified slightly for the sourcing heuristic. These modifications included removing the *other* aspect, replacing *evaluation of author* with *other evaluations of author*, and replacing *evaluation of documents* with *other evaluations of document*. These changes helped to clarify that all aspects of the sourcing heuristic should be evaluative and relate either to the author or the document.

The aspects of the corroboration heuristic were retained, but three changes to the aspects of the contextualization heuristic were instituted. The name of one category was modified slightly from *culture and setting awareness* to *cultural setting awareness* in order to refer to the surrounding culture and the emotional space, differentiating this category from the actual physical setting which is covered by the *time and location awareness* aspect.

The *biographic awareness* aspect originally referred to biographic information about the historical figures referred to in the texts (e.g., Benjamin Franklin’s political ambitions). This aspect was redefined to correspond to biographic information about the author of the documents and information regarding the time or manner of text production.
Table 8

**Aspects of Evaluative Heuristic Use**

<table>
<thead>
<tr>
<th>High School Intervention</th>
<th>This Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Evidence of the Sourcing Heuristic</strong></td>
<td></td>
</tr>
<tr>
<td>1) Author’s position</td>
<td>1) Author’s credentials</td>
</tr>
<tr>
<td>2) Author’s motivation</td>
<td>2) Author’s motivation</td>
</tr>
<tr>
<td>3) Author’s participation</td>
<td>3) Author’s participation</td>
</tr>
<tr>
<td>4) Evaluation of author</td>
<td>4) Other evaluations of author</td>
</tr>
<tr>
<td>5) Date of production</td>
<td>5) Date of production</td>
</tr>
<tr>
<td>6) Document type</td>
<td>6) Document type</td>
</tr>
<tr>
<td>7) Evaluation of document</td>
<td>7) Other evaluations of document</td>
</tr>
<tr>
<td>8) Other</td>
<td></td>
</tr>
</tbody>
</table>

**Evidence of the Corroboration Heuristic**

<table>
<thead>
<tr>
<th>1) Direct comparison</th>
<th>1) Direct comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>2) Direct contrast</td>
<td>2) Direct contrast</td>
</tr>
<tr>
<td>3) Claim of uniqueness</td>
<td>3) Claim of uniqueness</td>
</tr>
<tr>
<td>4) Claim of omission</td>
<td>4) Claim of omission</td>
</tr>
<tr>
<td>5) Other</td>
<td>5) Other</td>
</tr>
</tbody>
</table>

**Evidence of the Contextualization Heuristic**

<table>
<thead>
<tr>
<th>1) Time or location awareness</th>
<th>1) Time or location awareness</th>
</tr>
</thead>
<tbody>
<tr>
<td>2) Culture or setting awareness</td>
<td>2) Cultural setting awareness</td>
</tr>
<tr>
<td>3) Biographic awareness</td>
<td>3) Biographic awareness</td>
</tr>
<tr>
<td>4) Historiographic awareness</td>
<td>4) Historical awareness</td>
</tr>
<tr>
<td>5) Linguistic awareness</td>
<td>5) Linguistic awareness</td>
</tr>
<tr>
<td>6) Analogy</td>
<td>6) Analogy</td>
</tr>
<tr>
<td>7) Other</td>
<td>7) Other</td>
</tr>
</tbody>
</table>


The *historiographic awareness* aspect, a category specific to historical investigation was renamed *historical awareness*. The *historiographic awareness*
category focused on historical ways of knowing, whereas the new *historical awareness* category focuses on historical information pertaining to the time of the event that students might use to better understand the events (e.g., the Boston Tea Party). Since these heuristics were taught and used within a composition classroom, the focus was not on the historical ways of knowing specific to the discipline of history, but rather on strategies for making sense of events, which is applicable to various disciplines. The modified aspects better represent the types of knowledge that students bring to texts and ways they might best use that information in the developmental-level integrated reading and writing course and in their General Education coursework across disciplines.

The Evaluative Heuristics Scoring Rubric (Appendix D) was also developed by Nokes et al. (2007) and modified slightly by the researcher. The Evaluative Heuristic Scoring Rubric provides an organized space to record the number of appearances of each type of evaluative heuristic aspect in student writing, such as the Decision Essay, by marking a tally in the appropriate row. For example, if a student referred to how the author’s career as a soldier adds to his credibility, one tally mark was recorded in the *author’s credentials* row under sourcing. After the entire essay had been scored, the tally marks were added up for each evaluative heuristic. This instrument yielded four raw scores: number of instances of sourcing in writing, number of instances of corroboration in writing, and number of instances of contextualization in writing, and the total number of instances of evaluative heuristics in writing.

*Scoring Ranges for the Decision Essay*

Although there was no limit to the number of evaluative heuristics that a participant might use in their Decision Essay, both the 30-minute time limit and the
brevity of most essay responses limited the number of evaluative heuristics that could fit in the essay. It was unlikely that any participant could demonstrate more than 10 instances of evaluative heuristic use on this measure. Therefore, the ceiling for the total number of instances of evaluative heuristic use in writing on this measure was set at 10. The score from this measure is a single number value representing the number of instances of evaluative heuristic usage on the assessment. Based on frequency counts, participants were categorized as either High Use or Low Use. Scores from previous research were used to define High Use and Low Use categories for the Decision Essay.

*Sourcing heuristic.* One sample of high school students \((N = 6)\) was found to have used a total of 4 instances of the sourcing heuristic (Wasson, 2001). However, the definition of sourcing was broader than it is operationally defined for this study. A second sample of high school students averaged 1.5 instances of sourcing in their essays (Stahl et al., 1996). However, they had been asked to identify rather than evaluate information. In this study, participants who identified 2 or more instances of the sourcing heuristic were considered High Use. Although that number is lower than in previous research, the criteria for an instance of heuristic usage in this study was more rigorous, indicating evaluation, not just identification. If students demonstrated two instances of sourcing that suggests it was more than fortuitous. Intentional use of heuristics indicates skills consistent with experienced academic literacy and should be classified as High Use. Therefore, participants who demonstrated 2 or more instances of the sourcing heuristic on the Decision Essay were categorized as High Use, while participants demonstrating 0 or 1 instance of sourcing were categorized as Low Use.
**Corroboration heuristic.** The second content expert demonstrated 3 instances of corroboration in the Decision Essay. However, previous research indicates that graduate students demonstrated less than one instance of corroboration per decision essay (Rouet et al., 1997), while less experienced students have demonstrated fewer instances of corroboration. Half the high school students in Wasson (2001) exhibited use of corroboration. One-quarter of another sample of high school students were found to use corroboration in their essays (Stahl et al., 1996). The average use of corroboration was less than one instance of the corroboration heuristic. Based on these findings, participants were classified as Low Use if they used between 0 and 1 instance of corroboration. Participants who used 2 or more instances of the corroboration heuristic were categorized as High Use.

**Contextualization heuristic.** The second content expert exhibited 2 uses of contextualizing in the Decision Essay. Wasson (2001) who used a more generous coding scheme found a total of 9 instances of contextualization usage by high school students ($N = 6$), meaning an average of 1.5 uses even considering the very broad definition. In another study of heuristic use among high school students with an equally generous coding scheme, one-quarter of participants used contextualization in their essay (Stahl et al., 1996). Based on these and other evaluative heuristic studies which reported that contextualization usage by inexperienced readers was rare (e.g., Nokes et al., 2007; Wiley et al., 2009), the threshold for High Use was set at 1 instance of contextualization. Even one use of contextualization separates participants from the inexperienced readers/writers. In this study, participants who demonstrated 0 instances of
contextualization were categorized as Low Use, while participants who demonstrated 1 or more uses of the contextualization heuristic were categorized at High Use.

*Total scores for evaluative heuristic use.* Despite the low scores reported in previous studies, the researcher decided that three instances of evaluative heuristic use would demonstrate skill with evaluative heuristics, especially in light of the task not explicitly asking for or requiring evaluative heuristic use. The researcher decided that one use of evaluative heuristics could be incidental, while two uses of evaluative heuristics could be seen as purposeful use, but not consistent with experienced use. Participants who had scores ranging from 0 to 2 for total evaluative heuristic use on the Decision Essay were categorized as Low Use. Participants who demonstrated 3 or more instances of evaluative heuristic use on the Decision Essay were categorized as High Use.

*Training and Scoring Procedure*

During the third training meeting for the second scorer, the researcher reviewed the Decision Essay measure and the Evaluative Heuristic Scoring Rubric. The second scorer read three Decision Essays that had already been marked by the researcher. The scorers discussed the evidence of evaluative heuristic use. Then, the scorers scored three Decision Essays independently and discussed their evaluation. The second scorer had a number of questions about what constituted evaluative heuristic use, especially how much analysis the reader was to infer. Instead of following the original scoring plan to read a set of six essays independently and then compare ratings, the researcher altered the plan. The scorers read one essay independently and then discussed. This pattern of co-rating continued for six essays until the second scorer felt comfortable continuing on her own through a set of six essays.
Although there were few disagreements, both scorers noted questions for discussion. Both scorers entered tallies for any instances that they felt met the criteria for evidence of heuristic use. However, both scorers also marked any instances that they were uncertain about with a question mark, making notations to prepare for discussion. In many cases, both scorers had marked the same instances to discuss. Most often the scorers were able to reach agreement about how to resolve each question.

*Justify Trustworthiness*

The Justify Trustworthiness task (Appendix B) was used to measure the number of evaluative heuristics used in reading. This measure was administered preintervention and five weeks later postintervention.

Several multiple text researchers have used similar measures to assess student ratings of usefulness and trustworthiness (Rouet et al., 1996; Rouet et al., 1997; Wiley et al., 2009). After consultation with two instructors with prior experience teaching developmental-level reading and writing, the researcher decided not to include a measure of usefulness. At the university-level, usefulness and credibility need to be conjoined. If the source is not credible, it should not be used in academic writing. Participants in this study were asked to rank the seven Battle of Lexington documents based on their trustworthiness (*1* most trustworthy to *7* least trustworthy). Previous researchers matched participant rankings with those of experts, but at the university level less emphasis is placed on reproducing expert knowledge and more on enacting personal judgments about knowledge. Therefore, in this study, the numerical rankings were not scored. The rankings served only as a way for the participants to organize their response to each document as they considered the credibility of each text.
After they had read and taken notes on the document set, completed the Listing Inconsistencies task, and written a Decision Essay, participants were given the source information for each of the seven documents. The source information, in the form of citations, was presented on a simple table (Appendix B). Participants were asked to rank the trustworthiness of each document (1- most trustworthy to 7- least trustworthy) and write one to two sentences justifying their ranking. Students had fifteen minutes to evaluate the seven sources and record their ranking and their justification of the credibility of each document. Although students did not have access to the original document set, they were able to use their notes. Providing just the source information instead of the full text of each document ensured that students were not spending their limited time re-reading the documents, were focused on the source rather than irrelevant features like the writer’s style, and were able to complete the task under tight time constraints.

Discipline experts in academic literacy at the university level reviewed the Justify Trustworthiness task. Both suggested reformatting. Originally, a model row of cells formatted like those on the table appeared below the directions. The cells were labeled with what the student should fill in each (e.g., Write your one to two sentence justification here.). Both content experts suggested cutting the model row and adding labels to the boxes in the chart students would actually fill in (i.e., Justification or Rank). These changes were made, clarifying how the chart functions and limiting visual clutter.

Scoring: Evaluative Heuristics Rubric and Scoring Guide

The aspects of evaluative heuristics that participants used as a basis for their justifications were measured with the Evaluative Heuristics Scoring Rubric (Appendix
D), as discussed in the Decision Essay scoring section above. Training for the second scorer included reviewing the Evaluative Heuristics Scoring Guide (Appendix D) and norming with six assessments.

**Scoring Ranges for the Justify Trustworthiness Task**

Scores from this measure are a single number value representing the number of instances of evaluative heuristic usage on the assessment. Based on frequency counts, participants were categorized High Use or Low Use. Scores from previous research and one of the content experts were used to define High Use and Low Use categories for the Justify Trustworthiness task.

Although there is no limit to how many evaluative heuristics participants could use in their justification, practical limitations suggest that participants could not apply more than 21 evaluative heuristics. Therefore, the ceiling was set at 21. The fifteen-minute time limit and the small recording space limited the amount of evidence of the evaluative heuristics that each participant could present. In a prior study concerning scientific inquiry that used a similar measure, less than 10% of the justifications corresponded to the sourcing heuristic, while less than 7% corresponded to the corroboration heuristic (Wiley et al., 2009). There was no clear match for the contextualization heuristic in Wiley et al.’s coding scheme.

The Justify Trustworthiness task explicitly asked participants to justify their document credibility rankings, so the scores for this section were expected to be higher than for the Decision Essay. However, as with the Decision Essay, previous research indicates that while experienced academic readers use the evaluative heuristics to assess trustworthiness, inexperienced readers rarely use the evaluative heuristics (Bråten et al.,
The second content expert completed the Justify Trustworthiness task. Her scores will be discussed below along with the few published studies that used similar justification tasks.

*Sourcing heuristic.* Previous research indicates that while experienced academic readers use the sourcing heuristic, inexperienced readers rarely do (e.g., Britt & Aglinskas, 2002). The second content expert was credited with using the sourcing heuristic 12 times, indicating an average of 1.7 uses of the sourcing heuristic on each of the seven documents. Graduate students, a group with less expertise, used sourcing 68% of the time when justifying trustworthiness (Rouet et al., 1997). Sourcing is the most commonly used heuristic, so the researcher expected to see it used on a task that explicitly asks participants to evaluate sources. Since there were seven documents to be evaluated, using the sourcing heuristic with five documents would mean that the participant had used the sourcing heuristic with roughly 70% of the documents. Therefore, participants who demonstrated 5 or more instances of the sourcing heuristic (even if using multiple times with one document) were categorized as High Use. In this study, participants were categorized as Low Use if they demonstrated between 0 and 4 instances of the sourcing heuristic.

*Corroboration heuristic.* As noted previously, the Listing Inconsistencies task might prompt the use of corroboration since noting inconsistencies directs participants’ attention to comparing information across documents. The format of this measure visually represents the proximity of documents as the source information for each has been collected on one chart. However, the second content expert presented evidence of using 2 instances of corroboration across the seven documents. Despite the low use of
the corroboration heuristic by an expert, the explicitness of the task and the priming effect created by the Multiple Text Tasks create more likelihood for use, even among less experienced readers. Therefore, in this study, participants who demonstrated the use of between 0 and 2 instances of the corroboration heuristic were categorized as Low Use. Participants who demonstrated 3 or more instances of the corroboration heuristic were categorized as High Use.

*Contextualization heuristic.* Contextualization is the least used evaluative heuristic among inexperienced readers (Wiley et al., 2009). On this measure, contextualization may be more challenging because the context has largely been removed; the chart contains only source information. The second content expert had noted two instances of contextualization. Therefore, any use of contextualization could be considered impressive. In this study, participants who demonstrated 0 instances of the contextualization heuristic were categorized as Low Use. Participants who demonstrated 1 or more instances of the contextualization heuristic were categorized as High Use.

*Total scores for evaluative heuristic use.* The Justify Trustworthiness task is more explicit than the Decision Essay. In addition, it allows for brief, discrete evaluations of the documents. Therefore, participants would be expected to score higher than on the Decision Essay. The second content expert demonstrated a total of fourteen instances of evaluative heuristic use. That is an average of 2 instances of evaluative heuristic use per document. Since the developmental-population in this study is not expected to be experts, they were expected to use evaluative heuristics an approximate average of once per document to be considered high use. Therefore, participants who demonstrated a total of 7 or more instances of the evaluative heuristic use were
categorized as High Use. In contrast, participants who demonstrated the use of between 0 and 6 total instances of the evaluative heuristic use were categorized as Low Use.

**Training and Scoring Procedure**

During the second training meeting, the researcher reviewed the Justify Trustworthiness measure and the Evaluative Heuristic Scoring Rubric. The researcher presented three completed Justify Trustworthiness tasks as samples. The scorers discussed the evidence of evaluative heuristic use. Then, the scorers scored three Justify Trustworthiness tests independently and discussed their evaluation. The second scorer had a number of questions about what constituted evaluative heuristic use, especially how much analysis the scorer was to infer and how to handle potential misreadings. For example, several students were unfamiliar with the term *pro tem* which was used to describe the author of the first document: *Joseph Warren, president pro tem of the Massachusetts Provincial Congress*. Some students omitted it, referring to Joseph Warren as the president. Other students misread it in various ways, with one student replacing it with the term “pro-team”. It was decided that an analytic discussion of any aspect of the evaluative heuristics based on a reasonable misreading would be given credit.

Instead of following the original scoring plan to score a set of six Justify Trustworthiness tasks independently and then compare ratings, the researcher altered the plan. The scorers scored one task independently and then discussed. This pattern continued for six tests until the second scorer felt comfortable continuing on her own through a set of six. Then, the scorers continued co-rating the responses.
Although there were few disagreements, both scorers noted many questions for
discussion. Both scorers entered tallies for any instances that they felt met the criteria for
evidence of evaluative heuristic use. However, both scorers also marked any instances
that they were uncertain about with a question mark, making notations to prepare for
discussion. In many cases, both scorers had marked the same instances to discuss. Most
often the scorers were able to reach agreement about how to resolve each question.

 Procedures

Administration of Instruments

Consent and Demographic Information

Consent Forms (Appendix I) were distributed to all students during the class
period prior to the beginning of the study. The instructor invited the researcher to attend
the class meeting prior to the intervention (Day 0) in order to introduce the research
project. The researcher discussed the purpose of the study and students had an
opportunity to ask question about the study and about conducting research. The
researcher explained that students may grant or withhold consent. Students could
withhold consent to participate in the study without repercussions. All students engaged
in the same tasks, which were part of the instructor’s curriculum and submitted the same
work to the instructor. However, only data from students who had given consent were
collected by the researcher. The researcher answered student questions and gave students
10 minutes to read the consent form. The researcher distributed the Demographic
Questionnaire (Appendix J). Students completed the Demographic Questionnaire
(approximately 5 minutes). The researcher collected the Consent Forms and the
Demographic Questionnaires.
**Pretest and Posttest**

During the pretest prior to the instructional intervention, students completed the Topic Familiarity measure. The dependent variables were measured at pretest and posttest with the Multiple Text Tasks (Appendix A) which included reading and studying the Battle of Lexington document set for 24 minutes, Listing Inconsistencies for 5 minutes, writing the Decision Essay for 30 minutes, and completing the Justify Trustworthiness task within 15 minutes.

The students received two packets of materials. One packet secured with a staple included the Topic Familiarity measure followed by the directions for the Multiple Text Tasks and the seven Battle of Lexington documents. The second packet, which was secured with a paperclip, included two Notes pages for recording notes while reading the document set, the Listing Inconsistencies assessment, the Decision Essay prompt, four sheets of binder paper on which to write the Decision Essay, and the Justify Trustworthiness task.

Students had 1 minute to respond to the two-item Topic Familiarity measure and record their name on the front of the packet. They then turned the page on their packet to read the directions for the Multiple Text Tasks. Each of the seven documents of the Battle of Lexington document set was printed on a separate sheet of paper and included the document number and source information. After the 24-minute reading and notetaking period had elapsed, the Battle of Lexington document set (which includes the Topic Familiarity measure) was collected. Students kept the Notes pages for use in completing the three measures of the dependent variables.
The Listing Inconsistencies assessment was available on the second page of the second packet, behind the detachable Notes pages. Students had 5 minutes to list as many inconsistencies as they noticed. After 5 minutes, students were asked to turn the page and read the Decision Essay prompt. Following the Decision Essay prompt were four sheets of binder paper on which to write the Decision Essay. At the end of 30 minutes, students turned the page to the Justify Trustworthiness assessment. At the end of 15 minutes, the entire second packet including the Notes page was collated with a paper clip and collected by the instructor.

The process was repeated at posttest with a clean second packet of assessments. However, students received their same packet of Battle of Lexington documents with the Topic Familiarity measure they completed at pretest on top.

Treatment Description

Overview of Instruction

The explicit academic literacy intervention in recognizing inconsistencies and using evaluative heuristics to resolve those inconsistencies occurred over a four-week period (see Table 9). Students discussed academic literacy expectations at the university, supported by the anticipation guide and a PowerPoint presentation. Instruction was provided in recognizing inconsistencies using two Difficulty Paper assignments and the Introducing the Evaluative Heuristics information packet. Instruction in using the evaluative heuristics to resolve inconsistencies was provided using two Difficulty Paper assignments, the Introducing the Evaluative Heuristics information packet, and activities that built conditional knowledge about how and why to use the evaluative heuristics. Students practiced these skills with four education-themed texts.
### Table 9

**Overview of the Instructional Unit**

<table>
<thead>
<tr>
<th>Day</th>
<th>Instructional Focus</th>
<th>Out of Class Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Consent Form &amp; Demographic Questionnaire</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Pretest (Multiple Text Tasks)</td>
<td></td>
</tr>
</tbody>
</table>
| 2   | Academic literacy expectations discussion  
    - Anticipation Guide  
    - PowerPoint Lecture  
    Introduction of the Difficulty Paper (DP) | Bring 1st pair of education readings (Gatto & Sizer)  
    Begin DP #1, Part 1 |
| 3   | Evaluative Heuristics (Sourcing & Corroboration) | Complete DP #1, Part 1 |
| 4   | Discussion of DP #1, Part 1  
    - Question Types for the DP  
    Discussion of 1st pair of texts (Gatto & Sizer) | Complete DP #1, Part 2 |
| 5   | Conditional Knowledge of Sourcing Activity  
    Discussion of Gatto & Sizer texts  
    Discussion of DP, Part 2 | Work on DP #1, Parts 3 & 4 |
| 6   | Evaluative Heuristics (Contextualization) | Complete DP #1  
    Bring Dalrymple article |
| 7   | DP #1 is due.  
    Contextualization Application Activity  
    Preview Dalrymple text | Begin DP #2 |
| 8   | Contextualization Practice: Dalrymple’s Message  
    Discussion of 2nd pair of texts  
    (Dalrymple & Lockhart) | Complete DP #2, Parts 1 & 2 |
| 9   | Feedback provided on DP #1  
    Evaluation of a model DP #1  
    Discussion of DP #2, Parts 1 & 2 | Complete DP #2 |
| 10  | Posttest (Multiple Text Tasks)  
    DP #2 is due. |                      |
The instructional intervention was sequenced to encourage students to move toward independence using the gradual release of responsibility model (Fisher & Frey, 2008; Pearson & Gallagher, 1983). Early in instruction, the teacher bore the most responsibility for the instructional material, using instructional strategies like modeling and direct instruction. The instructional sequence moved into shared responsibility, with guided practice. For example, while students were working on Difficulty Paper #1, the class engaged in guided practice with identifying inconsistencies. Near the end of the unit, instruction focused on helping students bear the most responsibility by providing opportunities for independent practice by completing the second Difficulty Paper, primarily outside of class. Table 10 outlines the instructional sequence for the Difficulty Paper assignment as it relates to the gradual release of responsibility model.

Table 10

Gradual Release of Responsibility in the Instructional Intervention

<table>
<thead>
<tr>
<th>Principle of Instructional Scaffolding</th>
<th>Instructional Tool</th>
<th>Instructional Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Modeling</td>
<td>Model Difficulty Paper</td>
<td>Teacher leads class discussion of model Difficulty Paper</td>
</tr>
<tr>
<td>Guided Practice</td>
<td>Difficulty Paper #1</td>
<td>Guided practice identifying inconsistencies and selecting appropriate strategies to resolve them</td>
</tr>
<tr>
<td>Independent Practice</td>
<td>Difficulty Paper #2</td>
<td>Students independently practice identifying difficulties and creating plans to resolve those difficulties</td>
</tr>
</tbody>
</table>
The instructional intervention was designed to allow students to progress from declarative to conditional knowledge (Paris et al., 1983). Early instruction was explicit and included completing charts defining each heuristic. Students then engaged in guided practice to gain procedural knowledge. The last week of the instructional intervention was focused on building conditional knowledge. Table 11 offers an example of how these levels of knowledge were embedded in the evaluative heuristics instruction.

Table 11

*Example of Levels of Knowledge in the Instructional Intervention*

<table>
<thead>
<tr>
<th>Level of Knowledge</th>
<th>Instructional Tool</th>
<th>Instructional Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declarative</td>
<td>Introduction to Evaluative Heuristics handout</td>
<td>Defining each evaluative heuristics</td>
</tr>
<tr>
<td>Procedural</td>
<td>Introduction to Evaluative Heuristics handout</td>
<td>Guided practice using evaluative heuristics in a reading</td>
</tr>
<tr>
<td>Conditional</td>
<td>Conditional Knowledge of Sourcing handout</td>
<td>Guided practice in deciding which heuristics to apply in different literacy situations</td>
</tr>
</tbody>
</table>

*Texts*

The four primary texts for the intervention were evaluated on the following criteria: length, topic, complexity of viewpoints, and potential for conflicting perspectives. All four texts focus on the topic of education, are appropriate for university readers, and represent a variety of perspectives. The citations for these four expository
articles can be found in Appendix L. “We Don’t Want No Education” by Dalrymple (1995) is a scathing critique of how individuals have failed to take responsibility for themselves, thus creating a cult of stupidity that should be addressed with a back-to-basics curriculum. “The Seven-Lesson Schoolteacher” by J. T. Gatto (1992) discusses the author’s experience initiating elementary school students to the culture of school at the expense of education. In “The Mathematician’s Lament,” Lockhart (2002) decries the current state of school mathematics education with its emphasis on abstract terminology and mindless repetition of already solved proofs. Lockhart suggests that mathematics education should teach creativity and problem-solving skills. In “What High School Can Be,” Sizer (2003) explores the underlying cause of school failure and offers reform options. All four texts offer a variety of metaphors for education. The selected expository texts were paired to highlight conflicts across texts (e.g., different views on the problems that schools face, different approaches to educational reform).

