

Exploring Patterns of Historical Thinking through Eighth-Grade Students' Argumentative Writing

Jeffery D. Nokes

Brigham Young University, UT | USA

Abstract: Building upon research on writing and on students' development of historical thinking, the current study uses patterns in 427 eighth graders' writing about sources to assess the sophistication of their historical reasoning. A spectrum is proposed to represent five levels of increasingly sophisticated writing about the source of documents. Using examples of students' writing, this paper shows how students at lower levels fail to recognize sources of documents or misuse source information. In contrast, students at higher levels of the spectrum critically analyze sources of documents and a few even use source information persuasively in their writing. The spectrum, which categorizes the quality of students' sourcing, correlates positively with the frequency of sourcing in students' writing, suggesting a connection between strategy use and historical thinking. Suggestions for designing written assessments of students' historical thinking are proposed.

Keywords: Historical writing, historical literacy, historical reading, history teaching, assessments of historical thinking



Nokes, J. D. (2017). Exploring patterns of historical thinking through eighth-grade students' argumentative writing. *Journal of Writing Research*, 8(3), 437-467. doi: 10.17239/jowr-2017.08.03.02

Contact: Jeffery D. Nokes, History Department, 2160 JFSB, Brigham Young University, Provo, UT, 84602 | USA – jeff_nokes@byu.edu

Copyright: Earli | This article is published under Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 Unported license.

This study identifies patterns in students' writing about the sources of documents—writing that indicates various levels of sophistication in their historical thinking. In recent decades international educators and researchers have paid substantial attention to teaching historical reading, thinking, and writing (Eliasson, Alvén, Yngvéus, & Rosenlund, 2015; Körber, 2011; Lee & Ashby, 2000; Martinez, Mateos, Martín, & Rijlaarsdam, 2015; Seixas & Morton, 2013; Van Drie & Van Boxtel, 2008). The C3 Framework (National Council for the Social Studies, 2013), the influential Common Core State Standards (Common Core State Standards, 2010), and a growing body of research and emerging pedagogical practices (Lesh, 2011; Nokes, Dole, & Hacker, 2007; VanSledright, 2002; Wineburg, Martin, & Monte-Sano, 2013) have moved the United States into closer cadence with the world in terms of balancing content instruction with the nurturing of historical thinking and historical consciousness.

With the current interest in the design of reliable and valid assessments of historical thinking (Ercikan & Seixas, 2015; Smith & Breakstone, 2012; VanSledright, 2014), there is a growing worldwide movement to teach and assess historical thinking and historical consciousness through students' writing (Eliasson, et al., 2015; Monte-Sano, De La Paz, & Felton, 2014; Seixas, Gibson, & Ercikan, 2015; Waldis, Hodel, Thünemann, Zülsdorf-Kersting, & Ziegler, 2015). Global trends create the need to use assessments to identify the capabilities, deficiencies, and development of diverse aspects of students' historical thinking and historical consciousness so that educators can help students learn these vital skills and ways of viewing the world. This paper focuses on the use of a written assessment to diagnose students' ability to engage in sourcing, a basic historical thinking strategy that involves using information about an author, audience, and purpose to understand and use historical evidence in argumentative historical writing.

Although sourcing is a historical thinking strategy that students use with relative sophistication after some instruction (Nokes et al., 2007; Reisman, 2012), questions remain about how to assess sourcing as applied in students' argumentative writing. The Stanford History Education Group (2012) has prepared excellent resources for assessing students' use of sourcing, corroboration, and contextualization. However, these instruments provide less evidence of how students apply historical thinking when writing longer argumentative essays. In other words, it is fairly easy to design assessments to see whether students use sourcing but more difficult to design assessments to see how students use sourcing as part of larger historical argument. Questions remain about how strategy use (i.e. sourcing) translates into historical thinking and writing. For instance, what does it indicate about a student's historical thinking when he/she mentions the source of a document but does not write about how the source influences the document's content? How are educators to evaluate students' writing when students acknowledge but misuse source information in making an argument? Does such writing indicate higher or lower levels of historical thinking than not writing about a document's source at all? Such questions surround the cognitive moves required to synthesize an analysis of evidence into persuasive argumentative writing. These questions involve the translation of strategy-driven analytical reading

into writing that uses the analysis as leverage in making a case for an interpretation. This paper shares research using the writing of eighth-grade students, identifying patterns in their writing about sources of documents that indicate more and less sophisticated ways of thinking about the source of primary, secondary, and fictional accounts.

1 Theoretical Framework

This study is based on research on a) disciplinary writing; b) historical reading, thinking, and writing; c) the development of historical thinking; and d) the connection between writing and thinking.