The first pair of texts included the Sizer and Gatto texts, which present a similar problem, but they differ in perspective (Students’ vs. Teacher’s), solution (explicit vs. general), and hopefulness. The second pair of readings was comprised of the Lockhart and Dalrymple texts, which identify a similar problem (lack of student engagement), but suggest contrasting solutions (discovery education vs. a back-to-basics approach).

The Difficulty Paper Assignment

Each pair of texts was assigned in conjunction with the Difficulty Paper assignment (Appendix G). The Difficulty Paper is a four-part, elaborated reading strategy for use by developmental-level university readers. It was originally conceptualized for use with a single text by Salvatori (1996) and scaffolded for this
population by Levinson (2003). Levinson’s Difficulty Paper helps students to focus on
difficulty as a productive space for intellectual growth by explicitly scaffolding their
approach to reading to match the steps that expert readers go through. In Part 1, students
record questions and points of interest that arise as they read. For Part 2, students each
select one difficulty (inconsistency across texts) and frame it as a question. They are also
asked to come up with a plan of action for answering their question. In Part 3, students
answer their question, resolving the difficulty. Finally, Part 4 is a reflection on the
process. This assignment takes students through the steps experienced academic readers
employ as they identify and attempt to resolve inconsistencies. Although Levinson’s
Difficulty Paper helped developmental-level students to recognize inconsistencies in a
single text, those benefits did not transfer to recognizing inconsistencies across multiple
texts (Fisher, 2006).

Because Levinson’s Difficulty Paper assignment was originally designed for use
with a single text, it was modified slightly for use with multiple texts. Although students
in this study engaged in the same process for Part 1, they completed two Part 1s—one for
each text in the pair. For Part 2, students were asked to decide on one difficulty or
question, but the question should engage both texts. No changes were made to Part 3 and
Part 4. A model Difficulty Paper assignment for multiple texts was presented and
discussed in class (Appendix M). The Difficulty Paper assignment referred to throughout
the Procedures section is this modified Difficulty Paper for multiple texts.

Using the Difficulty Paper assignment as part of the instructional intervention
accomplished four goals. The Difficulty Paper assignment integrated reading and writing
instruction into a single assignment. Secondly, Part 1 of the Difficulty Paper explicitly
focused students’ attention on recognizing inconsistencies. Thirdly, Part 2 scaffolded students’ growing skill with planning to resolve inconsistencies, just as more experienced readers do. Finally, the Difficulty Paper assignment provided authentic opportunities for using the evaluative heuristics to resolve inconsistencies across texts.

The instructional sequence for the Difficulty Paper supported the gradual release of responsibility as students became more proficient and independent (Table 10). The early instruction with the Difficulty Paper was carefully scaffolded and included the use of models and support for categorizing and selecting questions for Part 2. Students then shared the responsibility within the class, with assignments like giving feedback on one another’s drafts of Difficulty Paper #1. Difficulty Paper #2 provided an opportunity for independent practice with a second pair of articles.

*Evaluative Heuristic Instruction*

Evaluative heuristic instruction was based on prior intervention research. The intervention focused on the three evaluative heuristics that expert readers use to resolve inconsistencies within and across texts: sourcing, corroboration, and contextualization (Wineburg, 1991, 1998). Wiley et al. (2009) demonstrated the importance of being explicit about the evaluative heuristics, so a handout entitled Introduction to the Evaluative Heuristics (Appendix H) which details the aspects of each evaluative heuristic (e.g., *type of document* and *date of publication* aspects of sourcing) was distributed. On the student handout, four heuristics were listed. Identifying Difficulty was listed as the first heuristic. Although it is not one of the evaluative heuristics, Wineburg (1998) did find evidence of experienced readers engaging in specification of ignorance (identifying difficulty). Adding Identifying Difficulty as an evaluative heuristic in the handout
allowed for an integrated presentation to students of the skills they should acquire during this instructional intervention.

Nokes et al. (2007) found benefits to practicing with the evaluative heuristics, so in this study students were asked to use the evaluative heuristics during structured, in-class activities, including with the education-themed texts. Students completed activities from the Introduction to the Evaluative Heuristics handout, including Contextualization Practice with the Author’s Message (Appendix H), a contextualization application activity with the lyrics to “We Don’t Need No Education” (Appendix O), and the Conditional Knowledge of Sourcing Activity (Appendix P). In order to avoid simple training like that provided by Britt and Aglinskas (2002), the Difficulty Paper assignment linked evaluative heuristic use to university-level writing assignments, provided contextual need, and encouraged intellectual sophistication.

Teacher Training

The researcher met with the instructor for an orientation meeting and two training sessions. The orientation meeting took place during the winter break prior to the spring semester in which the study occurred. During the 30-minute orientation meeting, the researcher explained the purpose of the study and outlined the curriculum. The researcher left the instructor with a packet of information to read over before the first training meeting, including a day-by-day instructional outline and the four education-themed texts.

The first training session occurred early in the spring semester, approximately 6 weeks before the intervention began. During the first training session which lasted 90 minutes, the researcher took the instructor through the instructional materials for
academic literacy (including the PowerPoint), how to identify inconsistencies, using the Difficulty Paper for multiple texts, and the three evaluative heuristics. The instructor was already familiar with the difficulty paper as an instructional tool, but this was her first exposure to the Difficulty Paper designed for multiple texts. The differences were discussed and the researcher and the instructor looked at a model Difficulty Paper for multiple texts and reviewed it in light of the evaluation criteria. The same multi-page handout for students—Introduction to the Evaluative Heuristics—that details the three evaluative heuristics (sourcing, corroboration, & contextualization) was used as the basis for training the instructor. The researcher and the instructor reviewed the tables that listed each aspect of the evaluative heuristics and the researcher provided some examples of the aspects as a model for the instructor. The researcher provided day-by-day lesson plans and a packet of the instructional materials for the instructor to review in preparation for the second training session.

In the ten days between the first and second training sessions, the instructor reread all four education-themed texts, looking for examples of inconsistencies and filling in charts for the aspects of each of the three evaluative heuristics. During the second meeting, the instructor and the researcher discussed the inconsistencies and evaluative heuristics the instructor had identified and reviewed the materials and directions for administering the Multiple Text Tasks. This meeting focused on the materials used for teaching students about evaluative heuristics, which include guidelines for class discussion, a multi-page handout for students, Introduction to the Evaluative Heuristics, and activities encouraging the use of the evaluative heuristics with texts. The researcher had scheduled two hours for this second meeting in order to provide ample time to
discuss the inconsistencies the instructor identified and the instructor’s examples of the evaluative heuristics from the texts, as well as to address the instructor’s questions. However, only one hour was needed. The researcher provided class sets of materials (e.g., anticipation guide, the Difficulty Paper assignment sheet, pretest and posttest), and scoring guides (for listing inconsistencies & for the evaluative heuristics).

The researcher also met briefly with the instructor a few days before she administered the Multiple Text Tasks at the beginning of the study to answer questions.

**Instructional Intervention**

*Overview*

An overview of the unit plan (Table 9) for the developmental-level integrated reading and writing course that met twice a week for 75 minutes lays out the explicit academic literacy instructional intervention students experienced. During the first week students completed the pretest and discussed the expectations for academic literacy at the university. During the second and third week, instruction focused on recognizing and resolving inconsistencies within and across multiple texts, using the Difficulty Paper assignment, the Introduction to the Evaluative Heuristics handout, and activities that encouraged students to practice these skills, as students read the first pair of education-themed texts. Week four focused on additional practice, particularly through the completion of a second Difficulty Paper assignment with a second pair of education-themed readings. Week 5 ended with the posttest. Each day of the intervention is discussed below.
**Day 0**

The class meeting before the study began was designated as Day 0. On Day 0, the students read and signed Consent Forms and completed the Demographic Questionnaire.

**Week 1**

On Day 1, as detailed in the Pretest and Posttest section above, students completed the pretest. Students completed the Multiple Text Tasks for the Battle of Lexington document set which includes the Topic Familiarity measure, the Listing Inconsistencies assessment, the Decision Essay assessment, and the Justify Trustworthiness assessment.

On Day 2, students filled out an anticipation guide (Appendix F) prior to the instructor delivering a 20-minute mini-lecture about expectations for university-level academic literacy (Appendix E). The anticipation guide, a schema activation and interest building activity for the mini-lecture, asked students to respond to five true-false statements about expectations for university-level reading. Answers to the five items were presented as part of the mini-lecture. The instructor also introduced the Difficulty Paper assignment (Appendix G) and reviewed a model Difficulty Paper (Appendix M). For their out-of-class assignment, students were assigned to download the first pair of education-themed readings (Gatto & Sizer) from iLearn, bring the texts to class, and begin reading and working on Part 1 of the Difficulty Paper—Identifying Difficulty Across Multiple Texts.

**Week 2**

Day 3 was dedicated to introducing the sourcing and corroboration heuristics and practicing the use of each heuristic with the first pair of education-themed texts (Gatto &
Sizer). Students read through the first three pages of the Introduction to the Evaluative Heuristics handout (Appendix H) and discussed the aspects of sourcing chart. They evaluated the source features of the Gatto text and completed the corroboration chart in the handout. For their out-of-class assignment, students completed Part 1 of the Difficulty Paper.

On Day 4, the instructor facilitated a discussion of the articles, including modeling various types of difficulty. The instructor discussed five types of questions (see Appendix N for handout). Working with a partner, students decided on one difficulty from each partner’s Difficulty Paper, Part 1 to share with the class, framed it as a question, and decided what type of question it was. Students each wrote their partner’s difficulty on the board. The instructor chose several examples of different types of difficulty (inconsistencies), emphasizing those that occur as a result of multiple perspectives and would, therefore, be appropriate for use in Part 2 of the Difficulty Paper (Plan of Action). After reviewing types of difficulties and seeing models, students selected a question (difficulty) that occurs across multiple texts for Part 2 of the Difficulty Paper and formulated a plan of action for answering that question. The out-of-class assignment was to complete Difficulty Paper, Part 2.

**Week 3**

Day 5 included a review of the sourcing and corroboration heuristics, including the Conditional Knowledge of Sourcing activity (Appendix P) as students discussed the Gatto and Sizer articles. Students also shared Part 2 of the Difficulty Paper and got feedback from a classmate after the instructor reviewed the criteria and the model for Part
2 of the Difficulty Paper. The out-of-class assignment was to revise and submit Parts 1 and 2 of the Difficulty Paper to iLearn and to work on Parts 3 and 4.

Day 6 was focused on the introduction to and practice with the contextualizing heuristic using the first pair of education-themed articles. Students received feedback on Part 1 and Part 2 of the Difficulty Paper. The out-of-class assignment was to print the Dalrymple article from iLearn, bring it to class, and to complete all four parts of Difficulty Paper #1. Difficulty Paper #1 was due on Day 7.

Week 4

On Day 7, students submitted their Difficulty Paper #1. Students completed an activity focused on application of the contextualization heuristic (Appendix O). The instructor introduced the next two education-themed readings (Dalrymple and Lockhart). Students previewed Dalrymple text using the sourcing heuristic. For their out-of-class assignment students were to read the second pair of education-themed articles and begin Difficulty Paper #2. Students were asked to complete and bring the readings to the next class meeting and begin Parts 1 and 2 of Difficulty Paper #2.

On Day 8, students engaged in practice with contextualization around Dalrymple’s message. The out-of-class assignment was to complete and bring Part 1 and Part 2 of Difficulty Paper #2.

Week 5

On Day 9, students received feedback from the instructor on Difficulty Paper #1. Difficulty Paper #1 was not assigned a grade by the instructor until both Difficulty Paper assignments had been submitted (Day 10). The class read, evaluated, and discussed one student’s sample Difficulty Paper #1 selected by the researcher as a model.
(approximately 30 minutes). Students also had approximately 25 minutes to share any difficulties (inconsistencies) they identified in the second pair of education-themed readings. Students had about 20 minutes to offer feedback on Part 1 and Part 2 (Plan of Action) of the Difficulty Paper #2. Students completed Difficulty Paper #2 outside of class.

On Day 10, students submitted Difficulty Paper #2 and completed the posttest. The instructor evaluated Difficulty Paper #2 along with Difficulty Paper #1, which students could revise as needed after getting feedback from the instructor and their classmates. The instructor’s evaluation of Difficulty Paper #1 and Difficulty Paper #2 formed the grade students received for this unit of instruction. The posttest was conducted in accordance with the Administration of Instruments section that precedes this Instructional Intervention section.

Data Collection and Analysis

Data Collection Procedures

The researcher collected data from the participants on the day prior to the start of the study, designated Day 0, in the form of Demographic Questionnaires. Participant exclusion decisions were made prior to administration of the pretest, based on limited English proficiency, chronic absenteeism, and withholding of consent to participate. During the intervention, data from students who missed two or more classes were to have been excluded from the analysis because these students would not have been exposed to the full benefit of the instructional intervention. However, no students were excluded because of absenteeism. Data were also collected at the pretest on Day 1 of the study
from the Multiple Text Tasks and at posttest on Day 10 of the study from the Multiple Text Tasks. Table 12 provides an overview of data collection.

Table 12

Data Collection

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Data Source</th>
<th>Data Collection</th>
</tr>
</thead>
</table>
| What is the effect of an explicit academic literacy instructional unit on the number of inconsistencies identified by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Listing Inconsistencies measure? | List Inconsistencies • Total number of inconsistencies identified | Pretest: Number of inconsistencies listed for the Battle of Lexington Document Set  
Posttest: Number of inconsistencies listed for the Battle of Lexington document set |
| What is the effect of an explicit academic literacy instructional unit on the number of evaluative heuristics used in writing by developmental-freshmen as measured by differences between their pretest and posttest scores on the Decision Essay measure? | Decision Essay • Subscores for number of times each evaluative heuristic is used in writing:  
  o Sourcing  
  o Corroboration  
  o Contextualization  
  • Total number of evaluative heuristics used in writing | Pretest: Number of evaluative heuristics used in the Decision Essay for the Battle of Lexington document set  
Posttest: Number of evaluative heuristics used in the Decision Essay for the Battle of Lexington document set |
| What is the effect of an explicit academic literacy instructional unit on the number of evaluative heuristics used in writing by developmental-freshmen as measured by differences between their pretest and posttest scores on the Justify Trustworthiness measure? | Justify Trustworthiness task • Subscores for number of times each evaluative heuristic is used in reading:  
  o Sourcing  
  o Corroboration  
  o Contextualization  
  • Total number of evaluative heuristics used in reading | Pretest: Number of evaluative heuristics used in the Justify Trustworthiness task for the Battle of Lexington document set  
Posttest: Number of evaluative heuristics used in the Justify Trustworthiness task for the Battle of Lexington document set |

Data Analysis Procedures

This pre-experimental study attempted to answer the following research questions:

1. What is the effect of an explicit academic literacy instructional unit on the number of inconsistencies identified by developmental-level freshmen as
measured by differences between their pretest and posttest scores on the Listing Inconsistencies measure?

2. What is the effect of an explicit academic literacy instructional unit on the number of evaluative heuristics used in writing by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Decision Essay measure?

3. What is the effect of an explicit academic literacy instructional unit on the number of evaluative heuristics used in reading by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Justify Trustworthiness measure?

To answer the first research question, quantitative data analysis included both descriptive and inferential statistics. Descriptive statistics included frequency counts for the total score on the Listing Inconsistencies measure at preintervention and postintervention. In this study, the relatively small sample size \( N = 31 \) and the small score range on the three measures of the dependent variables suggested a non-continuous (non-normal) distribution. Thus, a non-parametric test, such as chi-squared test, was an appropriate data analysis strategy (Creswell, 2008). The McNemar test for significance of change is a chi-squared test for dependent samples involving nominal data (Hinkle, Wiersma, & Jurs, 1988). The McNemar test is appropriate for pre-post designs in which the same sample is categorized before and after an intervening treatment. Thus, inferential statistics included the McNemar test for significance of change (Hinkle et al., 1988) for the total score on the Listing Inconsistencies measure to analyze the significance of the differences between pretest and posttest categorization.
McNemar’s test was used to test the null hypothesis that there would be an equal number of changes in both directions. In other words, the expected frequency in cell A (the number of students categorized as Low Use at pretest but categorized as High Use at posttest) would be equal to the expected frequency of cell D (the number of students categorized as High Use at pretest but categorized as Low Use at posttest).

The critical value of chi-squared for one degree of freedom at \( \alpha = .05 \) is 3.841. The test statistic was calculated. To interpret the results, if the chi-squared value exceeds the critical value, the null hypothesis would be rejected. The conclusion would be that there were more changes one way than another. A visual examination of the data would tell which way the change went.

To answer the second research question, quantitative data analysis included both descriptive and inferential statistics. Descriptive statistics included frequency counts for the subscores and total score on the Decision Essay measure at preintervention and postintervention. Inferential statistics included the McNemar test for significance of change (Hinkle et al., 1988) for the subscores and total score on the Decision Essay measure to analyze the significance of the differences between pretest and posttest categorization.

To answer the third research question, quantitative data analysis included both descriptive and inferential statistics. Descriptive statistics included frequency counts for the subscores and total score on the Justify Trustworthiness measure at pre-intervention and post-intervention. Inferential statistics included the McNemar test for significance of change (Hinkle et al., 1988) for the subscores and total score on the Justify
Trustworthiness measure to analyze the significance of the differences between pretest and posttest categorization.

Quantitative data analysis included both descriptive and inferential statistics. Descriptive statistics included frequency counts. Inferential statistics included the McNemar test for significance of change (Hinkle et al., 1988). Frequency counts from three measures (Listing Inconsistencies, the Decision Essay, and the Justify Trustworthiness task) of the three dependent variables (number of inconsistencies identified, number of times each evaluative heuristic is used in writing, and number of times each evaluative heuristic is used in reading) were collected. Based on frequency counts, participants were categorized as either High Use or Low Use.

Qualifications of the Researcher

The researcher holds a Master of Arts degree in English with certificates in Teaching Composition, Teaching Postsecondary Reading, and Educational Therapy. She has taught courses in postsecondary reading and developmental literacy at the research site for sixteen years. She has taught teacher preparation courses at several postsecondary institutions and has nine years experience as a mentor teacher. She was one of a group of seven instructors to have developed the current developmental-level integrated reading and writing course objectives and course curriculum. She has training and experience in designing curriculum for developmental-level literacy courses and administering diagnostic tests to individuals and groups. The researcher designed the instructional intervention in this study. In this study, the researcher trained a second scorer and co-scored the assessments measuring each dependent variable along with the second scorer.
The second scorer was an instructor with experience in working with developmental-level university freshmen at the research site. The second scorer holds a Master’s degree in Teaching Composition, a TESOL certificate, and a certificate in teaching postsecondary reading. She has been teaching courses in postsecondary reading and composition for 10 years at the university and at community colleges.

Summary

The purpose of this pre-experimental, within subjects pretest-posttest design study was to investigate the effectiveness of an explicit academic literacy intervention in two intact sections of a developmental-level integrated reading and writing course. Specifically, this study investigated the influence of providing explicit instruction in recognizing and resolving inconsistencies across multiple texts on participants’ ability to 1) recognize inconsistencies, 2) use evaluative heuristics to reconcile inconsistencies when writing, and 3) use evaluative heuristics to reconcile inconsistencies when reading. The four-week academic literacy instructional intervention utilized an anticipation guide, direct instruction and explicit practice in identifying inconsistencies within and across multiple texts and resolving those inconsistencies using evaluative heuristics, and two Difficulty Paper assignments. Three measures were administered at pretest and posttest—Listing Inconsistencies, the Decision Essay, and the Justify Trustworthiness task—to detect potential changes in participants’ abilities to engage in the academic literacy behaviors of experienced university readers.
CHAPTER IV

RESULTS

The purpose of this pre-experimental study was to investigate the effectiveness of an explicit academic literacy intervention with one group of developmental-level freshmen from two intact sections of a developmental-level integrated reading and writing course at a large, urban public university \( (N = 31) \). This study investigated the influence of providing explicit instruction in recognizing and resolving inconsistencies across multiple texts on participants’ ability to 1) recognize inconsistencies, 2) use evaluative heuristics to reconcile inconsistencies when writing, and 3) use evaluative heuristics to reconcile inconsistencies when reading.

At the pretest, participants completed the Multiple Text Tasks which included reading a seven document set concerning the Battle of Lexington, listing inconsistencies identified within the document set, writing a decision essay about which side fired the first shot at the Battle of Lexington, and ranking and justifying the credibility of each document in the set. The three measures—Listing Inconsistencies, the Decision Essay, and the Justify Trustworthiness task—were designed by the researcher to measure students’ ability to recognize inconsistencies and resolve them using evaluative heuristics.

During the four-week instructional intervention, participants received explicit instruction in recognizing inconsistencies within and across multiple texts and using evaluative heuristics (sourcing, corroboration, and contextualization). The instructional intervention began with an anticipation guide for which students were given a series of statements about academic literacy in order to activate prior knowledge in preparation for
acquiring new information (Readence et al., 2004). A PowerPoint lecture and class discussion based on the concepts presented in the Anticipation Guide for Academic Literacy Expectations was used to explicitly provide additional information about expectations for academic literacy.

The Difficulty Paper, an elaborated reading strategy, was used to make the problem-solving process of experienced readers and writers explicit to developmental-level readers/writers. The four-part Difficulty Paper assignment asked students to 1) identify any difficulties (inconsistencies) they noticed; 2) select one inconsistency and create a plan for resolving that inconsistency; 3) deploy strategies; and 4) reflect on learning outcome and choice of strategy.

Direct instruction in the three evaluative heuristics—in the form of the Introduction to Evaluative Heuristics packet and in-class activities—was provided to help students understand the evaluative heuristics and learn to use them flexibly as expert academics do. Expert utilization of evaluative heuristics has been well documented (e.g., Jacobson, 2001; Rouet et al., 1996; Smith et al., 1991; Wineburg, 1991, 1998; Wasson, 1991). At the end of the intervention, participants again completed the Multiple Text Tasks about the Battle of Lexington.

Quantitative data included pretreatment and posttreatment measures of three dependent variables—1) the number of inconsistencies identified, 2) the number of evaluative heuristics used in writing, and 3) the number of evaluative heuristics used in reading. Listing Inconsistencies was used to measure the change in the number of inconsistencies identified. The Decision Essay was used to measure the change in the number of evaluative heuristics used in writing. The Justify Trustworthiness task was
used to measure the change in the number of evaluative heuristics used in reading. Prior research and the performance of experienced academic readers on the measures were used to decide on cut scores for each subtest. Based on their scores, each participant was then categorized as High Use or Low Use for each subtest of the measures.

Table 13

Scoring Ranges for Classification as Low Use and High Use

<table>
<thead>
<tr>
<th>Measure</th>
<th>Low Use</th>
<th>High Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listing Inconsistencies</td>
<td>0 - 3</td>
<td>4 or more</td>
</tr>
<tr>
<td>Decision Essay</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sourcing</td>
<td>0 - 1</td>
<td>2 or more</td>
</tr>
<tr>
<td>Corroboration</td>
<td>0 - 1</td>
<td>2 or more</td>
</tr>
<tr>
<td>Contextualization</td>
<td>0</td>
<td>1 or more</td>
</tr>
<tr>
<td>Total Score</td>
<td>0 - 2</td>
<td>3 or more</td>
</tr>
<tr>
<td>Justify Trustworthiness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sourcing</td>
<td>0 - 4</td>
<td>5 or more</td>
</tr>
<tr>
<td>Corroboration</td>
<td>0 - 2</td>
<td>3 or more</td>
</tr>
<tr>
<td>Contextualization</td>
<td>0</td>
<td>1 or more</td>
</tr>
<tr>
<td>Total Score</td>
<td>0 - 6</td>
<td>7 or more</td>
</tr>
</tbody>
</table>

The Listing Inconsistencies test was scored by comparing responses to The Scoring Guide for Listing Inconsistencies and tallying the total number of matches. Based on the cut score, each participant was categorized as either High Use or Low Use in regards to their use of the strategy of identifying inconsistencies. Table 13 presents the
scoring ranges that correspond to Low Use and High Use for all three measures of dependent variables, including Listing Inconsistencies.

The Decision Essay was scored using a modified rubric designed to identify evidence of evaluative heuristic use in writing (Nokes et al., 2007). The Evaluative Heuristic Scoring Guide and The Evaluative Heuristic Scoring Rubric were used by the two scorers to identify the number of evaluative heuristics used in the essays. This instrument yielded four raw scores: number of instances of sourcing in writing, number of instances of corroboration in writing, number of instances of contextualization in writing, and the total number of instances of evaluative heuristics in writing. These raw scores were used to categorize each participant as High Use or Low Use.

The Justify Trustworthiness task was used to measure the number of evaluative heuristics used in reading. The Evaluative Heuristic Scoring Guide and The Evaluative Heuristic Scoring Rubric were used by two scorers to assess the number of evaluative heuristics used. This instrument yielded four raw scores: number of instances of sourcing in reading, number of instances of corroboration in reading, number of instances of contextualization in reading, and the total number of instances of evaluative heuristics used in reading. These raw scores were used to categorize each participant as High Use or Low Use.

This chapter contains the results of this study presented in three sections: Results, Scoring Anomalies, and Summary. The results are presented in response to the research questions.
Results

Research Question 1

What is the effect of an explicit academic literacy instructional unit on the number of inconsistencies identified by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Listing Inconsistencies measure?

The first research question was designed to investigate the change in participants’ ability to identify inconsistencies within or across multiple texts. At the beginning of the study, participants completed the Multiple Text Tasks, which is described above. The first measure in the Multiple Text Tasks is Listing Inconsistencies for which participants list as many inconsistencies as they can remember from the Battle of Lexington document set they have just read and taken notes on. Following the four-week instructional intervention in recognizing and resolving inconsistencies across multiple texts, participants again completed Listing Inconsistencies as part of the Multiple Text Tasks that was identical to the pretest. The frequency counts for preintervention and postintervention scores on Listing Inconsistencies are provided in Figure 2.

It was expected that more participants would be classified as experienced based on identifying more inconsistencies within and across the seven-document set on the Battle of Lexington after participating in the four-week instructional intervention.

Across all students, the total number of inconsistencies identified preintervention was 74 ($M = 2.39$, $SD = 1.61$). Postintervention, the total number of inconsistencies identified across all students was 102 ($M = 3.29$, $SD = 1.83$). Preintervention, six participants were classified as High Use because they identified 4 or more inconsistencies, while postintervention 14 participants were classified as High Use.
There was a 26% increase (from 19% to 45%) in participants classified as High Use for employing the strategy of recognizing inconsistencies.

Figure 2. Frequency Counts for Listing Inconsistencies

The McNemar test for significance of change was conducted for the total score on Listing Inconsistencies to analyze the significance of the difference between preintervention and postintervention categorization. For this analysis, the test yielded a p-value of 0.04. The chi-squared value exceeded the critical value; therefore the null hypothesis was rejected. It was concluded that there were more changes one way than another. A visual examination of the data showed which way the change went. Using McNemar’s test, a significant tendency was found for subjects who demonstrated inconsistency recognition in the range considered Low Use at pretest on the Listing Inconsistencies measure to be more likely to be categorized as High Use at posttest. The results in Figure 2, suggest that participants improved in their ability to recognize
inconsistencies within and across multiple texts after engaging with the instructional intervention.

To provide a fuller picture of the effect of the difference, the confidence interval was calculated. Confidence intervals provide information about a range in which the true value lies with a certain degree of probability, as well as about the direction and strength of the demonstrated effect. This enables conclusions to be drawn about the statistical plausibility and clinical relevance of the study findings. In this study, the confidence interval represents an interval estimate for the difference in proportion of participants who moved from Low Use to High Use ($p = 32.3\%$) and participants who moved the other way ($q = 6.5\%$). The true difference in proportions could be anywhere between .06 and .46 which represents a large range. Thus, the effect of the intervention, although statistically significant, does not indicate a large effect in practical terms.

**Research Question 2**

*What is the effect of an explicit academic literacy instructional unit on developmental freshmen’s ability to use evaluative heuristics as writers as measured by differences between their pretest and posttest scores on the Decision Essay measure?*

The second research question was designed to investigate the change in participants’ ability to resolve inconsistencies within or across multiple texts using evaluative heuristics in writing. The second measure in the Multiple Text Tasks administered preintervention and postintervention is the Decision Essay for which participants write an essay about who fired first at the Battle of Lexington using the notes they took while reading the document set. The Decision Essay yields four scores for evaluative heuristic use in writing—sourcing, corroboration, contextualization, and total
evaluative heuristic use. It was expected that more participants would be classified as High Use based on evaluative heuristic use in writing the Decision Essay about the Battle of Lexington after participating in the four-week instructional intervention.

The data related to the sourcing heuristic were examined first. Across all students, the total number of sourcing heuristics identified in the Decision Essay preintervention was 3 ($M = 0.10, SD = 0.30$). Postintervention, the total number of sourcing heuristics used in writing across all students was 14 ($M = 0.45, SD = 1.03$). Preintervention, zero participants were classified as High Use, while postintervention 2 participants were classified as High Use because they used the sourcing heuristic 2 or more times. Thus, 6% of the students moved from being classified as Low Use to High Use of the sourcing heuristic. Figure 3 presents the frequency counts for the number of sourcing heuristics used in the Decision Essay.

![Frequency Counts for the Decision Essay: Sourcing](image)

*Figure 3. Frequency Counts for the Decision Essay: Sourcing*
The McNemar test for significance of change was conducted for the sourcing score on the Decision Essay to analyze the significance of the difference between preintervention and postintervention categorization. For this analysis, the test yielded a p-value of 0.50. The chi-squared value did not exceed the critical value; therefore the null hypothesis was preserved.

Next, the data related to the corroboration heuristic were examined. Across all students, the number of corroboration heuristics identified in the Decision Essay preintervention was 4 ($M = 0.13, SD = 0.34$). Postintervention, the total number of corroboration heuristics used in writing across all students was 9 ($M = 0.29, SD = 0.64$). Preintervention, zero participants were classified as High Use, while postintervention 1 participant was classified as High Use because she used the corroboration heuristic 2 or more times. Thus, 3% of the students moved from being classified as Low Use to High Use of the corroboration heuristic. Figure 4 presents the frequency counts for the number of corroboration heuristics used in the Decision Essay.

The McNemar test for significance of change was conducted for the corroboration score on the Decision Essay to analyze the significance of the difference between preintervention and postintervention categorization. For this analysis, the test yielded a p-value of 1.00. The chi-squared value did not exceed the critical value; therefore the null hypothesis was preserved.
When the data related to the contextualization heuristic were examined, it showed that across all students, the number of contextualization heuristics identified in the Decision Essay preintervention was 9 ($M = 0.29, SD = 0.46$). Postintervention, the total number of contextualization heuristics used in writing across all students was 14 ($M = 0.45, SD = 0.72$). Preintervention, 9 participants were classified as High Use because they used the contextualization heuristic one or more times, while postintervention 10 participants was classified as High Use. There was a 3% increase (from 29% to 32%) in participants classified as High Use for the contextualization heuristic. Figure 5 presents the frequency counts for the number of contextualization heuristics used in the Decision Essay.

The McNemar test for significance of change was conducted for the contextualization score on the Decision Essay to analyze the significance of the
difference between preintervention and postintervention categorization. For this analysis, the test yielded a p-value of 1.00. The chi-squared value did not exceed the critical value; therefore the null hypothesis was preserved.

Figure 5. Frequency Counts for Decision Essay: Contextualization

Finally, across all students, the total number of evaluative heuristics (sourcing, corroboration, & contextualizing) identified in the Decision Essay preintervention was 16 ($M = 0.52, SD = 0.68$). Postintervention, the total number of evaluative heuristics used in writing across all students was 37 ($M = 1.19, SD = 1.30$). Preintervention, zero participants were classified as High Use, while postintervention 5 participants were classified as High Use because they used a total of 3 or more evaluative heuristics. This represents a change of 16% in participants classified as High Use in their total use of evaluative heuristics. Figure 6 presents the frequency counts for the total number of evaluative heuristics used in writing on the Decision Essay.
The McNemar test for significance of change was conducted for the total evaluative heuristic use score on the Decision Essay to analyze the significance of the difference between preintervention and postintervention categorization. For this analysis, the test yielded a p-value of 0.06. The chi-squared value did not exceed the critical value; therefore the null hypothesis was preserved. Using the McNemar test, no significant tendency was found for subjects who demonstrated heuristic use in the range considered Low Use on the Decision Essay preintervention to be more likely to be categorized as High Use postintervention.

![Graph](image)

**Figure 6.** Frequency Count for the Decision Essay: Total

In summary, participants used more evaluative heuristics postintervention than preintervention. However, none of the increases on the four subtests were statistically significant.
Research Question 3

What is the effect of an explicit academic literacy instructional unit on developmental freshmen’s ability to use evaluative heuristics as readers as measured by differences between their pretest and posttest scores on the Justify Trustworthiness measure?

The third research question was designed to investigate the change in participants’ ability to resolve inconsistencies within or across multiple texts using evaluative heuristics in reading. At the beginning of the study, participants completed the Multiple Text Tasks, which is described above. The third measure in the Multiple Text Tasks is the Justify Trustworthiness task for which participants are provided the source information (e.g., author, date, document type) for each of the seven documents they read about the Battle of Lexington. Participants are asked to rank the credibility of each document and write a justification for each of their decisions about credibility. Participants write one or more sentences evaluating the credibility of each of the seven documents about the Battle of Lexington using the notes they took while reading the document set. Following the four-week instructional intervention in recognizing and resolving inconsistencies across multiple texts, participants again completed the Justify Trustworthiness task as part of the Multiple Text Tasks that was identical to the pretest.

The Justify Trustworthiness task yields four scores for evaluative heuristic use in reading—sourcing, corroboration, contextualization, and total heuristic use. The preintervention and postintervention scores for each of the four categories of evaluative heuristic in reading are provided in figures. It was expected that more participants would be classified as High Use based on evaluative heuristic use in reading when justifying trustworthiness after participating in the four-week instructional intervention.
Examining first the sourcing heuristic use across all students, the number of sourcing heuristics identified on the Justify Trustworthiness task preintervention was 84 ($M = 2.71$, $SD = 2.19$). Postintervention, the total number of sourcing heuristics used in reading across all students was 94 ($M = 3.03$, $SD = 1.87$). Preintervention, 8 participants were classified as High Use because they used 5 or more instances of the sourcing heuristic, while postintervention 8 participants were classified as High Use. Thus, no additional students moved from being classified as Low Use to High Use of the sourcing heuristic postintervention. The frequency counts for use of the sourcing heuristic in reading on the Justify Trustworthiness measure are presented in Figure 7.

![Frequency Counts for Justify Trustworthiness: Sourcing](image)

**Figure 7. Frequency Counts for Justify Trustworthiness: Sourcing**

The McNemar test for significance of change was conducted for the sourcing score on the Justify Trustworthiness task to analyze the significance of the difference between preintervention and postintervention categorization. For this analysis, the test
yielded a p-value of 1.00. The chi-squared value did not exceed the critical value; therefore the null hypothesis was preserved.

With the corroboration heuristic, across all students the number of corroboration heuristics identified on the Justify Trustworthiness task preintervention was 2 ($M = 0.06$, $SD = 0.36$). Postintervention, the total number of corroboration heuristics used in writing across all students was 4 ($M = 0.13$, $SD = 0.43$). Preintervention, zero participants were classified as High Use, and postintervention zero participants were classified as High Use because none of them demonstrated 3 or more instances of the corroboration heuristic. Thus, no students moved from being classified as Low Use to High Use of the corroboration heuristic. The frequency counts for use of the corroboration heuristic in reading on the Justify Trustworthiness measure are presented in Figure 8.

Figure 8. Frequency Counts for Justify Trustworthiness: Corroboration
The McNemar test for significance of change was not conducted for the corroboration score on the Justify Trustworthiness task because there was no difference between preintervention and postintervention categorization. Therefore, the null hypothesis was preserved.

Next, the use of the contextualization heuristic was examined. Across all students, the number of contextualization heuristics identified on the Justify Trustworthiness task preintervention was 2 ($M = 0.06, SD = 0.25$). Postintervention, the total number of contextualization heuristics used in reading across all students was 7 ($M = 0.23, SD = 0.56$). Preintervention, 2 participants were classified as High Use because they used 1 or more instances of the contextualization heuristic, while postintervention 5 participants were classified as High Use. There was a 10% increase (from 6% to 16%) in participants classified as High Use for the contextualization heuristic. The frequency counts for use of the contextualization heuristic in reading on the Justify Trustworthiness measure are presented in Figure 9.

The McNemar test for significance of change was conducted for the contextualization score on the Justify Trustworthiness task to analyze the significance of the difference between preintervention and postintervention categorization. For this analysis, the test yielded a p-value of 0.45. The chi-squared value did not exceed the critical value; therefore the null hypothesis was preserved.
Finally, the total number of evaluative heuristics (sourcing, corroboration, & contextualization) used in reading was examined. Across all students, the total number of evaluative heuristics used in the Justify Trustworthiness task preintervention was 88 ($M = 2.84$, $SD = 2.11$). Postintervention, the total number of evaluative heuristics used in reading across all students was 105 ($M = 3.39$, $SD = 2.01$). Preintervention, 1 participant was classified as High Use, while postintervention 2 participants were classified as High Use because they used 7 or more total instances of evaluative heuristics. This represents a 3% increase in the participants categorized as High Use on their total use of evaluative heuristics. The frequency counts for the total number of evaluative heuristics used in reading on the Justify Trustworthiness measure are presented in Figure 10.
Figure 10. Frequency Counts for Justify Trustworthiness: Total

The McNemar test for significance of change was conducted for the total evaluative heuristic use score on the Justify Trustworthiness task to analyze the significance of the difference between preintervention and postintervention categorization. For this analysis, the test yielded a p-value of 1.00. The chi-squared value did not exceed the critical value; therefore the null hypothesis was preserved. Using the McNemar test, no significant tendency was found for participants who demonstrated evaluative heuristic use in reading in the range considered Low Use on the Justify Trustworthiness measure to be more likely to be categorized as High Use postintervention. Therefore, the null hypothesis was preserved.

In summary, participants used only slightly more evaluative heuristics in reading postintervention as measured by the Justify Trustworthiness task. The results were not statistically significant for evaluative heuristic use in reading.
Summary of Results

The first research question was designed to investigate the change in participants’ ability to identify inconsistencies within or across multiple texts. Listing Inconsistencies was used to measure the dependent variable—participants’ ability to identify inconsistencies across multiple documents. Analysis of the data suggests that participants increased their ability to identify inconsistencies within or across multiple texts. The results were statistically significant. However, the results were determined to be of small practical significance.

The second research question was designed to investigate the change in participants’ ability to resolve inconsistencies within or across multiple texts using evaluative heuristics in writing. Analysis of the data suggests that although participants increased the number of evaluative heuristics they used in writing, the gains were not statistically significant. At posttest, participants demonstrated an increased use of evaluative heuristics in writing for each subtest of the Decision Essay. Despite the increase in usage and the increase in number of participants categorized as High Use, none of the results reached statistical significance.

The third research question was designed to investigate the change in participants’ ability to resolve inconsistencies within or across multiple texts using evaluative heuristics in reading. At posttest, the gains in evaluative heuristic usage in reading were minimal and none of the results from the Justify Trustworthiness measure reached significance.
Although the data indicate that participants increased in their ability to identify inconsistencies, the change in evaluative heuristic usage failed to reach statistical significance.

Scoring Anomalies

The two scorers read the tests in sets of three and then switched sets, starting with Listing Inconsistencies and ending with the Justify Trustworthiness tasks. The scorers used the researcher-created Scoring Guide for the Listing Inconsistencies Protocol and The Evaluative Heuristics Scoring Guide, which was modified by the researcher from Nokes et al.’s heuristic scoring guide (2007). For the Listing Inconsistency measure the scorers wrote directly on the student’s list. The Evaluative Heuristic Scoring Rubric was used for scoring both the Decision Essay and the Justify Trustworthiness measures.

The scorers also used The Evaluative Heuristic Scoring Rubric to take notes about why they were assigning points or why they were uncertain about assigning points. For example, on the line for direct comparison, the scorer would make a mark if evidence of that evaluative heuristic were present, but would also list the documents being compared. If the scorer was uncertain about assigning a point, she might include a question mark and/or a brief note about her uncertainty.

After both scorers had finished scoring the full set of six, they reviewed their rubrics and discussed any questions or disagreements. The scorers found that they had several items to discuss for each set of six tests. Even with the scoring guides, a number of challenges arose: difficulty categorizing the specific aspect of an evaluative heuristic (e.g., author’s credentials vs. author’s motivation for sourcing); trying to decide how much to read into students’ explanations because students were inexperienced with this
type of writing; distinguishing a simple misreading (e.g., *president of the pro-team* for *president pro tem*) from an incorrect response (e.g., *Ezra Stiles was a participant in the battle*); and consistently applying the rubric.

Most of these discussions focused on notes and questions and were not considered scorer disagreements unless both scorers continued to disagree after discussion. To address disagreements, the scorers decided on several decision rules in addition to those outlined in the scoring guides to help clarify the awarding of points (e.g., If the student makes any mention of the novel being fiction, they are awarded a point for the *document type* aspect of sourcing). Throughout the discussion of these difficulties below, the term “assign” means that a scorer made a mark on their rubric—a tally, a tally with a question mark, or a note—to discuss with the other scorer, while the term “award” means that a decision was reached between scorers about giving a point for that item. In the examples from the measures presented below to illustrate the various challenges, students’ work has been transcribed with the errors preserved.

*Difficult to Categorize*

On the Justify Trustworthiness measure, the link to credibility is implied by the structure of the measure—students rank the credibility of each document and write a brief justification of their assessment of credibility. It is as if the student were saying that “I assigned a 2 to this article. It is credible because...” or “I assigned a 7 to this article. It is not credible because...” On the Decision Essay, in contrast, the student must be explicit about the connection to credibility. According to The Evaluative Heuristic Scoring Guide, “Credit for corroboration should only be given when it helps the individual make sense of the event” (Appendix D). To some extent the scorer must read
into how it helped the reader make sense of events. According to The Evaluative Heuristic Scoring Guide, “An individual uses contextualization when he or she discusses specific details about the event that helps him or her understand why or how the event took place” (Appendix D). As with corroboration, the scorer must decide if the particular instance helped the student understand how or why. For each type of evaluative heuristic the overall burden of proof is different, which can affect the categorization of points.

The Evaluative Heuristics Rubric includes space to tally each aspect of each evaluative heuristic, such as the document type aspect of sourcing or the linguistic awareness aspect of contextualization. The scorers sometimes disagreed how to categorize an aspect of evaluative heuristic use. When the disagreement was about different aspects of the same category (e.g., author’s credentials vs. author’s motivation), there was no need to resolve the difference, since the score was based on the number of sourcing heuristics used rather than the finer-grained aspects. However, some issues, such as how to categorize arguments about the minutemen’s agreement of the account presented in Document 2 were more challenging. About a half-dozen students were arguing for corroboration among the minutemen in their sworn statement, but to be eligible for a point for the corroboration heuristic, the corroboration would need to take place across documents. The students are pointing out corroboration within a single document to evaluate credibility, so the scorers awarded a point of the other evaluation of document aspect of sourcing.

The scorers encountered a categorization challenge on Student 814’s Justifying Trustworthiness posttest. Scorer 2 wanted to give a point for the linguistic awareness aspect of contextualization, but the instance does not fit the description for that aspect as
delineated by The Scoring Guide for Evaluative Heuristics. In his justification for Document 3, the student wrote, “I believe his story of his dad being shot, but because he says ‘I seem to remember’ it makes me feel uncertain of the source.” The student is paying attention to the word choice in Document 3. However, the description of the linguistic awareness aspect of contextualization is “a keen awareness of the different meanings of words over time” (Appendix D). The meaning of these words has not changed over time. After discussion, a point was awarded for the other evaluation of the document aspect of sourcing because the attention to the word choice is used to evaluate credibility (sourcing) rather than explaining how or why the event took place.

Other justifications were difficult to categorize because they were not covered in the scoring guides. For example, Student 753 used the absence of a publication date to determine credibility on the pretest for the Justify Trustworthiness task. Scorer 1 assigned a point for the date of production aspect of sourcing, but Scorer 2 disagreed because the student did not explain why the lack of a publication date would lower credibility. Scorer 1 argued that the lack of expected information, as this was the only document without a publication date, would lower credibility just by being absent. When Scorer 1 referred to The Evaluative Heuristic Scoring Guide, there was no support for giving a point for pointing out the absence of important information, unless it was between two specific documents in the document set (corroboration). Therefore, no point was awarded.

Challenges of the Demographic: Inexperienced Writers

In addition to hard to categorize explanations, the scorers also grappled with explanations that were difficult to decipher. The participants in this study are
inexperienced with the academic writing expectations of the university. They are not necessarily skilled in argumentative writing. Therefore, the scoring is anomalous in that each response must be considered on a case-by-case basis.

One way this inexperience manifested itself was in not providing enough explanation for the scorers to understand the student’s argument. For example, Student 329 wrote, “President Warren could have altered the story to make it seem more ideal or to just cover up some facts” (pretest for the Justify Trustworthiness task). The student does not give any reason why president pro tem of the Massachusetts Provisional Congress, Joseph Warren, would lie. Therefore, no point was awarded for the author’s credibility aspect of sourcing. This is a mark of inexperience writers; they often fail to adequately explain why.

Inexperienced writers use imprecise word choices, which leave the scorer to parse out the possible meanings. Student 979 wrote, “Testimony of actual participants in top 3 trustworthy because they were physically there at the battle” (posttest for the Justify Trustworthiness task). One point was awarded for the author’s participation aspect of sourcing. However, no point was awarded for the document type aspect of sourcing, because the student didn’t mention the swearing to tell the truth. The use of the word “testimony” does not adequately convey that the minutemen were under oath. A limited vocabulary is another mark of inexperience with academic writing.

The students’ lack of experience with argumentation may be the cause of debate between scorers. Student 527 wrote, “If it’s in a high school textbook, that’s being taught to students, it should have some basic facts” (pretest for the Justify Trustworthiness task). Scorer 2 originally assigned a point for the document type aspect of sourcing because she
agreed that we expect the textbooks in schools to at least maintain a minimal level of factual accuracy. However, Scorer 1 saw this as a general comment about what should be the standard for textbooks, rather than a claim about this textbook’s credibility. In addition, the rank assigned by the student was difficult to decipher, possibly a 4 (on a 1-7 range), so it was not clear whether the student was saying that it was credible because of basic facts or only somewhat credible because of basic facts. It is also not clear if the textbook should be credible or if the student assumes it is credible since the expectation of basic facts has been met. Because the claim is not clear, no point was awarded. This lack of clarity is another mark of inexperienced academic writing.

Scorers’ Prior Experience Contributes to Reading into Students’ Work

Scorers were selected because of their experience with this population. They are aware of the issues encountered by inexperienced readers and writers. However, that prior experience reading student writing, can also run counter to the goals of objective assessment, as these scorers are in the habit of giving these students the benefit of the doubt when grading essays.

During the scoring, the scorers found themselves reading into students’ work. The Listing Inconsistencies measure provided lines for noting each inconsistency (e.g., *What was the British purpose?* or *march through vs. engage in battle*). Yet, approximately six students did not make the inconsistency explicit on a single line. For example, Student 285 listed isolated concepts on each line. On line 4, the student wrote, “‘Regulars’ violently marched into Lexington.” And on line 8, the student wrote, “The troops had no intent to attack first” (pretest for Listing Inconsistencies). Despite the 3-item gap between line 4 and line 8, the scorers paired statements across lines. Therefore,
one point was awarded for identifying an inconsistency across documents for *What was the British purpose for marching on/though Lexington Green?* as listed in the Scoring Guide for the Listing Inconsistencies Protocol.

The scorers felt that the student had identified an inconsistency and should get credit for it. The scorers were consistent in applying this standard across all Listing Inconsistencies tests. However, this scoring approach was inconsistent with the design and intended use of the test. This could be seen as overreaching, as the student may not have consciously identified an inconsistency.

*Inexperienced Reader: How Wrong?*

The scorers had to decide if an error represented a misreading (minor or predictable errors) or an incorrect understanding of the document(s). Students’ struggles with the term “pro tem” typify what the scorers considered a misreading. It was not surprising to the scorers that students would be unfamiliar with the term. In Document 1, Student 684 appears to have misread *president pro tem*, writing, “This one [document] is semi-trustworthy as it gives good information but it is from the ‘pro-team’ so it could be biased” (posttest for the Justify Trustworthiness task). Despite the student’s recasting of the term pro tem, a point was awarded for the *author’s motivation* aspect of sourcing because the student did give a reason for potential bias.

*Difficulty Applying Consistent Scoring Standards*

The scorers found it difficult to maintain consistency because each student writes their reasoning/justification a little bit differently. For example, the following two examples are about justifying the trustworthiness of Document 3, the novel. Student 660 wrote, “I rank [Document 3] a 6 because it was a novel, so it could be the author’s
opinion on what happened at the Battle” (pretest for the Justify Trustworthiness task).

Scorer 1 was uncertain about assigning a point because the student points out that the novelist could make things up based on their opinion or perspective. However, Scorer 2 argued that the reasoning that it could be the author’s opinion was applicable to any piece of writing and the writer hadn’t shown why we should think that about this piece of writing. No point was awarded for the document type aspect of sourcing.

By comparison, Student 268 wrote, “I gave this one a [rank of] 6 because novels are also always altered and they don’t always get the story right” (pretest for the Justify Trustworthiness task). The scorers took this to mean that in a novel the author can alter the facts, or fictionalize them. One point was awarded for the document type aspect of sourcing. There is a very narrow difference between examples. In both cases, the scorers read into what the student wrote.

**Decision Rules**

One way the scorers attempted to maintain consistency was to create decision rules in an attempt to apply the same principle to every student response. Despite the creation of decision rules, their application remained challenging. Some examples of the razor thin reasoning between awarding a point and not awarding a point have already been presented, like the fiction example above. This section presents two additional scoring anomalies that challenged the decision rules.

**The Official Document Decision Rule and Degrees of Misreading**

When looking across several student examples, the challenge of applying the decision rule concerning official documents comes into focus. The scorers made a decision rule allowing for Document 1, the cover letter from Joseph Warren, President
pro tem of the Massachusetts Provincial Congress to Benjamin Franklin, the colonial representative in London, to be considered an “official document,” thereby earning a point for the document type aspect of sourcing. As a result, the scorers also created a decision rule allowing references to the President, and later, any president as a reliable official to be counted as an instance of the author’s credentials aspect of sourcing.

Even with the decision rule, the scorers tried to weigh out the roles of naiveté, inexperience with argument, and errors in reading comprehension. For example, Student 413 wrote, “The document is trustworthy because it is the president who is talking in the cover letter. The president’s words are credible because the president would not lie about what happened since his words are usually the promises to his country” (pretest for the Justify Trustworthiness task). Although naïve, the student explains why the President’s word should be credible. The student was incorrect about Joseph Warren. He was not the President and there was no U.S. President at this point in history. However, since “president pro tem” is confusing to most people, we awarded the student a point for the author’s credentials aspect of sourcing.

If the president pro tem of a colonial congress is to be treated like the president, then should the president of any group be considered equally credible as the President? In regards to document 6, Student 413 wrote about Ezra Stiles, the president of Yale College: “as a president, he wouldn’t have lied about what happened” (posttest for the Justify Trustworthiness task). Since this same student was given a point for the author’s credentials in referencing one president—President pro tem of the Massachusetts Provisional Congress, Joseph Warren—the scorers felt the student should also get a point for the president argument with Document 6. While reviewing the scoring, the researcher
noted that at the time, the presidents of colleges probably were seen as particularly trustworthy individuals, even though that may not be the case today.