1.1 Disciplinary Writing

Increasingly, researchers have argued that what counts as text, how individuals learn with texts, how texts are structured, and how texts are used varies strikingly across disciplines (Alexander & Jetton, 2000; Körber, 2011; Moje, Stockdill, Kim, & Kim, 2010; Shanahan & Shanahan, 2008). They argue that scientists, mathematicians, artists, historians, and scholars in each discipline use unique strategies to comprehend and use texts (Draper, Broomhead, Jensen, Nokes, & Siebert, 2010; Moje, 2008). Although researchers have paid greater attention to historical reading than historical writing, the consideration of specialized disciplinary literacies has recently begun to focus on the creation of disciplinary texts—“writing” within the disciplines (De La Paz, Ferretti, Wissinger, Yee, & MacArthur, 2012; Duke, Caughlin, Juzwik, & Martin, 2012). These researchers contend that like disciplinary reading the processes of writing differ across the disciplines. For instance, although argumentative writing is important within both biology and history, the things that count as evidence and how evidence is evaluated and employed in writing differ across these disciplines (Draper et al., 2010).

1.2 Historical Reading, Thinking, and Writing

Research during the past 25 years has yielded a great deal of information about the cognitive processes associated with historical reading, historical thinking, and historical writing. Of interest is the degree to which teachers can nurture disciplinary expertise within history students. Alexander (2003) shows that the Model of Domain Learning provides useful insights on the development of expertise. She argues that among other changes, students progress from surface level processing, which involves general reading/writing strategies, to deep level processing, which includes more discipline-specific strategies.

Applying Alexander’s ideas within the discipline of history, Seixas and his Canadian colleagues have identified six elements of historical thinking, including establishing historical significance, using historical evidence, and acknowledging the perspectives of historical actors (Seixas & Morton, 2006). Researchers in Sweden, the United Kingdom, and the Netherlands have identified metaconcepts, concepts that are

associated with historical thinking processes rather than the substantive concepts that represent historical content knowledge (Halldén, 1998; Lee & Ashby, 2000; Van Drie & Van Boxtel, 2008). For instance, in order to engage in sourcing, students must view texts as accounts and evidence. Instead, many novices conceive of texts as bearers of unadulterated information that they accept at face value. With instruction many students begin to view texts as accounts—information presented through a human author who has biases, insights, perspectives, an audience, and purposes. Students who view texts as evidence acknowledge that historical understandings are interpretive in nature and that textual evidence is required to support historical interpretations. Their writing differs from peers who have a less mature understanding of history. Researchers in Spain have observed the goal-driven selection and use of historical evidence by undergraduate students and historians (Limón & Carretero, 1998). Other Spanish researchers have discovered the value and challenge of “hybrid tasks” that involve writing based upon the reading of multiple historical documents (Mateos, Solé, Martín, Cuevas, Miras, & Castello, 2014). Their research, and other international studies illuminate and increasingly inform instruction in historical reading and writing.

In pioneering work in the United States, Wineburg (1991) identified sourcing and other heuristics used by historians to analyze historical evidence. Since then, sourcing, which is the application of source information to understand, critique, and use a document, has been assessed in numerous studies, many involving writing (De La Paz, 2005; Nokes et al, 2007; Reisman, 2012). Researchers have investigated how different types of instruction, such as computer software designed to allow students to practice sourcing (Britt, Perfetti, Van Dyke, & Gabrys, 2000); writing instruction integrated into historical thinking lessons (De La Paz, 2005; Monte-Sano, De La Paz, & Felton, 2014); teacher-delivered explicit strategy instruction (Nokes et al., 2007; Reisman, 2012), and cognitive apprenticeships (De La Paz, et al., 2012), among other instructional strategies, influence students’ engagement in historical thinking.

Although these studies present a wealth of information on students’ ability to use historical thinking strategies, they leave many questions unanswered. For instance, most studies focus on whether students engage in strategic thinking or not (see, for example Nokes et al., 2007). Less research has been conducted on how strategy use can contribute to deeper historical thinking. Certainly strategy use, such as sourcing, is related to sophisticated historical thinking, however, little research has been published that explores this relationship. For example, how should a teacher interpret a student’s use of seriously flawed evidence to bolster a claim (Wineburg, 1991)? The teacher might applaud the students’ efforts to support a claim with evidence or focus instead on his/her poor analysis of that evidence. Or, what does it mean when students think critically about a piece of historical evidence, but then fail to apply this critique to a larger historical question? How are teachers to interpret students’ historical thinking abilities when students critically analyze evidence but fail to use that analysis in their writing to bolster a historical argument?

Historians' writing includes a mixture of notes, critical summaries, descriptions, and narrations, but the keystone of historical writing is argumentative (Nokes & De La Paz, in press). Historians' writing is produced not merely to share information, as is often the case of writing composed by secondary history students (Duke et al., 2012), but to enter a conversation with academic colleagues (Graff, Birkenstein, & Durst, 2015). Through their writing historians share original interpretations that contribute to humanity's understanding of the past (Grafton, 1997). Their writing must serve two purposes: to share their fresh ideas, and to persuade their colleagues that their interpretations are accurate, meet disciplinary standards, and represent an important contribution to the body of historical knowledge (Grafton, 1997). Students engage in historical writing when they share and defend historical interpretations in a manner that matches, to the extent possible, the argumentative writing of historians (Nokes & De La Paz, in press). Argumentative historical writing is a "hybrid task" because it always involves the use of evidence, which must be read and analyzed. Thus, historical reading and historical thinking must be considered in any analysis of argumentative historical writing (Monte-Sano, et al., 2015).