With the next example, the application of the decision rule gets stickier. Student 285 wrote, “These are trusted men by Massachusetts. Highly respected men, but they were not there on the battlefield” (posttest for the Justify Trustworthiness task). Now the claim about the author is more general and includes the recipient.

The scorers agreed to award a point for the *author’s participation* aspect of sourcing. Scorer 1 also assigned a point for the *author’s credentials* aspect of sourcing. If being president was enough of a reason to award a point then the elected official reason could apply across positions. Student 285 points out that the sender and the recipient were trusted by Massachusetts, which an election or nomination would seem to support. Despite the initial disagreement, a point was awarded for the *author’s credentials* aspect of sourcing after discussion. This example shows the expansion of the decision rule beyond the President and a president, to “trusted men”.

In earlier examples the scorers were willing to overlook the conflation of Joseph Warren with the President, but other instances required the scorers to decide how much misreading was allowable. Student 517 wrote, “I believe [Document 1] is more trustworthy because it was a letter sent between congresses, and was also sworn to be sent” (posttest for the Justify Trustworthiness task). The student misunderstood that this was the cover letter, not the sworn depositions. The student also states that it was between congresses. It is true that it was sent on behalf of the Massachusetts Provincial Congress, but it was sent to Benjamin Franklin, not another congress. However, the student’s claim that it was sent between congresses does indicate that it was an official
letter, which would meet the criteria for the decision rule. Despite the student’s misrepresentation, a point was awarded for the document type aspect of sourcing.

Probably, the most debated decision was Student 484’s justification for Document 1 at posttest. The scorers debated this item extensively. Student 484 wrote, “Based on Benjamin Franklin, writing a letter. We can trust him because he was the president, and he can’t lie to his country.” The student’s explanation contains the following four inaccuracies: 1) The president CAN lie to his country; 2) This was before there was a United States, let alone a President of the U.S.; 3) Benjamin Franklin was never President of the U.S.; and 4) Benjamin Franklin is the recipient, not the author of the letter.

Although the decision rule was that the document being official should count for the document type aspect of sourcing, this justification presented additional challenges. Although the scorers agreed that claims about the president, even though there was no U.S. President yet, could count for the author’s credentials aspect of sourcing as it had in other instances, this misreading (recipient vs. author) disqualified the point for the author’s credentials aspect of sourcing. The student fundamentally misinterpreted the sourcing information: Benjamin Franklin was not the author of this document. Therefore, any conclusions drawn from that factual inaccuracy are faulty, including the claim that the letter is official. No point was awarded. Even with a decision rule that had been carefully considered, there were scenarios that made its application challenging.

The “Interesting the Reader” Decision Rule and the Unresolved Anomaly

An example of an unresolved anomaly comes from Student 329’s justification about the role making writing interesting plays in assessing credibility. Student 329 wrote, “Newspapers typically get their information from witnesses, but they can also alter
what they were given, just to make their story more interesting” (posttest for the Justify Trustworthiness task).

A point was awarded for the *author’s participation* aspect of sourcing (use of first-hand accounts). Originally, both scorers had given a point for *document type* aspect of sourcing (“alter what they were given, just to make their story more interesting”) since that type of text relies on public interest for purchase. After discussion, the scorers decided not to award a point since the student was not explicit about newspapers relying on revenue.

When the researcher was reviewing the data, however, she noticed that the student had been awarded a point for the *document type* aspect of sourcing for Document 3, the novel. In the justification for Document 3, Student 329 wrote, “The author could have used the Battle of Lexington as a starting point for his novel, but he could have changed some of the facts to make the story more interesting” (posttest for the Justify Trustworthiness task). The student is noting the fictional quality of novels and the reason why a novelist would “change some of the facts” was “to make the story more interesting” since this type of document relies on interesting the public in order to make a profit. Although the reason why—to make it interesting—is the same as in Document 5, the additional information about the fictional nature of the novel seems to have convinced the scorers to award a point. From the researcher’s point of view, the application of this decision rule seems inconsistent leading to a scoring anomaly.

*Summary*

The sheer quantity of anomalies raises concerns about the validity of the scoring. The examples discussed above present a picture of the amount of discussion and debate
required to reach somewhat consistent scoring. In some cases, the discussion of anomalies resulted in the creation of decision rules that were applied to later instances. Even with the decision rule, the scorers tried to weigh out the roles of naiveté, inexperience with argument, and errors in reading comprehension. Despite conscientious use of the Scoring Guide, creation of addition decision rules to guide scorers, and attempts to be consistent in applying scoring decisions, many items were scored case-by-case on the basis of scorers’ interpretation, meaning that a different set of scorers may have made other decisions regarding the awarding of points.

Summary

This chapter contained the results of the three research questions that were the basis of the present study. The raw scores on the subtests from the three measures were used to categorize participants as High Use or Low Use. The McNemar test for significance of change was used to answer all three research questions.

There were significant gains in participants’ ability to recognize inconsistencies across multiple texts. At posttest, fourteen participants were classified as High Use in their ability to identify inconsistencies. This represented a statistically significant change. However, the increase in evaluative heuristic use to resolve those inconsistencies was modest and failed to reach statistical significance for any subtest of the Decision Essay or Justify Trustworthiness measure. At posttest, participants demonstrated an increased use of evaluative heuristics in writing for each subtest of the Decision Essay. Despite the increase in usage and the increase in number of participants categorizes as High Use, none of the results reached statistical significance. The gains in evaluative
heuristic usage in reading were minimal and none of the results from the Justify Trustworthiness measure reached significance.

The number of student responses requiring discussion to arrive at a score indicates the challenges in scoring evaluative heuristic use. The scorers encountered responses that were difficult to categorize, responses that were difficult to parse because of the participants’ inexperience with academic writing, and challenges with consistently applying the scoring standards. The scorers supplemented the scoring guides with decision rules to help apply the rubric across students’ responses and were able to reach agreement in all but a handful of instances. However, the number of case-by-case decisions made by the scorers suggests the anomalous nature of some scoring decisions.
CHAPTER V

CONCLUSIONS

This chapter begins with a summary of the study. The rest of this chapter contains a summary of the findings, the limitations of the study, a discussion of the results of the study, and implications for future research and practice in the area of academic literacy.

Summary of the Study

Academic literacy is positively correlated with academic success at the university (Bosley, 2008; Pugh et al., 2000). In university contexts, students are expected to go beyond simple reading comprehension—to use reading to independently build knowledge, to apply what they learned from reading, often in the form of written work, and to solve novel problems.

However, a significant portion of students arrives at the university underprepared to meet these expectations for academic literacy. In California, more than 60% of the 40,000 freshmen admitted to the CSU require remediation (National Center for Public Policy and Higher Education, 2008). These students are placed in developmental-level reading and writing classes to help them attain the skills they will need to successfully navigate the academic literacy demands of the university.

Multiple text studies (e.g., Bråten et al., 2009; Britt & Aglinskas, 2002; Wineburg, 1991, 1998), which are closely aligned with the complex academic literacy practices of the university, can help guide the instruction in these developmental courses. These multiple text studies suggest that inexperienced, or novice, readers differ from experienced (expert and advanced) academic readers in their lack of awareness of the
complexity of academic literacy. In particular, inexperienced readers fail 1) to detect inconsistencies across texts, and 2) to employ conditional knowledge to strategically resolve inconsistencies.

Inexperienced readers tend to gloss over contradictory evidence provided in different texts (Britt & Aglinskas, 2002; Otero & Kintsch, 1992; Wineburg, 1991). In contrast, experienced readers notice inconsistencies, ask specific questions, and formulate action plans to resolve these inconsistencies (Wineburg, 1991, 1998). This strategic behavior is a key difference between expert and novice readers. Experienced readers utilize conditional knowledge, knowing when and why to apply a strategy (Paris et al., 1983). Inexperienced readers tend not to notice comprehension issues and, therefore, may not realize they should mobilize a strategic approach (Garner, 1994).

In a pair of landmark studies, Wineburg illuminated the academic reading behaviors of experienced readers. Wineburg (1991, 1998) identified three evaluative heuristics that expert academic readers use to resolve inconsistencies: a sourcing heuristic, a corroboration heuristic, and a contextualization heuristic. Sourcing refers to using source characteristics (e.g., author’s credentials or type of text) to evaluate a document’s credibility; corroboration is a strategy for comparing and contrasting information across documents; and contextualization refers to attempts to better understand events by reconstructing the context that surrounds them. Experienced academic readers use elements of these three heuristics flexibly to evaluate evidence and resolve inconsistencies (Rouet et al., 1997; Wasson, 2001; Wineburg, 1991, 1998).

Although undergraduates possess declarative knowledge (they can tell you that they should evaluate sources), they demonstrate a lack of conditional knowledge by not
utilizing evaluative heuristics when necessary. Inexperienced, developmental-level students are not engaging in the very activities that could contribute to their academic success at the university.

The primary purpose of this pre-experimental study was to investigate the effectiveness of an explicit academic literacy intervention with one group of developmental-level freshmen from two intact sections of a developmental-level integrated reading and writing course (N = 31). Specifically, this study investigated the influence of providing explicit instruction in recognizing and resolving inconsistencies across multiple texts on participants’ ability to 1) recognize inconsistencies, 2) use evaluative heuristics to reconcile inconsistencies when writing, and 3) use evaluative heuristics to reconcile inconsistencies when reading.

This study attempted to answer the following research questions:

1. What is the effect of an explicit academic literacy instructional unit on the number of inconsistencies identified by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Listing Inconsistencies measure?

2. What is the effect of an explicit academic literacy instructional unit on developmental freshmen’s ability to use evaluative heuristics as writers as measured by differences between their pretest and posttest scores on the Decision Essay measure?

3. What is the effect of an explicit academic literacy instructional unit on developmental freshmen’s ability to use evaluative heuristics as readers as
measured by differences between their pretest and posttest scores on the Justify Trustworthiness measure?

This study used a pre-experimental one group pretest-posttest design to investigate the effectiveness of an explicit academic literacy instructional intervention in the context of a developmental-level integrated reading and writing classroom. Participants received explicit instruction in recognizing inconsistencies within and across multiple texts and using evaluative heuristics, like those that experienced academic readers use to resolve these inconsistencies. A pretest was administered, followed by the four-week instructional intervention. Participants then completed the posttest. Measures of five student background variables—age, gender, ethnicity, language background, and familiarity with the topic of the document set for the Multiple Text Tasks—were administered prior to the start of the intervention.

The independent variable was the instructional intervention to improve developmental-level students’ academic literacy skills by identifying inconsistencies and using evaluative heuristics to resolve those inconsistencies. There were three dependent variables: 1) the number of inconsistencies identified, 2) the number of evaluative heuristics used in writing, and 3) the number of evaluative heuristics used in reading. Measures included Listing Inconsistencies, a researcher-designed measure of the number of inconsistencies identified; the Decision Essay to measure the number of evaluative heuristics used in writing, and the Justify Trustworthiness task, designed to measure the number of evaluative heuristics used in reading. Quantitative data from these three measures were collected preintervention and again postintervention.
Summary of the Findings

This section outlines the summary of the findings of the study. It is organized around the three research questions.

The first research question examined the effect of an explicit academic literacy instructional unit on the number of inconsistencies identified by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Listing Inconsistencies measure. The total number of inconsistencies across all students increased from 74 at pretest to 102 at posttest. Results show that the number of participants categorized as High Use increased from 6 at pretest to 14 at posttest. The McNemar test revealed a statistically significant tendency for subjects who demonstrated inconsistency recognition in the range considered Low Use at pretest on the Listing Inconsistencies measure to be more likely to be categorized as High Use at posttest. A confidence interval was calculated indicating a small effect in practical terms.

The second research question examined the effect of an explicit academic literacy instructional unit on developmental freshmen’s ability to use evaluative heuristics in writing as measured by differences between their pretest and posttest scores on the Decision Essay measure. Usage of the sourcing heuristic across all participants increased from 3 at pretest to 14 at posttest. The number of participants categorized as High Use increased from zero to 2. Across all students, the use of the second evaluative heuristic—corroboration—increased from 4 at pretest to 9 at posttest. The number of participants categorized as High Use increased from zero to 1. Participants’ aggregate use of the contextualization heuristic increased from 9 at pretest to 14 at posttest. The number of participants categorized as High Use expanded from 9 to 10. Across all Decision Essays,
the total use of evaluative heuristics increased from 16 at pretest to 37 at posttest, while
the number of participants categorized as High Use expanded from zero to 5.
McNemar’s test indicated that none of the changes for any subtest of the Decision Essay
were statistically significant.

The third research question examined the effect of an explicit academic literacy
instructional unit on developmental freshmen’s ability to use evaluative heuristics in
reading as measured by differences between their pretest and posttest scores on the
Justify Trustworthiness measure. Across all students, the use of the sourcing heuristic in
reading increased from 84 at pretest to 94 at posttest. The number of participants
categorized as High Use remained static with 8 at pretest and at posttest. The McNemar
test indicated that the change was not statistically significant. The aggregate use of the
corroboration heuristic in reading increased from 2 instances at pretest to 4 at posttest.
The number of participants categorized as High Use remained constant at zero. Across
all students, the use of the contextualization heuristic in reading increased from 2
instances at pretest to 7 at posttest. The number of participants categorized as High Use
expanded from 2 to 5. McNemar’s test indicated that the change was not statistically
significant. Overall, the total use of evaluative heuristics in reading increased from 88 at
pretest to 105 at posttest. The number of participants categorized as High Use increased
from 1 to 2. McNemar’s test indicated that the change was not statistically significant.

Discussion of Results

This section focuses on the discussion of the findings of the study in relation to
the research literature. The section is organized around recognizing inconsistencies and
using evaluative heuristics. Research Question 2 and Research Question 3 are both discussed in the second section.

The Impact of the Academic Literacy Intervention on Identifying Inconsistencies

Research suggests that inexperienced readers do not notice inconsistencies when they read (Britt & Aglinskas, 2002; Otero & Kintsch, 1992; Wineburg, 1991), whereas experienced readers focus on inconsistencies (Wineburg, 1998). Once an inconsistency is identified, the reader can identify the break down in comprehension and devise a plan to resolve it.

Prior intervention research (e.g., Britt & Aglinskas, 2002; Nokes et al., 2007; Wiley et al., 2009) has overlooked the recognition of inconsistencies, the trigger for deploying conditional knowledge. In prior studies, researchers have typically provided a specific controversy for the participants to resolve, such as asking participants to decide who fired first at the Battle of Lexington. Participants are asked to respond to an identified inconsistency instead of being asked to identify inconsistencies on their own (e.g., Bråten & Strømsø, 2006; Britt & Aglinskas, 2002; Stahl et al., 1996). Thus, there was a need for research that examines interventions that would help students learn to identify inconsistencies across multiple texts as the first step in deploying conditional knowledge, and the current study helps to fill this need.

There is a gap in the descriptive literature, as well. Aside from Wineburg (1991, 1998) who identified the phenomenon, only Rouet et al. (1996, 1997) asked students to engage in identifying inconsistencies. Rouet et al. used a very broad measure, prompting students to ask for additional information to address a lack of information about the provided inconsistency. The low incidence of recognizing inconsistencies could be
attributed to the vague task demands in prior studies. In the current study, a researcher-designed measure of recognizing inconsistencies, called Listing Inconsistencies, was utilized. Participants were explicitly asked to list any inconsistencies they noted after reading multiple documents about the Battle of Lexington, so the task demands were explicit.

Although previous research has suggested that inexperienced readers and writers do not notice inconsistencies (Baker, 1985; Garner, 1981; Otero & Kintsch, 1992; Scardamalia & Bereiter, 1991), the results from this study indicate that developmental-level students are able to identify implicit inconsistencies. Postintervention, there were fewer participants who identified zero, one, or two inconsistencies. In addition, one student had reached beyond the highest number of inconsistencies identified at pretest by posting a 7 on the Listing Inconsistencies measure. These results indicate that several students improved their ability to identify inconsistencies on the posttest. Additionally, more participants identified more than three inconsistencies postintervention. The change in categorization from Low Use at pretest (6 participants) to High Use at posttest (14 participants) was statistically significant.

One reason for the high number of inconsistencies recognized preintervention—74 inconsistencies total—may be the explicit request to identify inconsistencies. As noted above, this measure was designed to explicitly ask students to identify inconsistencies. Without prompting, these participants may have behaved more like the participants in prior research by not noticing inconsistencies. However, when directed to look for inconsistencies, they were able to identify inconsistencies. This may suggest
that they have procedural knowledge—they know how to identify inconsistencies—but they do not independently apply their knowledge.

*The Impact of the Instruction Intervention on the Use of Evaluative Heuristics*

Experienced academic readers evaluate information, frequently using the evaluative heuristics—sourcing, corroboration, and contextualization—to understand events and evaluate the credibility of evidence (Wineburg, 1998). In contrast, students rarely engage the evaluative heuristics (1991). These inexperienced academic readers fail to use strategies, which could improve their understanding of the event or their evaluation of the information.

The goal of a developmental-level course is to provide students with the same strategies that experienced readers would use to ensure their continued success at the university. Therefore, the current study included explicit instruction in the existence and use of the evaluative heuristics in order to help developmental-level students achieve academic literacy.

*Students Rarely Evaluate Information*

The results of this study support prior findings that students rarely evaluate source information in a manner consistent with university expectations for academic literacy (Flanagin & Metzger, 2007; Metzger et al., 2003). Examples of student responses, particularly from the pretest for the Justify Trustworthiness task, demonstrate a failure to use rigorous criteria for evaluation. Student 216 wrote, “Since this is from a newspaper it should be very trustworthy.” The student did not offer any explanation as to why the newspaper (Document 5) might be trustworthy. Participants in this study also failed to explain their justification. Student 791 provides another example, when writing about
Document 4: “This is a diary so somewhat reliable but also was and could have been very biases” (pretest for the Justify Trustworthiness task). The student’s description could be applied to any of the documents and is unsupported. The student tells neither why it is reliable, nor why it might be biased.

**Superficial Criteria**

Inexperienced academic readers use superficial criteria to evaluate information (Bråten et al., 2009; Calkins & Kelley, 2007; Flanagin & Metzger, 2007; Kolstø, 2001; Metzger et al., 2003; Twait, 2005). The current study revealed similar patterns of superficial evaluation, especially at pretest for the Justify Trustworthiness task. As in previous research (Bråten et al., 2009), some students cited usefulness as the justification for trustworthiness. In the justification for Document 1 which the student ranked as 7 (least credible), Student 451 wrote, “I didn’t exactly understand what was being said. I do know that the battle was briefly explained, but I wasn’t able to use the document for anything” (pretest for the Justify Trustworthiness task). This seems to indicate that the document is not credible because it was not useful to the student. Relevance was another criteria for evaluation used by participants in the current study. Student 753 wrote, “This is not so reliable because of the fact that it was published after the battle, but still good enough since it has relevant information in it” (pretest for the Justify Trustworthiness task). This suggests support for prior findings from Kolstø (2001), Twait (2005), and Wiley et al. (2009) that students rely on superficial criteria, like relevance.

Similarly, Wineburg (1991) recorded examples of students using writer’s style (and their own ease of reading) as a proxy for credibility. Student comments suggested that they failed to evaluate the claim and were, instead, persuaded by how direct the
statement was. Participants in the current study also used writer’s style to evaluate credibility. Student 451 noted that Document 3 “had good details about what happened” (pretest for the Justify Trustworthiness task).

Need to Provide Instruction in Evaluating Information

Few intervention studies have been conducted despite a solid research base demonstrating that evaluating sources of information is a key aspect of academic literacy (Davis, 2003; Fister, 1992; Fitzgerald & Galloway, 2001; Kolstø, 2001; Twait, 2005). One of the few intervention studies reported some success teaching Wineburg’s evaluative heuristics, yet their findings are inconclusive (Nokes et al., 2007). Participants had been divided up into four treatment groups, two of which showed gains at posttest. Therefore, it was not clear whether the gains were the results of multiple text instruction or explicit instruction in use evaluative heuristics. The current study helps to address this thin spot in the literature by contributing findings about an intervention study that provided explicit instruction in using evaluative heuristics.

Developmental-level Freshmen Used Evaluative Heuristics

As with Nokes et al. (2007), the current study shows that participants are using evaluative heuristics. Nokes et al. present an instructional intervention focused specifically on evaluative heuristics. After the instructional intervention, high school students demonstrated greater use of evaluative heuristics (the sourcing and corroboration heuristic, in particular) in their writing. On the Decision Essay in the current study, participants used more evaluative heuristics postintervention than they had preintervention. Similar to Nokes et al., in the current study, the number of instances of sourcing heuristic use increased from 3 preintervention to 14 postintervention and the
The number of instances of the corroboration heuristic use increased from 4 preintervention to 9 postintervention.

The results of the current study show miniscule gains in contextualization usage on the Decision Essay at posttest. However, unlike Nokes et al., students did utilize the contextualization heuristic in the current study. Nine participants were classified as High Use based on their use of the contextualization heuristic preintervention in the current study. The number of contextualization heuristics used at posttest increased from 9 to 14. However, the number of participants categorized as High Use grew by only one.

In the current study, the Justify Trustworthiness measure also shows the participants’ use of evaluative heuristics. The total use of evaluative heuristics on the Justify Trustworthiness measure increased from 88 instances of evaluative heuristic use at pretest to 105 instances of evaluative heuristic use at posttest. However, the growth was not robust across the subtests.

Student Samples of Evaluative Heuristic Use

The next subsection presents several examples of student responses in order to illustrate the varied evaluative heuristic use employed by participants. As noted in the previous chapter, the student responses have been reported without changes for spelling, grammar, or punctuation. The first three are examples of the corroboration heuristic. The next two examples are of the contextualization heuristic. Several illustrations of sourcing heuristic have already been presented in the Scoring Anomalies section in Chapter 4. The illustrations for corroboration and contextualization are followed with additional illustrations from other participants.
**Illustrative Student Samples: Corroboration**

Although there were few uses of corroboration, there are some clear illustrations of different aspects. Student 660 presents a concise illustration of the *direct comparison* aspect of corroboration. On the pretest for the Decision Essay, Student 660 points out that Document 2 and Document 4 corroborate that the British were not intending to attack Lexington. The student is specific about which documents are being compared and about the information being corroborated.

Student 873 provides an example of the *claim of omission* aspect of corroboration on the pretest for the Decision Essay. Student 873 wrote:

The only thing that confused me about the documents was when they talked about Paul Revere. He was never mentioned in the other documents. Although Paul Revere is supposed to be this important figure in history that almost everyone should know, he was not once mention until the last document. The student is pointing out that six of the seven documents do not refer to Paul Revere. One point was awarded for the *claim of omission* aspect of corroboration.

Student 814’s response included the *analogy* aspect of the contextualization heuristic. Student 814 writes:

The British on the other hand, have plenty of inconsistencies in their texts. In a text from Lieutenant John Barker from the British Army he states that when getting to Lexington, they came to find 300 troops on the field. But the newspaper, The London Gazette that was published on June 10 1775 says that ‘several guns were fired upon the king’s troops from behind a stone wall.’ But
the first text makes no mention of a stone wall. Its as if the British writers are making new stories every time they print something.

This is a solid example of the direct contrast aspect of the corroboration heuristic. The student contrasts a specific detail of two documents and points out an inconsistency across documents. The student then explains how that impacts the credibility of the information.

Illustrative Student Samples: Contextualization

In prior research, inexperienced readers and writers rarely demonstrated the use of the contextualization heuristic, which is commonly used by experienced readers (Nokes et al., 2007; Wiley et al., 2009). Although in the current study, the use of contextualization was not huge, it was used more frequently than prior research would indicate. On the pretest of the Decision Essay, Student 873 demonstrated the use of the cultural setting awareness aspect of contextualization. Student 873 wrote, “The people of Lexington hear that British troops are coming so they arm themselves for protection because they were scared. It is human nature to go on the defensive when they see that harm is going their way.” The student explains the emotional underpinning of the colonists’ actions. The student was awarded a point for the cultural setting awareness aspect of contextualization.

Although Student 814’s example presents a misreading, the student provides an Analogy to explain the lack of credibility. Student 814 wrote:

I feel that the Americans stories were a lot more reliable because their sources were from people who were actually there. The Americans (confederates) have at the most three first-hand accounts whereas the British only has one. But, in reality
no one will really know what really happened that day. Its as if this story is like a
game of telephone in regards to history. The next person who documents it may
alter it a little bit and the next after that. By the time you read the text the event
has altered so much that the whole thing could be false” (posttest for Decision
Essay).

The student called the colonists the “confederates,” combatants in the Civil War a century
later rather than the Revolutionary War. The student also misattributes the documents.
In the document set, Document 2 is a first-hand account of the battle from the American
perspective, while Document 4 is a first-hand account of the battle from the British side.
These are the only first-hand accounts of the battle. This misreading notwithstanding, the
student presents an analogy—playing telephone—to explain how events are reshaped in
the retelling over time, thus lessening the credibility of sources that are created further
from the moment of the event.

Illustrative Student Samples: Other Interesting Results

Another participant whose use of evaluative heuristics increased postintervention,
Student 684 demonstrates improvement in the use of the contextualization heuristic. On
the pretest for the Decision Essay, Student 684 wrote, “Not based on any of the articles, I
believe they [the British] fired first because they came in pursuit of a battle.” This
response does not include any support for the writer’s claim. Postintervention, Student
684 wrote, “What I want to say is that first of all the British are trying to colonize or
move/take over an area. I’m sure they knew that the patriots, farmers, and those living in
the area would not be happy with that, thus they would have to entice some violence in
order to move in, which is why they could have fired the first shot” (Decision Essay).
The student not only explains why, but does so by using the historical awareness aspect of contextualization.

On the posttest Decision Essay, Student 704 wrote her way into an argument, using the contextualization heuristic:

I believe that the regulars fired on Lexington first because in my opinion I don’t feel Major Pitcairn’s statement was strong enough to prove his men didn’t fire first. To me, I feel the regulars fired first, just like they had planned to do because after all, they’re the ones that marched to Lexington. I believe their intention was to start this battle which is why they chose to march to Lexington to perhaps catch them off guard and charge at them first. I also feel they released the first fire because they are the ones who killed and injured the troops of Lexington. Lexington only wounded one of the regular’s soldiers which I believe is because they were unprepared and since they were fired at first, they didn’t have time to fight back.

The student keeps writing, until she gets an argument to stick. The student’s first point about Major Pitcairn’s statement is confusing. Then, she provides information about the intention of the British soldiers, without any support. However, the last section picks up steam. One point was awarded for her argument that the colonists were unprepared and did not have time to fire back. The scorers awarded one point for the other aspect of contextualization.

In addition to providing us with an example of the direct contrast aspect of corroboration, Student 814’s posttest for the Decision Essay illustrates why there were so few instances of evaluative heuristic use on the Decision Essay. Student 814 writes:
The British on the other hand, have plenty of inconsistencies in their texts. In a text from Lieutenant John Barker from the British Army he states that when getting to Lexington, they came to find 300 troops on the field. But the newspaper, The London Gazette that was published on June 10 1775 says that ‘several guns were fired upon the king’s troops from behind a stone wall.’ But the first text makes no mention of a stone wall. It’s as if the British writers are making new stories every time they print something.

One reason there were so few evaluative heuristics expected in the Decision Essay is because it takes several sentences to make and support each claim. Student 814 wrote 2.25 hand-written pages during the thirty minutes allotted to write the Decision Essay. The paragraph discussing direct contrast was slightly more than half a page long. The time constraints can also create length constraints, limiting the number of arguments students can include in their essay and, therefore, the number of evaluative heuristics they can include.

Satisfying the Objectives of the Current Study

The current study used a more rigorous coding scheme based on evaluation, but still yielded results with increased evaluative heuristic use at posttest over pretest, particularly for the Decision Essay. Although previous studies have focused on participants identifying information, rather than evaluating information (e.g., Britt & Aglinskas, 2002), this study was designed to measure evaluation of information rather than just identification. The instruction was designed to help participants gain not just procedural knowledge, but also conditional knowledge. Participants were required to offer an argument and support, rather than the less cognitively taxing requirement of
identifying information. Although the results in this study did not reach statistical significance, there was slight improvement for in the number of evaluative heuristics students were credited with postintervention. Developmental-level freshmen are capable of meeting rigorous academic literacy expectations and instruction helps them to do so.

Another goal of the study—to help students build conditional knowledge—was not fully met. The few experimental or quasi-experimental studies addressing the efficacy of possible interventions have been limited to focusing on building procedural knowledge, not the conditional knowledge that developmental-level students need to be successful at the university. The instructional interventions have focused on helping students identify information (procedural knowledge), not evaluate or apply information (conditional knowledge).

Previous research with instructional interventions has not been successful in helping students to use the contextualization heuristic, suggesting that students are not able to access contextual knowledge and are, therefore, not using the most helpful aspects of the heuristics in a given situation (Britt & Aglinskas, 2002; Nokes et al., 2007; Wiley et al., 2009). Although participants in the current study did use more instances of contextualization, particularly in their Decision Essays, this instructional intervention still fell short. The low scores, such as a total of four instances of corroboration across all participants postintervention, suggest that students are not accessing all three of the evaluative heuristics. Therefore, students are not able to choose the best strategy for any given situation.
Research by Wasson (2001) highlights the difference between experienced (both expert and advanced) and inexperienced readers and writers. The disparity in the use of the contextualization heuristic provides insight into the vastness of the chasm: six inexperienced readers demonstrated a total of 9 instances of contextualization. The experienced readers were seven times more likely to demonstrate the contextualization heuristic than the inexperienced readers. This one finding from previous research helps illuminate the massive undertaking asked of students.

Another example from the Justify Trustworthiness task in the current study showcases the distance between the experienced academic reader and the inexperienced student. For this measure, students are asked to evaluate the credibility of each of the seven documents from the Battle of Lexington document set and to justify their evaluation of its credibility. There were 31 students each of whom were prompted to write justifications for 7 documents. Together that means that there were 217 opportunities to use evaluative heuristics. In fact, the opportunity is even greater than that, as a participant could use more than one evaluative heuristic to make a case for each of the 7 documents. Yet, all participants were credited with a total of 88 instances of evaluative heuristic use preintervention and a total of 105 instances of evaluative heuristic use postintervention. When explicitly prompted to evaluate information, students used evaluative heuristics less than half the time.

Summary

There seems to be consensus in the research literature that inexperienced readers do not notice inconsistencies. However, in the current study, participants demonstrated...
statistically significant improvement in recognizing inconsistencies postintervention. These results suggest that being more explicit about the task could be helpful both as an intervention and as a guide to more valid test construction.

The results of this study support several earlier findings from the research literature. Reinforcing what has been reported in previous studies, this study also found inexperienced readers and writers, who did not evaluate credibility (even when explicitly asked to do so) or who used only superficial criteria for evaluation. This study adds to the small number of intervention studies that have been conducted on the use of evaluative heuristics. In the current study, participants did use all three evaluative heuristics: sourcing, corroboration, and contextualization. There was growth in their usage between the pretest and posttest even though it did not reach statistical significance. Results indicated that providing more time, especially for writing the Decision Essay, might increase the number of evaluative heuristics students were able to include and might help clarify the scoring process.

Findings from this study show that student did evaluate information through the use of evaluative heuristics. Although the researcher had posited that increased usage of the contextualization heuristic would propel individuals toward flexible usage of the evaluative heuristics, that was not born out. One unique finding was that the inexperienced readers and writers in this study tended to use the contextualization heuristic more than in previous studies. From an instructional perspective, this is good news. However, the low usage of all heuristics across the board on both measures indicates that participants were not necessarily choosing the most effective option among the evaluative heuristics.
Participants increased their use of evaluative heuristics, but those results did not reach statistical significance. Looking at those finding within the context of the vast difference in skills, knowledge, and experience between academic literacy experts and first-year freshmen suggests that a four-week intervention may not provide enough time to take such large strides forward as a scholar.

Conclusions

One important finding is that developmental-level students, like other inexperienced students, do not use the same strategies as experienced readers. Previous research has indicated that inexperienced readers do not use the strategies, such as identifying inconsistencies and using evaluative heuristics, that experienced readers use to be successful in academic situations (Bråten et al., 2009; Britt & Aglinskas, 2002; Rouet et al, 1996; Wasson, 2001; Wiley et al., 2009; Wineburg, 1991). The findings of this study add further support to this claim. As already presented illustrations of participant pretest responses show, many students are not using evaluative heuristics. Instead, participants tended to use superficial criteria to assess credibility: Student 753 used relevance, Student 451 used writer’s style, and Student 451 used usefulness as the criterion. In addition, many participants failed to explain their justification, like Student 791.

Moreover, even when inexperienced readers did use evaluative heuristics they did so at much lower rate than more experienced readers. An example from prior research shows that one advanced reader employed 8 instances of corroboration, making 8 links across 4 documents, in an attempt to resolve a single inconsistency (Wineburg, 1998). In contrast, the inexperienced readers in this study rarely used the corroboration heuristic.
In fact, across 31 participants on two measures (the Decision Essay and the Justify Trustworthiness task) on pretest and posttest combined only 21 instances of corroboration in total were recorded. Findings from this study indicate that inexperienced students frequently failed to use the academic literacy strategies that could help them be successful at the university.

A key finding is that the results from this study indicated that participants improved in their ability to recognize inconsistencies, but experienced only minor improvement in their use of evaluative heuristics. Prior research has suggested that inexperienced readers and writers do not notice inconsistencies (Otero & Kintsch, 1992; see also Scardamalia & Bereiter, 1991; Baker, 1985; Garner, 1981). Analysis of the data collected in this study suggested that participants increased their ability to identify inconsistencies within or across multiple texts. Preintervention, participants identified 74 inconsistencies ($M = 2.39, SD = 1.6$), with six participants being categorized as High Use because they identified four or more inconsistencies. Postintervention, participants identified 102 inconsistencies ($M = 3.29, SD = 1.83$), with 14 participants being categorized as High Use because they identified four or more inconsistencies. Postintervention, there were fewer participants who identified zero, one, or two inconsistencies. Moreover, one student had reached beyond the highest number of inconsistencies identified at pretest by posting a 7 on the Listing Inconsistencies measure. The change in categorization from Low Use at pretest (6 participants) to High Use at posttest (14 participants) was statistically significant. The results from this study indicate that developmental-level students are able to identify implicit inconsistencies.
In contrast, analysis of the data from this study suggests that although participants increased the number of evaluative heuristics they used in writing, the gains were not statistically significant. At posttest, participants demonstrated an increased use of evaluative heuristics in writing for each subtest of the Decision Essay. Despite the increase in usage and the increase in number of participants categorized as High Use at posttest, none of the results reached statistical significance.

At posttest the gains in evaluative heuristic usage in reading were minimal and none of the results from the Justify Trustworthiness task reached significance. Although the data indicate that participants increased in their ability to identify inconsistencies, the change in evaluative heuristic usage failed to reach statistical significance.

Interestingly, even though the participants in this study did not report a high level of familiarity with the topic, they still were able to contextualize. The inexperienced readers in this study reported a somewhat low level of familiarity with the Battle of Lexington on the Topic Familiarity measure. Participants indicated that they had studied the Battle of Lexington as part of a class a mean of 2.52 times ($SD = 1.34$). Participants rated their mean familiarity with the topic as 1.71 ($SD = 0.53$) or Somewhat Familiar with the Battle of Lexington.

However, participants in this study demonstrated 32 instances of contextualization, a surprising finding. This finding seems to run counter to previous research in which demonstrations of contextualization by inexperienced readers were rare (Nokes et al., 2007; Wasson, 2001; Wineburg, 1991). On the pretest for the Decision Essay, 9 participants used the contextualization heuristic one time, meaning that all 9 were classified as High Use. At posttest, 4 participants used contextualization two times,
while six participants used contextualization one time, meaning that 10 participants were classified as High Use. On the pretest for the Justify Trustworthiness task, 2 participants used the contextualization heuristic two times, meaning that 2 participants were categorized as High Use. At posttest, three participants used the contextualization heuristic one time while 2 participants used the contextualization heuristic two times, meaning 5 participants were classified as High Use. While the use of contextualization is not huge, it is surprising since it requires the reader to bring knowledge to the text and participants are saying that they do not have much knowledge of the topic to bring with them to reading these texts. Yet, the developmental-level freshmen in this study were able to leverage what knowledge they do possess to contextualize.

The essay-writing task (the Decision Essay) seemed to elicit more examples of contextualization. This may be due to the format of the measure which allows students to further develop their thinking. The Justify Trustworthiness task only provides the source information, similar to a citation, so there is not much information available. Plus, participants are asked to write one or two sentences in relatively small boxes that would not allow for lengthy ruminations. Illustrations of participants’ responses suggest that some writers write themselves into analysis, sometimes with underdeveloped points discarded along the way, as was the case with Student 704’s Decision Essay at posttest. Her first point is confusing and her second point is unsupported, but her third point is better developed with analysis. From information about the colonists firing fewer shots, she infers that the colonists were unprepared for battle. Therefore, she suggests they were not intent on engaging the British, or they would have been ready. Thus, she finds support for the British firing first on the unprepared colonists. The Decision Essay
allowed her to try out several different arguments before she was able to develop convincing support for one. Similarly, Student 814’s Decision Essay at posttest shows that explaining how the corroboration heuristic impacts credibility takes several sentences and time to write it. The Justify Trustworthiness task limits both the time available to develop arguments and the space to try out and explain different potential points.

Although participants used the contextualization heuristic, which the researcher had posited would lead to conditional knowledge (e.g., flexible application of the evaluative heuristics), the findings indicate that conditional knowledge was not attained. Previous research indicates that sourcing is the most commonly utilized heuristic, while instances of the contextualization heuristic are rare (Britt & Aglinskas, 2002; Nokes et al., 2007; Stahl et al., 1996; Wiley et al., 2009). One surprising finding was that 9 participants used the contextualization heuristic at pretest on the Decision Essay and were, therefore, categorized as High Use. Postintervention 10 participants were categorized as High Use. The instructional intervention seems to have had minimal impact on participants’ use of contextualization.

However, the researcher had posited that increasing use of the contextualization heuristic would help students move towards mastery of the evaluative heuristics. Contextualization seems like the most challenging heuristic because the student must bring in their prior knowledge and we know from prior research that inexperienced readers struggle to build background knowledge from reading (Scardamalia & Bereiter, 1991). Since use of contextualization was so rare in previous students, the researcher theorized that increased instruction in contextualization would allow developmental-level freshmen a full complement of strategies to choose from to resolve inconsistencies,
increasing the likelihood that participants could choose the best evaluative heuristic for each situation. Although participants in this study demonstrated an unexpectedly high number of uses of the contextualization heuristic, only about one-third of participants were classified as High Use for contextualization at posttest.

The lack of research in this area may have contributed to flawed scoring ranges which the researcher used to guide categorization. This instructional intervention was designed to help participants gain not just procedural knowledge (how to use the evaluative heuristics), but also conditional knowledge (selecting the most appropriate heuristic for any situation). One use of the contextualization heuristic, which was sufficient to categorize a participant as High Use, does not indicate intentional use, nor does it suggest that an inexperienced reader has chosen the best heuristic to use in any given situation. In fact, low usage across all evaluative heuristics, especially corroboration, suggests that participants were not using the best evaluative heuristic option or any evaluative heuristic at all. The total instances of evaluative heuristic use (sourcing, corroboration, and contextualization) increased from pretest to posttest, but this does not indicate students had mastered the use of evaluative heuristics nor does it suggest they were choosing the most effective heuristic for any situation. At pretest, 31 students demonstrated 16 total instances of evaluative heuristic use on the Decision Essay. The total increased to 37 total instances of evaluative heuristic use at posttest, showing that participants averaged one use of evaluative heuristics per essay. That is improvement, but it does not match expert-levels of use as presented by Wineburg (1991, 1998).
On the Justify Trustworthiness task, which explicitly asked participants to evaluate the credibility of each document, participants \((N = 31)\) demonstrated 88 total instances of evaluative heuristic use at pretest. If every student had used one evaluative heuristic on each of the seven documents, there would have been 217 instances of evaluative heuristic use. Even if they had only demonstrated one instance of evaluative heuristic use on five of the seven documents, that would have totaled up to 155 uses of evaluative heuristics. At posttest, participants demonstrated 105 total instances of evaluative heuristic use, which is still below the conservative estimate of 155 uses. These low total numbers make it clear that participants did not use evaluative heuristics to evaluate the credibility of every document. This low use suggests that students are not choosing the best evaluative heuristic for each situation. Therefore, conditional knowledge was not attained.

Although previous studies set a lower bar for students by asking them to identify rather than evaluate information, participants in this study demonstrated comparable levels of achievement on a more complex task. Several researchers used less rigorous coding schemes (Britt & Aglinskas, 2002; Rouet et al., 1996; Wasson, 2001; Wiley et al., 2009). For example, Britt and Aglinskas (2002) used a software template for filling in source characteristics as the instructional intervention and the measure of heuristic use. That intervention was somewhat successful in helping students identify the type of information (procedural knowledge) used by experienced readers, but did not address students’ use of that information to resolve an inconsistency (conditional knowledge).

Wasson (2001), who used less rigorous criteria for Wineburg’s heuristic, found that inexperienced readers \((n = 6)\) demonstrated 29 instances of sourcing heuristic use on
a think aloud ($M = 4.83$). This is a significantly higher mean score than participants in this study ($N = 31$) attained on the posttest for the Decision essay: 14 uses of the sourcing heuristic ($M = 0.45$). However, a closer examination of Wasson’s data shows that 25 of the 29 instances of sourcing among novices were for “superficial sourcing,” meaning participants were identifying information rather than evaluating information. In Wasson’s study, inexperienced readers ($n = 6$) exhibited 4 instances of “deep sourcing” ($M = 0.67$) a closer match to the evaluation required in this study, which is more in line with the findings in this study ($M = 0.45$).

In another study of twenty high school history students, 5 instances of corroboration ($M = 0.25$) were recorded in essay writing tasks (Stahl et al., 1996). However, the coding scheme required only identification of information rather than evaluation. Students were credited with corroboration for mentioning two documents. Similarly, another study of inexperienced readers in science reported that 18 of 60 participants ($M = 0.3$) referenced corroboration on a justify trustworthiness task (Wiley et al., 2009). In this study, participants demonstrated 9 uses of corroboration on the posttest for the Decision Essay ($M = 0.29$). Even though the coding scheme for this study was more rigorous, the average use of the corroboration heuristic was consistent with studies that had broader definitions.

When the coding schemes for contextualization are consistently rigorous, participants in this study performed better than inexperienced readers in a previous study. On an essay-writing task, 7% of high school history students ($N = 246$) used the contextualization heuristic (Nokes et al., 2007). In this study, 10 out of 31 participants,
or roughly 30% of participants, demonstrated use of the contextualization heuristics on the Decision Essay at posttest.

The current study used a rigorous coding scheme based on evaluation, and findings indicate that despite setting a high bar, developmental-level freshmen met it at rates similar to those achieved in other studies.

Transporting the framework for evaluating sources from the discipline of history into other disciplines is a promising line of inquiry. Wineburg’s original research focused on identifying the differences between the historical thinking of novice and experienced readers (1991). However, the evaluative heuristics are a good way to evaluate information, not just historical information. The goal of academic literacy courses is to help students become familiar with the strategies more experienced readers use to attain success at the university. Examples of student writing from this study demonstrate the value of learning to use the evaluative heuristics and show how they are consistent with the goals of a developmental-level academic literacy course. The two examples presented below are from the Decision Essay at posttest. Both illustrations suggest that use of the evaluative heuristics helped these students to think analytically.

Student 814 presented a telephone analogy to explain how retelling events can reshape them over time and, thereby, negatively impact credibility. Creating an analogy represents complex thinking because the student must select a key feature of the event in the texts, compare it with their experience of the world, and recast it in terms that would be familiar to their reader.

Student 684 was able to look beyond just the events presented in the document set about the Battle of Lexington and take into account the bigger picture of colonization and
its effects. Previous research had indicated that inexperienced readers tend to take a piecemeal approach rather than casting a wide strategic net for any help to understand and evaluate textual information (Garner, 1981; see also, Bråten & Strømsø, 2011; Kletzien, 1991; Strømsø et al., 2003). Student 684 was not taking a limited, piecemeal approach to understanding the events portrayed in the document set.

In both cases, use of the contextualization heuristic allows students to demonstrate the evaluative analysis that is a goal of academic literacy courses. This higher-order thinking will serve students well throughout their university careers.

Another important finding from this study is that there is a need to intervene more. Participants showcased several instances of evaluative heuristic use that demonstrates analytical thinking consistency with university-level expectations. Student 814 provided an illustration of close reading by using a linguistic feature (“I believe”) to evaluate credibility on the Justify Trustworthiness task at posttest. Student 684 showed improvement from baseless opinion preintervention on the Decision Essay to use of the historical awareness aspect of contextualization to justify her position postintervention. These two examples, along with the illustrations of complex analytical thinking presented in the two illustrations in the section above show the power of evaluative heuristics to help developmental-level freshmen to convey the complex thinking expected of advanced academic readers and writers.

However, there are also numerous examples of participants not engaging in those expert behaviors. For example Student 329 failed to explain his reasoning, writing only “President Warren could have altered the story to make it seem more ideal or to just cover up some facts” (pretest on the Justify Trustworthiness task). Student 527 attempted
to evaluate the credibility of Document 7, the textbook, but lacked clarity, leaving the reader to wonder how credible the student thought the document was and what the student meant when he wrote that the document “should be” credible (pretest on the Justify Trustworthiness task).

The evaluative heuristics are valuable tools for university students, but the chasm between inexperienced and advanced readers identified by Wasson (2001) and Wineburg (1991) remains. Despite improvement in the total number of evaluative heuristics used, developmental-level freshmen need additional instruction in and practice with evaluative heuristics in order to be more like successful, experienced academic readers and writers.

The final significant finding from this study is that data analysis, especially scoring will need to be refined for future investigations. The Scoring Guide may be adequate for tasks limited to identifying information, but the high number of scoring anomalies suggests that it does not provide sufficient guidance for scoring when the task requires evaluating information. For instance, in accordance with the Scoring Guide, Student 753 was not awarded a point on the pretest for Justifying Trustworthiness task for pointing out the absence of a publication date negatively impacts credibility. Seeing what information is not available is the type of evaluation that experienced readers should engage in. One specific change to the Scoring Guide that should be considered is to clarify that the absence of information can affect trustworthiness.

In another instance, Student 111 made use of background knowledge acquired in his History 120 class to evaluate information during the posttest for the Decision Essay. The Scoring Guide specifically disallows using background knowledge and pinpoints knowledge gained from a previous class as a non-example of corroboration. At the
university, students are expected to build their background knowledge through coursework and to apply that background knowledge. Therefore, another specific change to the Scoring Guide that should be considered is allowing points to be awarded for corroborating information with specific, credible outside sources, like other courses.

In addition to issues with the adapted Scoring Guide, the scoring ranges used to classify participants as High Use and Low Use should be revisited. For instance, based on the findings the corroboration cut score for the Justify Trustworthiness task may need to be lowered from three or more instances of corroboration being the baseline for classification as High Use to two or more instances being the baseline for classification as High Use.

However, an even better potential solution may be to design a hybrid assessment that measures identification of information as many previous studies did AND measures evaluation of information. This might address some of the questions surrounding the scoring anomalies and allow a more concrete foundation for the scoring guide revisions. In any case, the results of this study suggest that the data analysis procedures need to be strengthened.

Limitations

There were several limitations to this study. Threats to validity included the small sample size, the use of the same test preintervention and postintervention, issues with data analysis, the lack of measure of fidelity of treatment, and time constraints. The pre-experimental one group pretest/posttest design did not include a control group who completed the pretest and posttest without experiencing the instructional intervention. Because participants were completing the same Multiple Text Tasks with the same
document set at pretest and posttest, it is possible that repeated exposure to the measure could have improved student scores, which would present a threat to validity. The researcher did not include a measure of fidelity of treatment such as observing the instruction. Therefore, the researcher cannot be certain that the intervention was performed exactly as designed.

Issues with data analysis also impact the validity of the study. Chapter 4 outlined some of the scoring anomalies. The scoring guide, though seemingly detailed, was not sufficiently detailed to guide the scores; additional decision rules were created ad hoc and still many decisions needed to be hashed out between the scorers, leading to subjective decisions. Chapter 3 discussed the challenges the researcher faced to set scoring ranges to categorize participants as Low Use or High Use on each measure. The literature did not provide an adequate road map, so the researcher had to make several subjective decisions.

Time constraints may also have limited the effectiveness of the intervention. This study attempted to deliver a lot of instruction into a four-week window. It is difficult to change student behaviors in such a short time. With the compact time frame, there was inadequate practice time for mastering new skills, especially with corroboration. In addition, allowing participants to be absent three times before dropping them from the study meant that some participants may have missed three of the 10 class periods allotted for this study.

There were also threats to the generalizability of the findings. The small sample size and previous use of the difficulty paper may limit generalizability. This study utilized a relatively small sample size. The two sections (N = 31) represented five
percent of the developmental-level integrated reading and writing sections at the research site. In addition, the number of participants just reached the minimal, recognized sample size of thirty participants for research (Creswell, 2008). The small sample size limits the generalizability of findings from this study to a larger population. In addition, the students in the sample were familiar with a single-text version of the difficulty paper assignment from the fall semester. Although the difficulty paper is a common assignment at the research site, it is not a common assignment in other university settings. Therefore, other populations may need more time to master the difficulty paper assignment before benefitting from the practice in identifying inconsistencies and planning to resolve them.

Implications for Future Practice

This study provides data on an intervention to help inexperienced readers become more successful academic readers. Although descriptive research has identified the skills and knowledge that inexperienced readers lack (e.g., Bråten et al., 2009; Wineburg, 1991), few intervention studies have been conducted to find ways to meet these students’ instructional needs. This study contributes to the literature on possible instructional interventions. The modest success of the recognizing inconsistencies instruction make it a starting point for other instructional interventions.

Prior intervention research has overlooked the recognition of inconsistencies which is the key to successful reading at the university level (Britt & Aglinskas, 2002; Nokes et al., 2007; Wiley et al., 2009). Once an inconsistency is identified, the reader can then specify what information he or she needs and devise a plan for resolving the inconsistency. Noticing an inconsistency is the trigger for deploying conditional
knowledge. The results suggest that the intervention worked. Offering explicit instruction in recognizing inconsistencies and providing opportunities to practice identifying inconsistencies could be incorporated into the curriculum of developmental-level academic literacy courses.

A more specific implication for instruction is that the Difficulty Paper should be adopted. The difficulty paper assignment that was modified for use with multiple texts is one instructional tool for helping inexperienced readers learn to recognize inconsistencies. Therefore, the Difficulty Paper assignment should be considered for inclusion in courses focused on helping developmental-level freshmen achieve academic literacy.

Since evaluative heuristic use did increase at posttest, especially on the Decision Essay measure, one implication for practice is to provide instruction in the evaluative heuristics. Instruction could focus on declarative and procedure knowledge of evaluative heuristics and provide opportunities for practice. The student samples of evaluative heuristic use from this study show inexperienced readers/writers analyzing information. An approach like evaluative heuristics instruction that encourages students to engage with texts and think critically should be considered for inclusion in a curriculum targeted toward inexperienced readers and writers.

Another implication would be to increase the duration of instruction. Cognitive Flexibility Theory posits that students move from inexperienced to advanced and finally to expert. Educators should recognize that the developmental trajectory for moving from inexperienced to experienced may be challenging and time consuming. The results of this study suggest the participants improved in using evaluative heuristics to resolve
inconsistencies. However, the failure to reach statistical significance suggests that a more
time-intensive intervention is needed.

One specific instructional practice that might help students to use more evaluative
heuristics would be to help them see that experts not only use the evaluative heuristics,
but use them exponentially more often than inexperienced readers. From the students’
perspective, using a single evaluative heuristic means that they are now using the new
strategy. They might not realize the frequency and flexibility with which experienced
readers/writers engage evaluative heuristic usage. Showing students examples of the
types of justifications that experts use and sharing the data that highlights the disparities
in experienced reader/writer usage of evaluative heuristics versus inexperienced
reader/writer usage of evaluative heuristics would help make the difference in amount of
use explicit for students. Increased awareness can help students monitor their use of
evaluative heuristics in relation to the end goal—fluid problem solving.

A final implication for improving the effectiveness of instruction would be to
devote more instructional and practice time to corroboration, in particular. Low rates of
corroboration usage suggest that more instruction should be focused on the corroboration
heuristic. The current intervention presented corroboration. However, instruction in both
sourcing and contextualization included multiple activities and additional practice time.
Increased instructional focus might help students gain mastery with corroboration. Once
students are more comfortable with corroborating information, they may be able to be
more selective about which heuristic they use, eventually attaining the flexible
application that experts enjoy.
Recommendations for Future Research

There remains a need for research that examines interventions to help students learn to identify inconsistencies across multiple texts. This study made some progress in this research area. However, additional exploration of the topic would ensure effective instruction.

Previous research tended to present participants with researcher-selected inconsistencies. The results of this study indicate that participants can become proficient at self-identifying inconsistencies, a step towards independent reading. Therefore, one implication for future research is to include measures of inexperienced readers/writers’ ability to identify inconsistencies.

Another related recommendation for future researchers is being more explicit about the task they are measuring, like explicitly asking students to list inconsistencies. Previous measures for noting inconsistencies have been vague (e.g., Rouet et al., 1996). In this study, the researcher-designed measure—Listing Inconsistencies—was explicit about the task students were asked to do. Being more direct will increase the validity of the scores.

Further study of the Listing Inconsistencies measure would strengthen research in the area of recognizing inconsistencies within and across multiple texts. Specifically, future researchers might explore the value of recording inconsistencies during the document reading and notetaking period, rather than asking participants to list inconsistencies after having read and taken notes on the document set. This approach to measurement would be more consistent with how experienced readers read.
Further study is needed to understand the unique needs of developmental-level university freshmen. Previous multiple studies have drawn samples from high school students for whom building procedural knowledge may be more appropriate (Britt & Aglinskas, 2002; Nokes et al., 2007; Wineburg, 1991) and undergraduates who could be upper-classmen already apprenticed to the advanced academic literacy expectations of the university (Bråten et al., 2009; Wiley et al., 2009). Thus, upperclassmen may have already acquired some strategies related to the evaluative heuristics. First-year students identified as developmental remain an understudied, yet needy, population with regard to multiple text reading behaviors. Although this study provided some information about this group, many questions remain. Further study is needed to understand the unique needs of this population when reading multiple texts.

Specifically, future research might look at the effects of the language background of the participants. In this study, information on language background was collected to describe the sample. This population brings diverse language experiences to the college classroom. It would be interesting to explore ways that language background might interact with this instructional intervention. For instance, the data from this study could be reanalyzed comparing groups with different language background. Another option would be to conduct a study with a larger sample to see how strategies, like identifying inconsistencies and using evaluative heuristics, compare across groups of participants with differing language backgrounds.

Further study is needed to understand why the inexperienced college students in this study used the contextualization heuristic more than previously studied populations. This study found that these students did use the contextualization heuristic, even before
the intervention began. Further study is needed to understand why the developmental-level freshmen in this study used the contextualization heuristic more than participants in previous studies.

Another area for future research is in improving the intervention in order to foster conditional knowledge. In this study, developmental-level university freshmen improved in their recognition of inconsistencies and improved slightly in the number of evaluative heuristics they used postintervention. However, conditional knowledge was not achieved. A suggestion for future research would be to replicate this study with a longer instructional intervention to allow participants more opportunities to practice with these new strategies.

One implication for future research is to investigate why so few evaluative heuristics were used on the Decision Essay measure. One possible line of inquiry is the time constraints for the Multiple Text Tasks. Participants were given 30 minutes to write an essay. Although the time constraints in some studies were even more stringent (e.g., Rouet et al., 1997), studies of experienced readers and writers afforded more time to work through the documents and the tasks (e.g., Wineburg, 1998). In Wineburg’s study, the advanced reader spent nearly two hours carefully examining documents to make a decision about President Lincoln’s stance (1998). For inexperienced academic writers, 30 minutes is a tight window to decide what they want to say, get all their ideas down on paper, and revise for clarity. The time constraints could also have contributed to the difficulties encountered when scoring the measures. Perhaps with more time, students could more explicitly explain their reasoning. Future studies could be designed to compare results based on different time allotments.
Certainly, the scoring was messy which suggests that the Scoring Guide needs to be revised. Because the researcher was looking for evaluation, which is amorphous, even the very detailed rubric and scoring guides were of limited utility. Further research could investigate additional markers to be used to determine credibility (i.e., decision rules) to help clarify the awarding of points. Another avenue for future research would be to do an analysis of the students’ responses looking for patterns that might help refine the scoring guide.

Revisions to the instrumentation should also be considered. The student samples included in the Discussion of the Findings and the Scoring Anomalies sections both provide multiple examples of students misreading the documents. The Scoring Anomalies section, in particular, chronicled the difficulties the scorers encountered in deciding what the student was arguing for. One recommendation for future measurement of the phenomenon is a hybrid assessment that included identifying information and evaluating information. A measure that asks participants to first identify information and then evaluate it would help to ensure that participants understand the facts. In addition, awarding points for identification of factual information would provide the scorers with a factual foundation for parsing the participants’ explanations of trustworthiness. However, the measure should retain the evaluative component, like the current measures for using evaluative heuristics because that is the better match for the rigors of college-level coursework. Future studies might pilot a hybrid measure.

Another recommendation is to revise the cut scores for being classified as High Use or Low Use. In this study, participants needed only to demonstrate one use of contextualization to be considered High Use. Although it is wonderful that so many
students were able to show evidence of using the contextualization heuristic, the high number of individuals who were classified as High Use at pretest indicates that the classification needs to be more selective.

Future researchers might also consider a different data analysis strategy. In this study, participants were classified as High Use or Low Use based on the number of inconsistencies identified or the number of evaluative heuristics used. This data analysis procedure was selected because of the small sample size and the small range of potential scores on each of the measures. Repeating this study with a larger sample would yield information about the utility of this data analysis method. A few recommendations mentioned above have focused on better instrumentation, including a hybrid assessment that would yield a larger potential scoring range, thus allowing for more differentiation between performances.

One implication for advancing understanding of academic literacy instruction is to use Cognitive Flexibility Theory (CFT) as a framework for research. CFT takes into account the complexity of the academic literacy expectations university students face. It provides a lens for seeing challenges as a natural part of learning at the university. CFT offers guidance for developmentally appropriate practice. As such, one recommendation for future research would be to adopt CFT as the theoretical framework.

This study explored the value of Wineburg’s (1991, 1998) evaluative heuristics as academic literacy strategies to be used across domains. Most of the previous multiple text research had focused on the domains of history (e.g., Nokes et al., 2007) and science (e.g., Wiley et al., 2009), so there was little empirical research addressing whether the benefits of evaluative heuristic usage would generalize to other domains. In the current
study, the underlying goals—recognizing inconsistencies and using evaluative heuristics to resolve them—were found to be compatible with the goals of an academic literary course. Future research could investigate the appropriateness of the evaluative heuristics to other content areas.

The larger question remains unanswered: How to help students attain flexible strategy deployment? Using more contextualization heuristics in itself was not the answer to being more flexible in applying the evaluative heuristics. Further research into how to support student learning around flexible strategy deployment is needed.
REFERENCES


Intersegmental Committee of the Academic Senates of the California Community Colleges, the California State University, and the University of California. (2002). *Academic literacy: A statement of competencies expected of students entering California’s public colleges and universities*. Sacramento: Academic Senate for California Community Colleges.


Strømsø, H. I., Bråten, I., & Britt, A. M. (2010). Reading multiple texts about climate change: the relationships between memory for sources and text comprehension. Learning and Instruction, 20, 192-204.


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, 513-527. doi: 10.1016/j.learninstruc.2007.11.001


Appendix A

Data Collection
<table>
<thead>
<tr>
<th>Research Question</th>
<th>Data Source</th>
<th>Data Collection</th>
<th>Data Analysis</th>
<th>What the Data Will Reveal</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the effect of an explicit academic literacy instructional unit on the number of inconsistencies identified by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Listing Inconsistencies measure?</td>
<td>Listing Inconsistencies</td>
<td>Pretest: Number of inconsistencies listed for the Battle of Lexington document set</td>
<td>McNemar’s Test for Significance of Change</td>
<td>If participants are able to identify more inconsistencies, like experienced readers do.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Posttest: Number of inconsistencies listed for the Battle of Lexington document set</td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is the effect of an explicit academic literacy instructional unit on the number of evaluative heuristics used in writing by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Decision Essay measure?</td>
<td>Decision Essay</td>
<td>Pretest: Number of heuristics used in the Decision Essay for the Battle of Lexington document set</td>
<td>McNemar’s Test for Significance of Change</td>
<td>If participants use evaluative heuristics in writing tasks like experienced readers do.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Posttest: Number of heuristics used in the Decision Essay for the Battle of Lexington document set</td>
<td></td>
<td>Which evaluative heuristics are participants using like experienced readers/writers?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is the effect of an explicit academic literacy instructional unit on the number of evaluative heuristics used in reading by developmental-level freshmen as measured by differences between their pretest and posttest scores on the Justify Trustworthiness measure?</td>
<td>Justify Trustworthiness</td>
<td>Pretest: Number of heuristics used in the Justify Trustworthiness task for the Battle of Lexington document set</td>
<td>McNemar’s Test for Significance of Change</td>
<td>If participants use evaluative heuristics in writing tasks like experienced readers do.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Posttest: Number of heuristics used in the Justify Trustworthiness task for the Battle of Lexington document set</td>
<td></td>
<td>Which evaluative heuristics are participants using like experienced readers/writers?</td>
</tr>
</tbody>
</table>
Appendix B

Multiple Text Tasks, based on the Battle of Lexington document set
Multiple Text Tasks

Topic Familiarity

Today you will be reading about the Battle of Lexington which took place during the Revolutionary War. Please indicate your familiarity with this topic by circling your response to the two items that appear below.

1) Rate your familiarity with the Battle of Lexington.

Not at all Familiar  Somewhat Familiar  Familiar  Very Familiar

2) Please indicate the number of times you have studied the Battle of Lexington for a class.

Never 1 Time 2 Times 3 Times More Than 3 Times
Directions for the Multiple Text Tasks

This packet contains 7 documents that discuss the Battle of Lexington which was fought during the Revolutionary War. Read through the documents carefully and try to figure out what you think happened at the Battle of Lexington. You will be asked to write a short essay supporting your decision. You may take notes on your Notes Page to help you figure out what happened. You will have 24 minutes to carefully consider the information in the documents and to take notes. At the end of 24 minutes, this booklet will be collected and instructions for the next task will be distributed. You may keep your notes for the other tasks, but the documents will be collected.
In 1775, Benjamin Franklin was the colonial representative in London. After the events in Lexington and Concord, the Massachusetts Provincial Congress put together 21 sworn depositions about the events and sent them to Franklin with the following cover letter:

In Provincial Congress, Watertown  
April 26, 1775  

To the inhabitants of Great Britain:  
Friends and fellow subjects: Hostilities are at length commenced in the Colony by the troops under command of General Gage; and it being of the greatest importance that an early, true, and authentic account of this inhuman proceeding should be known to you, the Congress of this Colony have transmitted the same, and from want of a session of the honorable Continental Congress, think it proper to address you on the alarming occasion.

By the clearest depositions relative to this transaction, it will appear that on the night preceding the nineteenth of April instant, …the town of Lexington…was alarmed, and a company of the inhabitants mustered on the occasion; that the Regular troops, on their way to Concord, marched into the said town of Lexington, and the said company, on their approach, began to disperse; that notwithstanding this, the Regulars rushed on with great violence, and first began hostilities by firing on said Lexington Company, whereby they killed eight and wounded several others; that the Regulars continued their fire until those of said company, who were neither killed nor wounded, had made their escape.

These, brethren, are marks of ministerial vengeance against this colony, for refusing, with her sister colonies, a submission to slavery. But they have not yet detached us from our Royal Sovereign. We profess to be his loyal and dutiful subjects, and so hardly dealt with as we have been, are still ready, with our lives and fortunes, to defend his person, family, crown, and dignity. Nevertheless, to the persecution and tyranny of his cruel ministry we will not tamely submit; appealing to Heaven for the justice of our cause, we determine to die or be free.

• Joseph Warren, [President pro tem]
Document #2

We NATHANIEL MULLIKEN, PHILIP RUSSELL, [followed by the names of thirty-two other men present on Lexington Green on April, 19, 1775],… all of lawful age, and inhabitants of Lexington, in the County of Middlesex,…do testify and declare, that on the nineteenth of April instant, about one or two o’clock in the morning, being informed that … a body of regulars were marching from Boston towards Concord,…we were alarmed and having met at the place of our company’s parade [Lexington Green], were dismissed by our Captain, John Parker, for the present, with orders to be ready to attend at the beat of the drum, we further testify and declare, that about five o’clock in the morning, hearing our drum beat we proceeded towards the parade, and soon found that a large body of troops were marching towards us, some of our company were coming up to the parade, and others had reached it, at which time the company began to disperse, whilst our backs were turned on the troops, we were fired on by them, a number of our men were instantly killed and wounded, not a gun was fired by any person in our company on the regulars to our knowledge before they fired on us, and they continued firing until we had made all our escape.

- Lexington, April 25th, 1775, NATHANIEL MULLIKEN, PHILIP RUSSELL, [and the other 32 men] [Duly sworn to by 34 minutemen on April 25th before three justices of the peace]
Document #3

Major Pitcairn screamed at us: “Lay down your arms, you lousy bastards! Disperse, you lousy peasant scum!”...At least, those were the words that I seem to remember. Others remembered differently; but the way he screamed, in his strange London accent, with the motion and excitement, with his horse rearing and kicking...with the drums beating again and the fixed bayonets glittering in the sunshine, it’s a wonder that any of his words remained with us...We still stood in our two lines, our guns butt end to the ground or held loosely in our hands. Major Pitcairn spurred his horse and raced between the lines. Somewhere, away from us, a shot sounded. A redcoat soldier raised his musket, leveled it at Father, and fired. My father clutched at his breast, then crumpled to the ground like an empty sack...Then the whole British front burst into a roar of sound and flame and smoke.

- Excerpt from the novel, *April Morning*, by Howard Fast, published in 1961
Document #4

19\textsuperscript{th}. At 2 o’clock we began our march by wading through a very long ford up to our middles; after going a few miles we took three or four people who were going off to give intelligence; about five miles on this side of a town called Lexington, which lay in our road, we heard there were some hundreds of people collected together intending to oppose us an stop our going on; at five o’clock we arrived there, and saw a number of people, I believe between 200 and 300, formed in a common in the middle of the town; we still continued advancing, keeping prepared against an attack though without intending to attack them; but on our coming near them they fired one or two shots, upon which our men without any orders, rushed in upon them, fired and put them to flight; several of them were killed, we could not tell how many, because they were got behind walls into the woods; We had a man of the 10\textsuperscript{th} light Infantry wounded, nobody else hurt. We then formed on the Common, but with some difficulty, the men were so wild they could hear no orders; we waited a considerable time there, and at length proceeded on our way to Concord.

- Entry for April 19\textsuperscript{th}, 1775, from the diary of Lieutenant John Barker, an officer in the British army
Lieutenant Nunn, of the Navy arrived this morning at Lord Dartmouth’s and brought letters from General Gage, Lord Percy, and Lieutenant-Colonel Smith containing the following particulars of what passed in the Province of Massachusetts-Bay and several parties of rebel provincials…

Lieutenant-Colonel Smith finding, after he had advanced some miles on his march, that the country had been alarmed by the firing of guns and ringing of bells, dispatched six companies of light-infantry, in order to secure two bridges on different roads beyond Concord, who, upon their arrival at Lexington, found a body of the country people under arms, on a green close to the road; and upon the King’s Troops marching up to them, in order to inquire the reason of their being so assembled, they went off in great confusion, and several guns were fired upon the King’s Troops from behind a stone wall, and also from the meeting-house and other houses, by which one man was wounded, and Major Pitcairn’s horse shot in two places. In consequence of this attack by the rebels, the troops returned the fire and killed several of them. After which the detachment marched on to Concord without any further happening.

- Newspaper account from *The London Gazette, June 10, 1775*
Document #6

There is a certain sliding over and indeterminateness in describing the beginning of the firing. Major Pitcairn who was a good man in a bad cause, insisted upon it to the day of his death, that the colonists fired first... He does not say that he saw the colonists fired first. Had he said it, I would have believed him, being a man of integrity and honor. He expressly says he did not see who fired first; and yet believed the peasants began. His account is this—that riding up to them he ordered them to disperse; which they not doing instantly, he turned about to order his troops to draw out as to surround and disarm them. As he turned he saw a gun in a peasant’s hand from behind a wall, flash in the pan without going off: and instantly or very soon two or three guns went off by which he found his horse wounded and also a man near him wounded. These guns he did not see, but believing they could not come from his own people, doubted not and so asserted that they came from our people; and that thus they began the attack. The impetuosity of the King’s Troops were such that a promiscuous, uncommanded but general fire took place, which Pitcairn could not prevent; though he stuck his staff or sword downward with all earnestness as a signal to forbear or cease firing. This account Major Pitcairn himself gave Mr. Brown of Providence who was seized with flour and carried to Boston a few days after the battle; and Gov. Sessions told it to me.

• From the diary of Ezra Stiles, president of Yale College, entry for August 21, 1775
In April 1775, General Gage, the military governor of Massachusetts, sent out a body of troops to take possession of military stores at Concord, a short distance from Boston. At Lexington, a handful of “embattled farmers,” who had been tipped off by Paul Revere, barred the way. The “rebels” were ordered to disperse. They stood their ground. The English fired a volley of shots that killed eight patriots. It was not long before the swift-riding Paul Revere spread the news of this new atrocity to the neighboring colonies. The patriots of all of New England, although still a handful, were now ready to fight the English.

- From *The United States: A Story of a Free People*, a high school textbook by Samuel Steinberg, Allyn and Bacon, publishers
Listing Inconsistencies

Directions: Please list all the inconsistencies you noticed in the Battle of Lexington document set you read and studied. An inconsistency is a difference in fact or opinion across documents. You have 5 minutes to complete this task.

1. __________________________________________________________________________

2. __________________________________________________________________________

3. __________________________________________________________________________

4. __________________________________________________________________________

5. __________________________________________________________________________

6. __________________________________________________________________________

7. __________________________________________________________________________

8. __________________________________________________________________________

9. __________________________________________________________________________

10. __________________________________________________________________________
Decision Essay

Directions: Now that you have read the documents discussing the Battle of Lexington and decided what you think happened at Lexington Green on that April morning in 1775, you have a chance to convince me of your decision.

On the attached binder paper, write a 200-word draft essay that answers the question that appears below in bold. Your essay should explain your decision. Be sure to write about the documents.

You have 30 minutes to write your Decision Essay. Feel free to cross things out or make revisions, but do not worry about recopying. Please write in pen.

ESSAY PROMPT: Who fired first at the Battle of Lexington?
**Justify Trustworthiness**

**Directions:** In this task you will rank the trustworthiness of each document you read and explain your ranking. The chart below includes some information to remind you of each of the seven documents about the Battle of Lexington that you read earlier. There is a box to the right of each source for you to place your ranking of the document (from 1-7) on the basis of its trustworthiness. Give each document a number, assigning a 1 to the document you think is the most trustworthy and a 7 to the document you think is the least trustworthy. Beneath the source information is a place to write a sentence or two that tells *why* you assigned that rank to each source. You have 15 minutes to complete this task.

<table>
<thead>
<tr>
<th>Document #1</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>President pro tem of the Massachusetts Provincial Congress, Joseph Warren’s cover letter for the 21 sworn depositions he sent to Benjamin Franklin who was the colonial representative in London. The letter was dated April 26, 1775.</td>
<td></td>
</tr>
<tr>
<td><strong>Justification:</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Document #2</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexington, April 25th, 1775, NATHANIEL MULLIKEN, PHILIP RUSSELL, [and the other 32 men] [Duly sworn to by 34 minutemen on April 25th before three justices of the peace]</td>
<td></td>
</tr>
<tr>
<td><strong>Justification:</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Document #3</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excerpt from the novel, <em>April Morning</em>, by Howard Fast, published in 1961</td>
<td></td>
</tr>
<tr>
<td><strong>Justification:</strong></td>
<td></td>
</tr>
<tr>
<td>Document #4</td>
<td>Ranking</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Entry for April 19th, 1775, from the diary of Lieutenant John Barker, an officer in the British army</td>
<td></td>
</tr>
<tr>
<td><strong>Justification:</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Document #5</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newspaper account from <em>The London Gazette</em>, June 10, 1775</td>
<td></td>
</tr>
<tr>
<td><strong>Justification:</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Document #6</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>From the diary of Ezra Stiles, president of Yale College, entry for August 21, 1775</td>
<td></td>
</tr>
<tr>
<td><strong>Justification:</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Document #7</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>From <em>The United States: A Story of a Free People</em>, a high school textbook by Samuel Steinberg, Allyn and Bacon, publishers</td>
<td></td>
</tr>
<tr>
<td><strong>Justification:</strong></td>
<td></td>
</tr>
</tbody>
</table>
Appendix C

Scoring Guide for Listing Inconsistencies
Scoring Guide for Listing Inconsistencies

This guide will define and give examples of inconsistencies that participants may provide as responses on the Listing Inconsistencies measure for the Battle of Lexington document set. However, you should feel free to revisit the document set at any time to verify an inconsistency.

**An inconsistency is a contradiction between two or more interpretations or accounts, within or across documents.**

On the next page, you will find a list of 26 categories of inconsistencies. These are considered categories because there are several possible inconsistencies within each. For example, a participant could identify the inconsistency between Document 2 (drumbeats alerted colonists that British were coming) and Document 3 (drumbeats occurred during, not before the battle); the inconsistency between Document 2 (drumbeats alerted colonists that British were coming) and Document 5 (guns and bells alerted the British about the colonists’ presence); Document 3 (drumbeats occurred during, not before the battle) and Document 5 (guns and bells alerted the British about the colonists’ presence). The participant could earn a point for each contradiction explicitly identified for a total of 3 points.

Although participants can point out inconsistencies between specific documents (i.e., Document 2 says drumbeats alerted the colonists, but Document 5 says the ringing of bells alerted the British), they do not need to identify the documents that produce the inconsistency. *Drumbeat alert vs. no drumbeat* would also receive credit. Responses that indicate one document is inconsistent with other documents (i.e., *Only Document 7 brings up Paul Revere*) may be scored as identifying an inconsistency across documents.

However, each response should make clear that there is a contradiction. Participants may format their response as a question (i.e., *Was there a drumbeat?*) or a statement that points out the a difference or a contrast (i.e., *one document says drumbeat, another says bells*). Participants can point out categories of inconsistencies (i.e., *Were drums sounded during the battle?*) or specific instances of an inconsistency (i.e., *drumbeat alert vs. no drumbeat*). But, there must be an explicit inconsistency.

Participants can also receive credit for identifying an inconsistency within a single document (i.e., *In Document 1, would the colonists die fighting for or against the king?* OR *was someone arrested for flour possession?).

Participants should **not** receive credit for listing a fact or a detail (i.e., *flour?* OR *a drumbeat sounded during the battle* OR *Document 2 says there is a drumbeat*).
Procedure for Scoring the Listing Inconsistencies Assessment

1) Write your rater code at the top of each assessment.

2) Place a check (✓) to the left of any correctly identified inconsistency and an Ø to the left of any incorrectly identified inconsistency.

3) Count the total number of check marks; each check mark counts as one point.

4) Write the total number of points earned beneath your rater code.

26 Potential Categories of Inconsistencies for the Battle of Lexington Document Set

Time of day that the battle occurred
Location of the battle
Was there a wall or buildings on Lexington Green?
The number of colonists present
Did the colonists disperse? Why did they disperse?
Term used to refer to the colonists
Term used to refer to the British
Who fired first?
How many shots were fired?
How many were killed?
How many were injured?
Were any horses wounded?
Were drums sounded during the incident?
Were any other persons besides soldiers present on Lexington Green?
What was the British purpose for marching on/through Lexington?
What was the colonists’ intent?
Did the colonists flee?
Did the British issue commands? What were those commands?
What was the British demeanor?
Was General Gage there?
Was Major Pitcairn there?
Was Paul Revere involved?
How many depositions were there?
In Document 1, who does “you” refer to?
In Document 1, would colonists die for their king or their freedom from the king?
In Document 6, was someone arrested for having “flour”?
Appendix D

Evaluative Heuristics Scoring Guide & Rubric
THE EVALUATIVE HEURISTICS SCORING GUIDE

For Identifying Students’ Use of Heuristics in Writing and Reading

Adapted from Nokes et al., 2007

The Evaluative Heuristic Scoring Rubric is used to keep a tally of the number of times a student uses a particular evaluative heuristic in either the Decision Essay or the Justify Trustworthiness task.

**Instructions for Using the Evaluative Heuristics Scoring Rubric**

1) Record the assessment code

2) Circle your rater code

3) Circle the appropriate subtest: Decision Essay or Justify

4) When the student uses an evaluative heuristic, make a tally mark in the “tally of occurrences” column that corresponds with the aspect of the evaluative heuristics (i.e., Author’s credentials).

**Note:** A separate Evaluative Heuristic Scoring Rubric should be used for the Justify Trustworthiness task and the Decision Essay. When the student uses an evaluative heuristic for the Justify Trustworthiness task, record the number of the document to which the heuristic was applied in the “tally of occurrences” column.

5) After you have scored the entire assessment, count the tally marks and write the total for each aspect identified in the “total” column. Then, fill in the box with the subtotals for each evaluative heuristic (i.e., Sourcing). At the bottom of the page, record the total number of uses of evaluative heuristics.

**Instructions for Identifying Evaluative Heuristic Use**

This guide will define and give examples of each aspect of the evaluative heuristics that a student might use.

**SOURCING:**

Sourcing only occurs when the consideration of the source helps the individual make sense of the document—through improved understanding of the events or evaluation of credibility as evidence. If the student analyzes two documents together, give two marks. For example, a student might write “Both Document 2 and Document 7 are biased because the authors of both documents wanted to blame the other side for the event.” You would record a “2” & a “7” in the Author’s Motivation row. When evaluating the students’ essays, the following items will be viewed as evidence of the use of sourcing.
• **author’s credentials:** Any reference to the **occupation, profession, level of training, or other credentials of the author** of the document in order to suggest that the document is more or less reliable or in order to understand what the document says qualifies as sourcing. Examples: “Since Shaw was an officer in the British army, he would have known...”; “The historian who wrote this must have studied a lot to become a historian so...”

• **author’s motivation:** Any reference to why an author might have written the document in order to suggest that the document is more or less reliable or in order to understand what the document says, qualifies as sourcing. Examples: “Colonel Jackson had a lot to gain by telling his commander about his success, so he may have exaggerated...”; “The author was probably trying to convince people that the Americans did not start the battle, so he wrote...”

• **author’s participation:** Any reference to the author’s level of participation in an event to suggest that the document is more or less reliable or in order to understand what the document says, qualifies as sourcing: Examples: “Jones was a witness of the battle, so he knew what happened when he wrote.”; “Smith only heard about the incident by word of mouth, so he is less reliable than an eye-witness.”

• **other evaluations of the author:** Any other consideration of the author to suggest that the document is more or less reliable or in order to understand what the document says, qualifies as sourcing. Examples, “It sounds like the author wanted the reader to think that the battle was a good thing, but he doesn’t use a very good argument.”; “Because Simpson admits that he made mistakes, this letter seems more truthful.” “The author admits that he can’t remember, so...”

• **date of production:** Any reference to when a document was created, to suggest that the document was more or less reliable or in order to understand what the document says, qualifies as sourcing. Examples: “He didn’t write this until many years after the event, so he might have forgotten...”; “This was written in his journal the day of the event, so it was probably fresh on his mind.”

• **document type:** Any reference to the type of document in order to suggest that it is more or less reliable or in order to understand what the document says, qualifies as sourcing. Examples: “This statement was sworn before a justice of the peace, so it was probably truthful.”; “People usually write in their journals to keep a record for themselves, so it wouldn’t make sense for him to write lies in his journal.”
• other evaluations of document: Any specific statement giving a reason why the document is more or less reliable could be considered sourcing. Examples: “textbooks tend to exaggerate the good about a country and leave out the bad, so I really don’t trust this textbook account... Any other reference to the source suggesting that it is more or less reliable would qualify as sourcing.

Sourcing Reminders:

1) Any reference to eyewitness accounts or being there, counts as Author’s Participation

2) Any reference to the sworn statement, counts as Document Type

3) Misinterpretations supported by the text still count as sourcing. For example, a student would get credit for Author’s Participation if he/she said “General Gage lead the battle, therefore he witnessed it.” The documents only suggest that he lead the British troops, not that he was actually at the Battle of Lexington (but they don’t explicitly say he wasn’t).

NOT SOURCING:

Students should NOT be given credit for sourcing if:

• they show an awareness of an aspect of a text, but do not tell why the type of text is important (except in the case of an eyewitness account or sworn statement). For example, if they wrote “Document 3 is most reliable because it is a novel” but they do not explain why historical novels are reliable. However, if they give any legitimate reason, count it as sourcing.

• they make any general claim, even one that you might agree with, such as “media lies” without an explanation. However, if they explained why the media might lie in this case (e.g., the British newspaper wanted to appeal to British pride), count it as sourcing.

• they comment on the content of the passage. For example “this document is a good one because it has lots of details or gives good information.” Or “the newspaper account is not very good because it is hard to understand and very confusing.”

• they comment on the syntax of a passage. For example “this document has a lot of misspelled words in it.” Or “This document is one single run-on sentence.”
CORROBORATION:

An individual uses corroboration when he or she compares or contrasts information found in two or more specified documents. Credit for corroboration should only be given when it helps the individual make sense of the event. When evaluating the students’ essays, the following items will be viewed as evidence of the use of corroboration.

- **direct comparison**: When the writer makes a **direct connection between similar information** that was found in two or more documents, it qualifies as corroboration. Examples: “almost every one of the people said that the shot was fired from behind the stone wall, so I am pretty sure…”; “both Simpson and Smith wrote that…”

- **direct contrast**: When the writer points out **information that was different in two documents**, it qualifies as corroboration. Examples: “The textbook says that 8 Americans were killed, but the letter says that only 5 were killed…” “Jones’ account of what happened was very different from the others’ accounts.”

- **claim of uniqueness**: When the writer points out that **information was found in only one source**, it qualifies as corroboration. Examples: “Only Valdez wrote anything about hearing the command to fire. None of the others mentioned that.”; “The textbook was the only source that talked about how the Spanish felt.”

- **claim of omission**: When a person claims that a **source left out an important detail that was found in other sources**, it qualifies as corroboration. Examples: “It seems like an eye-witness would have written about hearing the command, yet Smith didn’t say anything about it…”; “Since Harper didn’t include any information about the bad effects of the program, you can tell that he is biased against…”

- **other**: If a writer makes any other connection between information in two or more different sources, it qualifies as corroboration.

NOT CORROBORATION:

Students should **NOT** be given credit for corroboration if:

- information from a document is being compared or contrasted with the **student’s background knowledge** rather than information from another text. For example, “document 2 disagrees with what I learned in 8th grade about…” Or “My dad is a police officer and what document 3 says is really true about police work”

- documents are linked in a **vague manner**. For example, “In the documents it says that…”
CONTEXTUALIZATION:

An individual uses contextualization when he or she discusses specific details about the event that helps him or her understand why or how the event took place. When scoring the students’ essays, the following items will be viewed as evidence of the use of contextualization.

- **time or location awareness**: If the writer demonstrates a keen awareness of the chronology of an event, or specific features of the physical location of an event, it qualifies as contextualization. If the writer includes an awareness of both time and place make two marks on the tally sheet. Examples: “Since the Boston Tea Party took place in New England in October, it was probably chilly…”; “Gettysburg is a hilly area with some forests around it, so it would have been hard to see enemy troops…”

- **cultural setting awareness**: If the writer demonstrates a keen awareness of the cultural values or common attitudes of the time period, or emotions that participants in an event may have been feeling, it qualifies as contextualization. Examples: “After marching all night the soldiers must have been tired, so they…”; “In the 20’s most people thought that a women’s place was in the home, so it is surprising…”

- **biographic awareness**: If the writer demonstrates a keen awareness of the values and influences surrounding the production of the text, especially those affecting the author it qualifies as contextualization. If the student questions whether a historian used effective methods to study an event before writing, it is contextualization. Examples: “This was written shortly after the Civil War during Reconstruction when most people, including historians had a pretty racist way of viewing life in the South, so…”; “The historian who wrote this didn’t have access to all of the information that has come out recently, so…”

- **historical awareness**: If the student demonstrates a keen awareness of the historical events (or people) surrounding/contributing to the events, it qualifies as contextualization. “The slavery debates in England riled up the landowners in the colonies”; “Since it was Roosevelt who the idea, and he was so popular at the time, it probably was accepted with little debate.” “I’m sure that as soon as Stanton walked into the room, the whole atmosphere changed and…”

- **linguistic awareness**: If the writer demonstrates a keen awareness of the different meanings of words over time, it qualifies as contextualization. Examples: “In Lincoln’s time it was generally accepted to call African Americans ‘Negroes’ so I don’t think he was being disrespectful when he said…”; “The word ‘misdemeanor’ meant something different when it was written in the Constitution, so…”

- **analogy**: If the writer attempts to connect with information in the past by comparing it to current events or personal experiences, it qualifies as contextualization. Example: “The debate over Prohibition was probably a lot like the debate today over legalizing marijuana, with more traditional people for Prohibition, and more liberal and younger people opposing it.”
• **other**: If the writer makes other attempts to understand an incident or a document by attempting to place themselves or their reader in the event, it qualifies as contextualization. Example: “We don’t shoot our own people.”

**NOT CONTEXTUALIZATION:**

Students should NOT be given credit for contextualization if:

• they inappropriately project today’s values or culture on the people of the past. For example, if a student argues that lots of women have short hair, so it shouldn’t have been shocking for a woman to get her hair cut in 1920, this is an inappropriate analogy and should not be counted as contextualization. This misrepresentation is known as “presentism.”

**Scoring Packet will include:**

Document Set
(Scoring Guide for Listing Inconsistencies)
(6 Sample Listing Inconsistencies assessments)
The Evaluative Heuristics Scoring Guide
3 Model Decision Essays with Evaluative Heuristics Rubrics
3 Sample Decision Essays to score together
3 Model Justify Trustworthiness tasks with Evaluative Heuristics Rubrics
3 Sample Justify Trustworthiness tasks to score together
Evaluative Heuristics Scoring Rubrics
<table>
<thead>
<tr>
<th>Evidence</th>
<th>Tally of occurrences (Document #)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOURCING</td>
<td></td>
<td></td>
</tr>
<tr>
<td>author’s credentials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>author’s motivation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>author’s participation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>other evaluation of author</td>
<td></td>
<td></td>
</tr>
<tr>
<td>date of production</td>
<td></td>
<td></td>
</tr>
<tr>
<td>document type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>other evaluation of document</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CORROBORATION</td>
<td></td>
<td></td>
</tr>
<tr>
<td>direct comparison</td>
<td></td>
<td></td>
</tr>
<tr>
<td>direct contrast</td>
<td></td>
<td></td>
</tr>
<tr>
<td>claim of uniqueness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>claim of omission</td>
<td></td>
<td></td>
</tr>
<tr>
<td>other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONTEXTUALIZATION</td>
<td></td>
<td></td>
</tr>
<tr>
<td>time or location awareness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cultural setting awareness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>biographic awareness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>historical awareness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>linguistic awareness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>analogy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>other</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TOTAL NUMBER OF HEURISTICS USED: ___________
Appendix E

PowerPoint Slides for the Academic Literacy Expectations Mini-Lecture
Slide 1

Academic Literacy Expectations in College
Responding to the Anticipation Guide

Slide 2

#1 Expert Readers Preview
- In 2 studies investigating how professors read, Wineburg (1991 & 1998) found that expert readers evaluated the source before reading the text.
  - Sourcing Heuristic (Strategy)
- These findings were supported by Wasson’s (2001) research with historians.

Slide 3

Evaluating Sources
In Wyatt et al.’s study (1993):
- All 15 expert readers evaluated the text
- All 15 expert readers evaluated the information vs.

Jolliffe (2007) indicates:
- In a survey, 35% of Freshmen reported evaluating sources
Slide 4

#2 Expert Readers Form Questions

- Expert readers express more doubt and formulate more questions than students (Wineburg, 1991).
- Students and university professors read 8 documents.
- Professors tended to use the Identify Difficulty Heuristic.

Slide 5

#3 Expertise is Domain Specific

A thorough review of 100 years worth of research into expertise by Ericsson, Krampe, & Tesch-Römer (1993) indicates that expertise is domain specific.

For example, a physicist may be an expert reader of physics articles (in his domain), but struggle with reading psychology articles.

Slide 6

#4 Expert Readers Embrace Confusion

Expert readers explicitly note difficulty (Wyatt et al., 1993)

- Wyatt et al. conducted research into the reading behavior of expert readers (15 social scientists) reading journal articles

- 13 expert readers noted difficulty
  * 3 experts noted difficulty 1 time
  * 3 experts noted difficulty 2-4 times
  * 7 experts noted difficulty 5+ times

Supported by MacDonald & Posthuma (1996) and Scardamalia & Bereiter, 1991.
In college, students are expected to
• read longer texts,
• from multiple sources,
• representing multiple perspectives
• on complex concepts

(Blåten & Strømsø, 2006)

• Williamson (2008) conducted research into the difficulty level of texts, using Lexile measurement.
• He analyzed a database of high school, college, workplace, citizenship, & military texts.
• Increase in difficulty of HS texts: 100 L
• Increase in difficulty between HS & College texts: 260 L

• College level reading assignments require the reader to flexibly adapt to the situation and use reading independently (Myers & Savage, 2005; Palmer 1999)
Literate Expertise

• In 1991, Scardamalia & Bereiter published a review of the literature on expertise in reading and writing.

• One of their findings was that expert readers put more into the reading process and get more out of the process.

Expert Readers Engage in Knowledge Transforming

• Knowledge Telling
  – Regurgitate information

• Knowledge Transformation
  – Dialectical process of reading
    • Incorporating information from the text into background knowledge
    &
    • Utilizing background knowledge to better comprehend the text.

A metaphor for reading at the university is...

• Criss crossing the knowledge landscape.
• For example:
  • Always Running is the knowledge area
  • Landmarks include gang affiliation & Viviana

• By approaching the landmark from different directions, we get different perspectives
Experienced Academic Readers...

- Experienced readers treat texts as "analyzable cognitive objects" (Flavell, 1981).
- In other words, experienced academic readers mentally pick up, examine, and evaluate books and articles from different points of view.

Experienced Academic Readers...

- Understand academic expectations
- Formulate questions within & across texts
  - Recognize inconsistencies
- Use evaluative heuristics

The Evaluative Heuristics Used by Expert Readers

1. Identifying Difficulty
2. Sourcing
3. Corroborating
4. Contextualizing
### Evaluative Heuristics Transcend Discipline

The use of evaluative heuristics has been documented among:

- Readers & writers (Scardamalia & Bereiter, 1991)
- Historians (Wineburg, 1998)
- Scientists (Jacobson, 2001)
- Doctors (Coulson, Feltovich, & Spiro, 1997)
- Immunohematologists (Smith et al., 1991)
Appendix F

Anticipation Guide for Academic Literacy Expectations: Answer Key
Anticipation Guide for Academic Literacy Expectations: Answer Key

Directions: Read through the six statements below. Check whether you agree or disagree with each and write 1-2 sentences explaining why.

1) Expert readers preview before reading.

Agree ___   Disagree___   Yes, Evaluate Source Information

2) Expert readers form fewer questions when reading than inexperienced readers.

Agree ___   Disagree___   No, More Questions, more tentative explanations

3) Once someone becomes an expert reader, he/she is an expert reader in any subject.

Agree ___   Disagree___   NO. Discipline & task specific skills/knowledge

4) Encountering confusion while reading indicates that you are an inexperienced reader.

Agree ___   Disagree___   No, good to be aware/metacognitive
                        No, may be textual features to engage w/
                        No, Ac. readers value questions ➔ productive

5) College students are expected to synthesize across multiple texts.

Agree ___   Disagree___   Yes. More reading, more texts, more connections

6) Reading expectations are the same for high school and college.

Agree ___   Disagree___   No, Type of reading
                        No, Difficulty of reading
                        No, Type/difficulty of task, i.e., Essay
                        No, Independence
Appendix G

Difficulty Paper Assignment and Feedback Sheet
Difficulty Paper Assignment Sheet

The purpose of this assignment is to help you pay greater attention to what your mind does as you read, and to allow you to explore texts in greater depth. In particular, I want you to learn about difficulty in reading—to recognize that in difficulty lies rich promise for interesting discoveries. All readers encounter difficulty, and we are exploring how to recognize it and use it to advance our understanding.

Part 1: Identifying Difficulty Across Multiple Texts  
DUE:

A) Read through the first text and notice any places that make you stop and think. Look for sections that particularly confuse or interest you. For example, are you confused about the author’s message or tone? Are you interested in the importance of a particular detail the author includes?

B) Now write a 1-2 page detailed description of your experience: what, specifically, did you focus on as you read? What, specifically, did you find interesting or confusing about these sections? What might you want to know more about? Try to be as specific as possible about which sections you are focusing on and what your mind was doing as you read these sections.

C) Repeat Steps A & B for the second text.

Part 2: Creating a Plan of Action  
DUE:

1. Looking back at what you wrote in Part 1, clarify one main question that you want to explore further using both texts.

2. Formulate a plan of action (2 pages) in which you devise some strategy you will use to answer your question. Perhaps you wish to use annotation to focus on one particular theme or idea as it runs throughout the texts, or you might decide to compare/contrast the authors’ experiences with your own. Be sure you explain not only what the strategy involves, but also what you hope to accomplish by using it.

Part 3: Implementing Your Plan  
DUE:

Here’s where you put your plan into action and answer your question. First, re-read the texts—all the way through—using whatever approach you have chosen. Then write 2 pages in which you reflect upon the texts with a fresh perspective. What new insights did you gain? You may wish to consider the following: In what ways did your understanding change or shift after re-reading the texts?

***Note: please include at least 4 quotes from the texts to support your ideas.

Part 4: Evaluating Your Plan  
DUE:

Write a page reflecting on your plan: How effective was your new approach? What might you do differently next time? How can you apply these strategies to future encounters with challenging texts?

Please turn in all 4 parts, including a Work Cited page to iLearn.
Difficulty Paper: Feedback Sheet

Reader/Writer: ________________________________

Part 1: Identifying Difficulty Across Multiple Texts

<table>
<thead>
<tr>
<th>Text 1 (title)</th>
<th>Text 2 (title)</th>
</tr>
</thead>
<tbody>
<tr>
<td>____ 1-2 pages</td>
<td>____ 1-2 pages</td>
</tr>
<tr>
<td>____ is specific about which sections you are focusing on</td>
<td>____ is specific about which sections you are focusing on</td>
</tr>
<tr>
<td>____ specifically explains about what your mind is doing as you read these sections; why you are interested and/or confused</td>
<td>____ specifically explains about what your mind is doing as you read these sections; why you are interested and/or confused</td>
</tr>
</tbody>
</table>

Part 2: Creating a Plan of Action

____ clarifies one main question about both texts to investigate

_____ the question focuses on one thing

_____ the question will allow for a 2 page response

_____ selects specific strategies

_____ details each step of the plan in well developed paragraphs

_____ explains what you hope to accomplish by using each strategy

Part 3: Implementing Your Plan

_____ reflects some new insight or shift in perspective

_____ includes at least 4 quotes from the texts

Part 4: Evaluating Your Plan

_____ reflects on effectiveness of chosen strategies

_____ reflects on future usefulness of those strategies as you continue to encounter difficult reading in college

_____ reflects on value or purpose of the assignment

Presentation

_____ includes all four parts and a work cited page, labeled and uploaded to iLearn
Appendix H

Introduction to the Evaluative Heuristics
Introducing the Evaluative Heuristics

Although you bring a lifetime of reading skills to college with you, you may have noticed that academic reading is different than other types of reading. Expert academic readers—scholars, researchers, graduate students—all use specific strategies and approaches to successfully navigate challenging texts. You have probably used these same strategies or heuristics. The difference is the frequency with which these evaluative heuristics are mobilized. Successful academic readers use these evaluative heuristics regularly when reading. They may use several heuristics concurrently, but they use only the most helpful features of each heuristic in each reading situation.

The purpose of this handout is to introduce those heuristics to you. The purpose of the instruction over the next few class sessions is to help you learn when and how to use each heuristic, so you can use those same expert reading strategies to hone your critical thinking and writing.

We will be working with four heuristics. All four have been identified by research as important elements to readers’ success in academic reading situations.

**Heuristic 1: Identifying Difficulty**

While most people look to avoid difficulty, academic readers rejoice in uncovering moments of difficulty. Researchers look for gaps in the research literature, scientists look for holes in their theory, philosophers look for breakdowns in arguments. Research suggests that undergraduates tend to gloss over difficulty, while experienced academic readers focus in on it. When students run into questions, they are sometimes frustrated. Undergraduates might think they don’t know enough to sort through challenging information. They might feel that they are stuck whereas these moments of difficulty open up a variety of possibilities for experienced academic readers. When they notice a contradiction between texts or find they have a question, they mobilize a set of strategies. They define the problem or the question, they form a plan for answering the question, and they decide what resources they might be able to use to resolve the difficulty.

The first strategy experienced academic readers employ is actively identifying difficulty. Then, they come up with a plan of action in order to resolve that difficulty, which often includes using the other three evaluative heuristics. To help us learn this process we’ll use the Difficulty Paper assignment whose parts mirror the problem-solving steps that expert readers utilize. As we discuss the parts of the Difficulty Paper assignment, you may recognize steps you’ve taken to solve past problems.

**Steps in the Difficulty Paper**

**Part 1: Identifying the Difficulty** - Being aware of your reading process

**Part 2: Creating a Plan of Action** - Defining the question and deciding on the steps to resolve the difficulty

**Part 3: Implementing Your Plan** - Completing the steps in order to answer your question

**Part 4: Evaluating Your Plan** - Refining the strategy for next time
Heuristic 2: Sourcing

When experienced academic readers preview a text, they tend to note the source information. Although you may have used source information to help you evaluate information, you probably did not start by assessing the source characteristics. But experienced academic readers often evaluate the source before they start reading the content of the text.

The chart below includes some of the features they evaluate. Please fill in the last column, explaining why a reader should evaluate each source feature.

<table>
<thead>
<tr>
<th>Source Feature</th>
<th>Definition or Criteria</th>
<th>Reason to Evaluate It</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author Credentials</td>
<td>Information about the author’s profession, level of training, or other credentials</td>
<td></td>
</tr>
<tr>
<td>Author Motivation</td>
<td>Information about why an author might have written something</td>
<td></td>
</tr>
<tr>
<td>Author Participation</td>
<td>Information about the author’s level of participation in the events he discusses</td>
<td></td>
</tr>
<tr>
<td>Date of Production</td>
<td>Information about when the document was created or published</td>
<td></td>
</tr>
<tr>
<td>Document Type</td>
<td>Information about the type of document</td>
<td></td>
</tr>
<tr>
<td>Any other feature you</td>
<td></td>
<td></td>
</tr>
<tr>
<td>can think of?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any other feature you</td>
<td></td>
<td></td>
</tr>
<tr>
<td>can think of?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

You can use the specific features of a text to argue for or against its credibility. For example, *the author had a lot to gain by telling his boss about his success, so he may have exaggerated which means this document isn’t very trustworthy*. Or, *people usually write in their journals to keep a record for themselves, so it wouldn’t make sense for him to write lies in his journal*. 
Heuristic 3: Corroboration

You may have heard the term “corroborate” if you watch any of the legal procedurals on television, like Law & Order or CSI. The detectives corroborate an alibi by checking with someone who can verify that it is true or a second witness must be found to corroborate the testimony of the first witness. In academic situations, corroboration refers to evaluating how much support there is for a claim across multiple texts or people.

College readers use other documents, texts they have previously read, prior learning, and knowledge of the discipline to compare and contrast information. Academic reading and writing demand verifiable, supportive evidence in order make sure that information from a text is trustworthy.

On the chart below, fill in the final column with examples of corroboration from the articles you are using for your Inquiry Project. You will need to use at least 2 articles, but you could use all four to complete this task.

<table>
<thead>
<tr>
<th>Type of Corroboration</th>
<th>Explanation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Comparison</td>
<td>The reader makes a direct connection between similar information that was found in two or more documents.</td>
<td></td>
</tr>
<tr>
<td>Direct Contrast</td>
<td>The reader points out information that was different in two documents.</td>
<td></td>
</tr>
<tr>
<td>Claim of Uniqueness</td>
<td>The reader shows that information was found in only one source.</td>
<td></td>
</tr>
<tr>
<td>Claim of Omission</td>
<td>The reader shows that a source left out important information that was found in other sources.</td>
<td></td>
</tr>
</tbody>
</table>
Heuristic 4: Contextualization

Readers and writers use contextualization to place themselves or their reader in the specific context of the event that is taking place. They discuss specific details about the event that helps them understand why or how the event took place.

Because of the depth of analysis required for many academic tasks, academic readers and writers rely on contextualization to gain new insights.

<table>
<thead>
<tr>
<th>Type of Contextualization</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time or location awareness</td>
<td>The reader demonstrates awareness of the chronology or specific features of the physical location of an event.</td>
<td></td>
</tr>
<tr>
<td>Cultural setting awareness</td>
<td>The reader demonstrates awareness of the cultural values or common attitudes of the time period, or emotions that participants in an event may have been feeling.</td>
<td></td>
</tr>
<tr>
<td>Biographic awareness</td>
<td>The reader demonstrates awareness of the values and influences surrounding the production of the text, including biographical information about the author.</td>
<td></td>
</tr>
<tr>
<td>Historical awareness</td>
<td>The reader demonstrates an awareness of how surrounding historical events may contribute to an understanding of the text.</td>
<td></td>
</tr>
<tr>
<td>Linguistic awareness</td>
<td>The reader demonstrates awareness of the meanings of words.</td>
<td></td>
</tr>
<tr>
<td>Analogy</td>
<td>The reader creates an analogy, comparing things, in order to more fully understand an idea.</td>
<td></td>
</tr>
</tbody>
</table>
Contextualization Practice: Author’s Message

What do you think the author’s purpose or message is?

Below are 6 ways experienced academic readers attempt to contextualize information in a text. They look for information about each type of context. Then, they consider how each piece of information helps them understand the writer’s message.

Please find one example for each type of context and connect it back to what you have identified as the author’s message.

<table>
<thead>
<tr>
<th>Type</th>
<th>Example (Cite page #)</th>
<th>Connection to Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time or location awareness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cultural setting awareness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biographic awareness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical awareness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linguistic awareness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analogy</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix I

Informed Consent for Research Participation
Informed Consent for Research Participation

Purpose and Background
Patty Baldwin, a graduate student in the School of Education at University of San Francisco and a Lecturer as San Francisco State University, is conducting a study on the effectiveness of strategies in helping students resolve controversies across multiple texts. The purpose of this study is to test the effectiveness of an instructional program to help with evaluating information. I am being asked to participate because I am a student enrolled in an Integrated Reading and Writing (IRW) course.

Procedures
If I agree to be a participant in this study, the following will happen

1) I will complete a short questionnaire giving basic information about me, including age, gender, ethnicity, and language background.

2) I will complete pretest related to reading and writing;

3) I will participate in an instructional unit on evaluating information, including class participation, in-class activities, homework, and two writing assignments; and

4) I will complete a posttest related to reading and writing.

Risks/Benefits

1) I am free to decline to participate in this study or to discontinue participation in the research study at any time. My decision as to whether or not to participate in this study will have no influence on my course grade or status as a student in this class. If I decide not to participate in the research, my data will not be used in data collection for the study. However, I will still be expected to complete all the same work, which is part of the curriculum of the class I am enrolled in.

2) Participation in research may mean loss of confidentiality. Study records will be kept as confidential as possible. All participants’ data will be coded to protect their identity. No individual identities will be used in any reports or publications resulting from this study. Study information will be coded and kept in locked files at all times. Only the researcher will have access to the files.

3) There will be no direct additional costs to participating in this study. There will be no direct benefit to me from participating in this study. Whether or not I agree to participate in this study, I will complete all the same work which is part of the curriculum for the course in which I am enrolled.

4) The anticipated benefit of this study is to improve instruction in IRW courses.
Questions

If I have any questions or comments about participation in this study, I should first talk with my instructor. I may also contact the researcher. I may contact Patty Baldwin at 415-XXX-XXXX or pattiey@sfsu.edu.

If for some reason I do not want to do this, I may contact the IRBPHS, which is concerned with the protection of volunteers in research projects. I may reach the IRBPHS office by calling (415) XXX-XXXX and leaving a voicemail message, by e-mailing IRBPHS@usfca.edu, or by writing to the IRBPHS, Department of Counseling Psychology, Education Building, University of San Francisco, 2130 Fulton Street, San Francisco, CA 94117-1080.

Consent

I have access to the “Research Subject’s Bill of Rights” and I have been given a copy of this form to keep.

PARTICIPATION IN THIS RESEARCH IS VOLUNTARY. I am free to decline to participate in this study or to withdraw from it at any point. My decision as to whether or not to participate in this study will have no influence on my course grade or status as a student in this class.

My signature below indicates that I agree to participate in this study.

___________________________________
Participant’s Name

___________________________________
Participant’s Signature

___________________________________
Researcher’s Signature

___________________________________
Date

___________________________________
Date
Appendix J

Demographic Questionnaire
Demographic Questionnaire

1) How old are you?
   17 years   18 years   19 years   20 years   older than 20 years

2) What is your gender?
   Female     Male     Other

3) Please put a checkmark beside the ethnicity that best describes you.
   ____ African American
   ____ American Indian or Native Alaskan
   ____ Asian American
   ____ Filipino
   ____ Mexican America/Mexican
   ____ Other Latino
   ____ Pacific Islander
   ____ White/Caucasian
   ____ Other

4) What is your primary language for speaking? _____________________________

5) What is your primary language for reading and writing? ___________________

6) Do you speak any other languages proficiently? ___________________________

   Please list those languages.
   ______________________________________________________________

7) Do you read and write proficiently in any of those other languages? _________

   Please list which languages. _________________________________________

8) How long have you lived in the United States?
   2 years or less   3-5 years   6-10 years   11 or more years
Appendix K

Letter of Permission from Instructor to Participate in Study
November 2012

Dear Ms. XXXX,

My name is Patricia Baldwin and I am currently a doctoral candidate in the School of Education at the University of San Francisco. As part of the degree requirements, I am doing a study on the effects of an academic literacy instructional intervention on the ability of developmental-level freshmen to recognize and resolve inconsistencies across multiple texts.

This project includes a pretest, four weeks of in-class instruction, and a posttest, totaling 10 class periods. The instructional intervention will include reading several education-themed texts, completing an anticipation guide, listening to a PowerPoint lecture about post-secondary academic literacy expectations, engaging in discussion, participating in structured in-class activities relating to the skills experienced academic readers and writers demonstrate in recognizing and resolving inconsistencies across multiple texts, and completing two difficulty paper assignments. This study will collect background data through a demographic questionnaire and a topic familiarity measure. Data on the Multiple Text Tasks will also be collected at pretest and posttest. It is my hope that the explicit instruction in recognizing and resolving inconsistencies will help developmental-level freshmen engage in the advanced academic literacy skills that experienced readers and writers use to be successful at the university.

Participation in this study in entirely voluntary and the results will be kept confidential and anonymous. Your signature on the enclosed permission letter indicates that you acknowledge and authorize the research that is to be conducted with the permission of the English Department chair and the consent of students in your two sections of integrated reading and writing. Please keep a copy of this letter and the permission form, sign one copy of the permission letter, and return it in the pre-addressed, stamped envelope.

Sincerely,

Patricia Baldwin
Doctoral Candidate, University of San Francisco
pattiey@sfsu.edu or (415) XXX-XXXX
Permission

My signature below indicates that I authorize Patricia Baldwin to conduct a research study on the effects of an academic literacy instructional intervention in the two sections of integrated reading and writing that I teach at San Francisco State University. I give permission for her to contact students and gather data. I am agreeing to administer the pretest and posttest and teach the four-week instructional intervention in my classes.

Signature:

Date:
Appendix L

Education Readings for the Instructional Intervention
Education Reading Citations for the Instructional Intervention


Appendix M

Model Difficulty Paper
Model Difficulty Paper

Part 1: Identifying Difficulty in “Education in a Multicultural Society”

In “Education in a Multicultural Society: Our Future’s Greatest Challenge,” Lisa Delpit outlines the problems associated with culturally tone-deaf education and recommends that education be recast as culturally sensitive. I think this is an important area for improving teacher practice and challenging to many educators, myself included, because we are limited by our own cultural perspective. I appreciate that she discusses several ethnic groups and moves the discussion from individual teacher responsibility towards systemic changes that would need to be envisioned. As I read this chapter, a number of points captured my interest and raised questions from me.

Early in the article, Delpit notes that teachers can “easily misread students’…abilities” (167). This seems like a human moment to me. Cultural experience, along with SES, gender, race, age, and sexual orientation, can direct individuals to incorrect assumptions. Interestingly, she notes on the second page how students can misread their instructors: “the second statement sounds to many of these youngsters like the words of someone who is fearful (and thus less deserving of respect)...” (p. 168). I realize that the adults—teachers—should shoulder more responsibility for their preconceptions. However, I wonder if Delpit is aware that she suggests we should examine teachers’ cultural assumptions, while accepting students’ assumptions at face value.

Delpit uses Native American’s cultural prohibition against speaking for someone else as an example of culturally incompatible instruction (p. 170). I appreciate the specific example which helps make the argument more concrete. Plus, I didn’t know that
about Native American communities. Is it widespread throughout most Native American communities? Or, is this true of specific communities/tribes? I see the potential for conflict between cultural values and educational expectations, and empathize with the individuals who are caught in that tension. However, Delpit doesn’t offer a solution. Many would argue that students should be taught to summarize as it is a useful skill in school and workplace settings. How could you help Native American students resolve the tension and acquire a skill that schools should teach, like summary writing?

I think Delpit sort of addresses this issue later in the article when she points out that teachers might not call on a Native American student in order to “avoid causing them discomfort” (p. 172). Here a teacher is aware of a cultural sensitivity on the part of the student and tries not to put the student in an uncomfortable position. However, the teacher still disprivileged the student by not allowing them to fully participate in the classroom community. What does Delpit suggest that the teacher do? Awareness isn’t enough, so how should the teacher learn culturally inclusive practices? Why aren’t we teaching them in teacher education programs? Do we even know what culturally inclusive practices look like?

I noticed a place that reminded me of Lockhart’s Mathematician’s Lament. Delpit provides an example of teaching dance through a series of workbooks (p. 173) that echoes Lockhart’s claim that schooling can squeeze the joy, life, and creativity out of any enterprise. I did wonder at Delpit’s choice of art. Did she choose dance to tap into a preconception of African-American’s superior skill in that area? Is she falling into cultural insensitivity by trading on a stereotype? Or is she presenting a vivid example
that many in her audience can readily identify with? What if she’s doing both: reinforcing a stereotype, while simultaneously increasing our awareness of it?

I whole-heartedly agreed with the point that “child-centered” approaches can yield less instruction (p. 174). Discussion can be a way to bring in other perspectives, try out ideas, and refine our thinking. But, it is not an adequate solution for reading. Educators have to walk a fine line. We are asked to note a child’s strengths, not dwell on their weaknesses, but offer instruction to address those weaknesses and strengths. Culturally sensitive practice is not necessarily instruction. We can make people feel good, or seen, without necessarily providing them the education they need. How do we balance making students feel good about their skills with providing needed instruction to build those skills?

**Part 1: Identifying Difficulty for “Beyond Language”**

In “Beyond Language: Ebonics, Proper English, and Identity in a Black-American Speech Community,” John Ogbu discuss conflicting beliefs about acquiring standard English to which one African-American community subscribes. Ogbu conducted a multi-year ethnographic study in the African-American community of Lafayette. This article presents data gleaned from numerous interviews of adults and children in the community. This article brought up a lot of questions for me, including some that arise in relation to Delpit’s work.

I noticed that the abstract mentions 1996, which is a decade prior to the publication of Delpit’s book. Are there differences between the author’s perspective that are merely a reflection of a decade of cultural and educational change? Are both arguments dated? Ogbu’s research took place between 1988 and 1990, so it is over
twenty years old. There’s been a whole new generation in that community since this study was conducted. It takes a long time to get something published, so Delpit’s work could be more than six years old. John Ogbu identifies himself as African American. Is Lisa Delpit also African American? I thought she spoke with an authority and fearlessness that suggests she is an African American.

I had more than two pages worth of questions and points of interest with Delpit, yet I have even more questions about this article. I’ll have to be selective to keep it around three pages which is already a little bit long.

Ogbu suggests that we can’t just look at the students. We need to see them within their community cultural context to really see them (p. 149). I agree that we are the products of our environments, as does Delpit. However, instructors do not have the luxury of multi-year research projects to learn about each community? And many instructors teach in diverse classrooms with students from multiple cultural and language communities. How can a teacher build adequate knowledge within the time and resource limitations he/she must necessarily work?

It does seem like international students and immigrant students are willing and eager to learn the “different cultural rules” (p. 149). Why don’t they feel that they will lose their original language/culture by acquiring fluency in English/American culture?

“Diglossia” (p. 150) was a new term for me. Why didn’t he use bi-dialectalism or another term that seems closely related to the point he is trying to make? I found diglossia confusing to keep track of because it was a new vocabulary word and because it refers to bi-lingualism, not just the competing dialects Ogbu wishes to focus on.
Several interviewees echo Parent 28L: “‘White people are born to talk proper English,’ whereas ‘Black people have to learn it’” (p. 162). Does this mean that the community believes they will never be native speakers of standard English? Will they always feel that they have an “accent”? Do they really mean it is inborn/innate or just easier to learn?

I loved the example of the white double standard. Black people speak slang, but white people get new names for their “slang,” like “valley talk” (p. 163). So, white dialects are not dismissed as slang, but named and categorized as alternative “real” languages.

One child notes that his mother code switches in certain situations. In addition to switching to standard English, she speaks in a “little, high-pitched voice” (p. 165). Does her voice really change, or does it just sound different because the words are different? If it is different, could it be because of gender expectations (i.e., women are less threatening)? Or, is it just her phone voice? Or, is she literally representing herself as someone else?

Several interviewees suggest that teachers should be “firm in correcting children’s speaking and writing” (p. 169). Is that culturally sensitive teaching?

It is crazy to me that the community ostracizes those who use standard English (p. 170). Especially, when they agree that students should learn standard English? Why must speakers of standard English be seen as traitors?

If slang were the yard stick to measure success in school (p. 179), instead of standard English, would middle- and upper-class white families hire slang tutors?
Part 2: Plan of Action for Ogbu & Delpit Texts

I found a lot of points of intersection between the two readings. Certainly, cultural sensitivity is a main idea in each article. However, I’m interested in a slightly different question. I noticed the dates of both readings. Although I consider Delpit’s work current, Ogbu’s study could be dated. I have questions about how representative their claims are for 2012. My overarching question is: Do their articles represent the experience of current African American communities in education?

In order to answer my questions, I plan to use several strategies. First, I want to research Ogbu and Delpit. Learning about their age and background will help me decide how they fit into the communities at the times they are writing. For instance, if Ogbu was already in his 40’s, he would have already had an entrenched world view. If he was younger, it is possible that his preconceptions influenced his findings about the perceptions of the African-American community to a lesser degree.

Secondly, I should look at the rest of Delpit’s book to gain insight into her perspectives. I’d also like to know if her work is based on research, or just her personal experience, or just her personal insight. I might also get a sense of when she wrote the book: in the early 2000’s or over a longer stretch of time representing a different time in our cultural history.

Thirdly, I’d like to use JSTOR (or an equivalent library database) to find current research on the African American community perceptions of standard English and schooling. I’ll start by looking for additional work by these two authors. I can also use the references from Ogbu (and maybe Delpit’s book) to see if any of those researchers have done more current work in the field. Lastly, I could just do a general search of
“perceptions of schooling in the African-American community” and “perceptions of standard English usage in the African-American community.” Using the advanced feature, I can set the dates of publication to only return items published in the last two years.

Because I am not part of the African-American community, my personal experience won’t help me much with this question. I’d love to do some sort of survey research, but there’s not a lot of time to complete this assignment and I need to spend a good chunk of the available time on library research. But, I could possibly set up an interview. I know a lot of teachers. I’ll start by trying to brainstorm a list of teachers who might have insight in to this phenomenon. Then, I’ll try to contact them via email, letting them know why I’m interested in talking with them and sharing a few of my questions. If I can find a person to interview and set up an interview before the assignment is due, I’ll add an interview to my sources.

The place my personal experience might help is with current educational contexts. Since I’m a teacher, I can use my prior knowledge of how school works. As I read through the articles again, I can take notes on specific examples and illustrations with regards to how current they seem. Is that something that would still happen?

With my notes from external research and revisiting the currency of illustrations within the texts, I should be able to decide if the issues and perspectives that Ogbu and Delpit bring up are still applicable to our current educational context.

Part 3: Implementing My Plan

In working through the steps of my plan, I discovered several interesting insights. I learned about both authors, including that both are well-recognized in the field of
education and both have published more recent works dealing with the same themes. Although the research may be dated, their theories are still alive in academic debate. Not only are both Ogbu and Delpit’s theories still relevant, the unexpected connections between them are more relevant than the contrasts.

While researching Ogbu and his theory, I found that his work is still relevant today. Brandes, Dundes, and Nader (2003) memorialized Ogbu for UC Berkeley’s Academic Senate. Goldsmith (2003) wrote an in-depth exposé on Ogbu’s current research for Oakland’s weekly East Bay Express and Foster (2004) published a critique of Ogbu’s theoretical work in a peer-reviewed, scholarly journal. Ogbu was born in 1939 in Nigeria (Brandes, Dundes, & Nader). This originally suggested that he might have a mindset different than that of 2012. However, his experiences as an immigrant and his position outside of America could offer a unique perspective. He came to UC Berkeley in 1961, conducted fieldwork in Stockton, California (Brandes, Dundes, & Nader). So, much of his experience as an educator has taken place in California and he has studied different communities. These experiences all add to his credibility. He has been recognized by UC Berkeley for contributions to several fields, particularly anthropology and education. Brandes, Dundes, and Nader noted that he applied anthropological theories to minority education in the US providing us with additional tools for research. His theory of cultural ecology, delineating the systemic forces and the community forces that operate on individuals, continues to be discussed in peer-reviewed journals (Foster, 2004) and his theory about voluntary vs. involuntary minorities is part of the current discourse on education and social justice (Foster).
In researching Delpit and her theory, I came to better understand her philosophy. T.M.B. (1995) reviewed her first book for the prestigious *Harvard Educational Review*. Goldstein (2012) interviewed Dr. Delpit for *The Nation*, a respected journal concerned with political and cultural analysis. Delpit is critical of deficit models of education (Goldstein, 2012), where students are seen as lacking and need to be fixed. She does believe that we need to explicitly teach standard English and how to be successful to all children (Goldstein) as do the parents in Ogbu’s study. She argues that “We see through our beliefs,” meaning teachers are blinded to students because of the cultural preconceptions they bring into the class with them (TMB). Although I agree that we are influenced by our experiences and our beliefs, I would suggest that the “we” refers to all people—parents, students, community members—as well as to teachers. She blames the systemic factors, singling out teachers, rather than examining the community factors.

Looking at Delpit’s book suggests that she doesn’t have strong research to support her claims. Her first book was published in 1995, seventeen years ago. Much of her evidence is in the form of excerpts from conversations with parents, teachers, and students (T.M.B., 1995), which a less rigorous version of Ogbu’s research methodology. She includes what she learned on her “personal journey” as a mother and an educator (Delpit, p. 73). Her experiences with her daughter Maya who has ADHD have also informed her theory (Goldstein, 2012). She draws on her personal experiences in Papua New Guinea and Alaska (T.M.B.). Although these experiences offer an interesting perspective, her research was not as rigorous as Ogbu’s. She studied very different populations, including those outside the US, whereas Ogbu has conducted research in
different minority communities within the US and “how they compare to racial and
ethic minorities in India, Israel, Japan, New Zealand, and Britain” (Goldsmith, 2003).

Although the original texts are dated, the thinking is still current today. The
groundbreaking essay that formed the basis of Other People’s Children was published in
1988 (Goldstein, 2012). The book was published in 1995. However, she has a new book
out just this year: “Multiplication is for White People” (2012). She is an Eminent Scholar
and Executive Director of the Center for Urban Educational Excellence at Florida
International University in Miami (Goldstein). She currently holds a prestigious position
in a nationally recognized organization. Delpit and her views on culturally sensitive
teaching remain very much a part of the discussion of education in America.

Despite the dated research in “Beyond Language,” Ogbu’s work also continues to
be part of the current discussion of education. In fact, in the late 1990’s Ogbu was
approached by African American parents to examine academic disparities in the upper-
middle class Shaker Heights neighborhood of Cleveland, Ohio (Foster, 2004). Clearly,
his reputation has spread well beyond the academy if parents in Ohio wanted to recruit
him. He and his research assistant “observed 110 different classes, from kindergarten all
the way through high school” and spent one year interviewing parents, teachers, students,
and community members (Goldsmith, 2003). He published his finding in Black Students
in an Affluent Suburb: A Study of Academic Disengagement (2003). In line with
“Beyond Language,” he concluded that students knew they needed to work hard for
academic success, but didn’t do it. This more recent research with a different African-
American community at a higher socio-economic status, adds to his previous findings.

Plus, a number of scholars, including one of Dr. Delpit’s colleagues at Georgia State
University are “taking a few more swipes at [his] premise” (Goldsmith, 2003), suggesting that the other scholars in the field were following his work. Even those that take issue with his conclusions, support his theoretical work. In fact, Foster straight out states that “the problems were not in his theory” (2004, p. 377). And his methodology is still setting the bar for ethnographic studies of minority education communities. Clearly, his work remains relevant.

The surprising finding for me was that these two theorists are engaged in an academic debate, rather than a practical debate. By that, I mean they seem to be clashing as opposing forces, but further investigation shows they are not far apart ideologically. In fact, Ogbu’s theory of the community and system factors explains the difference between them. Ogbu posits that both community and systemic factors influence student’s performance. His research has focused on explaining the community factors. Delpit eschews the role of the community focusing exclusively on the systemic factors, especially teachers. However, if you put their work together, you get a fuller view of the systemic AND community factors. In addition, they both agree that teaching standard English is crucial and the instruction in how to be successful will be the key to helping all students be successful. Overall, their adversarial posturing masks the basic compatibility of their views: the need to bring school and community forces together to educate all our children.

Part 4: Evaluating My Plan

Overall, my plan helped me arrive at some interesting insights about the two authors, including the similarities in their seemingly different stances. I think my plan
was successful although I did not complete all the steps. In the future, I’d add a time management component.

I conducted research on Ogbu and his theory. The currency of articles and blog posts suggested that his work is still part of the discussion about African American education. In particular, his theory was recently discussed as a model for future theories by Foster (2004). His theory about voluntary vs. involuntary immigrants is very much alive in the discussion of educational attainment. Finding out a little about his life and seeing how his work is still part of the debate, helped me see his relevance.

Similarly, looking at Delpit’s book showed me how little direct research she has. Her work is based mostly on personal experience and anecdote. These personal experiences are valuable contributions to our understanding of teaching and culturally responsive pedagogy, but further research is needed. However, publishing a new book shows that her perspective is still valued within the community of scholars.

In both cases, finding out a little about the background of each writer helped me to contextualize their ideas. Looking at the type of work they’ve produced, research vs. theoretical musing, helped me to understand their perspectives. Reviewing their more recent work helped me to answer my question by showing how relevant their work still is.

I got the information I needed from the first three steps, barely using JSTOR, since the book publication dates of their more recent work showed current contributions by each author. Therefore, I didn’t feel the need to conduct an interview. Plus, I used a lot of time researching and reading the sources I did use in my Difficulty Paper. I wouldn’t have had time to set up and conduct an interview.
Although I didn’t enact my prior knowledge as a teacher, I did make use of my experience with academic debate. I drew on my understanding of the rules of engagement and the need to make a name for yourself through publication of your work to understand the relevance of both authors and to determine that there are underlying similarities, which aren’t being explored, because of the pressure to produce “new” or different work. Controversy raises one’s profile and, therefore, one’s reputation.

My plan was very successful. I answered my question and gained new insight. However, in the future I would also include time allotments for each step. Once, I get into researching, I find that I can spend a lot of time trolling the internet. I might have gotten more out of the JSTOR databases, if I’d looked at them earlier. Although I didn’t need the interview, I may want to do something similar in the future and will need to apportion my time better to ensure that I can complete any steps I need to get the answer to my question. It worked out here, but if I need to go further into my plan in the future, I’ll need to manage my time better.

Overall, I answered my question and learned the importance of contextualizing the thinking of the authors I read. I also noticed that a little investigation can help bring the whole big picture into focus. I might research the authors of challenging texts, even when I’m not asked to do a difficulty paper because it certainly helps to see their relevance to today’s world.
Works Cited


Appendix N

Question Type Handout
### Question Types for the Difficulty Paper

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Explanation</th>
<th>Example (From “The Three Little Pigs”)</th>
<th>Your Turn:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Right There</strong></td>
<td>A question whose answer is “right there” in the texts. All you have to do is find it and copy it down.</td>
<td>What kind of building material did the first pig use?</td>
<td></td>
</tr>
<tr>
<td><strong>Pulling It Together</strong></td>
<td>A question whose answer is in the texts, but a reader has to “pull it together” from different parts.</td>
<td>Which building material was the most stable? What is the tone of the parable?</td>
<td></td>
</tr>
<tr>
<td><strong>Text &amp; Me</strong></td>
<td>A question whose answer is not solely in the texts. The reader has to combine information from the texts with information from his/her own prior knowledge to come up with a reasonable answer. This kind of thinking is called inferential.</td>
<td>Which pig would have survived the mortgage crisis?</td>
<td></td>
</tr>
<tr>
<td><strong>Research</strong></td>
<td>A question whose answer is not in the texts. The reader would need to do some research in addition to using the texts to answer the question.</td>
<td>How do wolves find their prey?</td>
<td></td>
</tr>
<tr>
<td><strong>On My Own</strong></td>
<td>A question whose answer is not in the texts. The texts simply serve as a springboard for the reader’s discussion of a topic related to the readings.</td>
<td>What kind of house would I build?</td>
<td></td>
</tr>
</tbody>
</table>
Appendix O

Contextualization Application Activity Handout
**Contextualization Application Activity: “We Don’t Need No Education”**

Below you’ll find lyrics for “We Don’t Need No Education” from Pink Floyd’s 1979 rock opera, *Another Brick in the Wall, Part II*. Read through and annotate the lyrics. Then, in a small group, work to contextualize the song. What do you know about the band, the song, the late 1970’s? Think about how what you know might have impacted the song. After you’ve met with your small group, we’ll discuss your findings as a full class.

The lyrics can be found here:

Appendix P

Conditional Knowledge of Sourcing Activity
Conditional Knowledge of Sourcing: What to Focus On

Directions: Read each situation carefully and decide which source feature is the most important for making a decision about the credibility of the text. In the space to the left, please note why you selected that source feature.

1. You are reading a review of the new version of the iPad in order to decide whether or not to purchase it.
   a) date
   b) author
   c) sponsoring organization

2. You are reading an article on the economy in order to understand the growing mortgage crisis.
   a) publisher
   b) date
   c) title

3. You are reading a website giving diet advice because you are concerned about your friend’s new diet plan.
   a) date
   b) author
   c) sponsoring organization

4. You are reading a journal article on advances in genetics for your Ethics class.
   a) title of journal
   b) title of article
   c) author

5. You are reading an article about Taylor Swift's new love interest in order to join a conversation on Facebook.
   a) title of magazine
   b) title of article
   c) author