1.3 The Development of Historical Thinking

Researchers in the United Kingdom and elsewhere have explored students' development of historical thinking, documenting trends that are observable as students gain experience and develop increasingly sophisticated ideas about history (Lee, 2005). These researchers suggest that patterns exist in the way young people think about history (Lee, Ashby, & Dickinson, 1993), a phenomenon that Halldén (1998) attributes to youngsters' physical and cognitive contexts. Lee and colleagues identify trends in students' progress in historical thinking, contending that young people's ideas about the past, allow or inhibit what they call "cognitive moves." For instance, only when students understand the idea of evidence are they likely to use evidence to support written claims in a historically appropriate manner (Ashby, Lee, & Shemilt, 2005).

By studying students between the ages of 7 and 14, Lee and colleagues (1993) discovered trends in historical thinking associated, to some degree, with age. Young students often viewed history as primarily factual but unknowable, because, as one fifth grader put it, "nobody alive today was there" (Lee & Ashby, 2000, p. 205). Some slightly older students still viewed history as factual but considered it knowable, though often unknown because of a lack of evidence. By eighth grade, some students began to look more critically at evidence, accusing writers of "lying" or of being biased. For these youngsters, historical thinking still consisted of knowing facts rather than constructing interpretations. One eighth grader explained, "I don't think we could find out definitely [what happened in the past] because there are only biased stories left" (Lee & Ashby, 2000, p. 206). By the end of elementary school, some students began to use naive strategies for working with evidence, such as counting the opinions found in the various pieces of evidence and ruling with the majority. By middle school a small proportion of students viewed history as interpretive, "a matter of opinion," as one sixth

grader explained. Further, this same student, as others did, began to consider the use of documents in establishing criteria for one's opinion (Lee & Ashby, 2000, p. 207). Of note, Lee and Ashby discovered that students often developed skills without the anticipated conceptual development. In other words, skills, such as sourcing, emerged more rapidly than did an understanding of why the skill was necessary for historical thinking. And although historical thinking generally improved with age, a range of abilities existed at each age.

For purposes of this study, the notion of historical development is again considered, with a search for patterns in eighth-grade students' ability to synthesize their evidence-based interpretations into a coherent written historical argument. The process of historical writing involves not only sophisticated historical thinking, but a knowledge of how to display that thinking in writing (Berninger & Swanson, 1994). Questions arise about how students can demonstrate certain historical thinking skills in their writing, and how writing can help teachers diagnose missteps in historical thinking. This study investigates how students (a) use texts to defend a historical interpretation, and (b) write critically about documents' reliability and usefulness in solving a historical dilemma.

1.4 Writing and Thinking

Researchers on writing have considered the ties between writing and cognitive activities both within and outside of the discipline of history. In a series of studies, Wiley and Voss (1996, 1999) found that undergraduate students produced richer, more sophisticated writing, when assigned to write an "argument-based essay" than when asked to write a "history" or a "narrative." Further, these researchers found that students engaged in a higher level of critical analysis when they were given multiple texts than when given a single document to analyze. In Spain, Martinez and colleagues (2015) found similar epistemological advantages when teachers assisted older students as they engage in reading/writing hybrid tasks (Martinez et al., 2015). They found that teacher support was a vital factor in students' success. Swedish researchers used students' writing to assess their historical consciousness (Eliasson et al., 2015), a construct that includes the students' use of historical evidence to develop a frame of reference to orient their identity and to comprehend current issues and events. The instrument they designed used students' writing to assess their ability to use the metaconcepts of continuity and change to write persuasively about Sweden's demographic future. It should be noted that in the case of Wiley and Voss's research as well as that at Eliasson and his colleagues, students' understanding of the writing prompt, the nature of various types of writing, and their metaconceptual understandings appeared to be vital factors in their success.

Klein (1999) contends that producing a particular genre of written text, such as argumentation, comparison/contrast, analogy, or personal writing, is a cognitive activity. However, in an early review of research he indicates that merely assigning students to produce a certain genre of writing does not always lead to the intended cognitive activities. In order for the target thinking to occur, students have to

understand the nature of the task they are asked to do and engage in that task (Penrose, 1992). For instance, if students are assigned to write a critical analysis and they instead produce a summary, the target thinking is not achieved.

The current study integrates these three fields of research: a) historical reading, thinking, and writing; b) the development of increasingly sophisticated historical thinking skills; and c) the link between writing and cognition. Additionally, recent research on the assessment of historical thinking brings these fields together, providing instruments that help teachers gauge students' historical thinking through various writing prompts (Ercikan & Seixas, 2015; Smith & Breakstone, 2012; VanSledright, 2014).

This study focuses on the following questions: (a) What patterns exist in eighth graders use and/or misuse of sourcing to analyze historical evidence? (b) What range of sophistication do eighth graders exhibit in their use of sourcing in argumentative writing? What are the common strengths and weaknesses they exhibit in their historical writing? (c) How can written assessments help teachers diagnose potential barriers to the historical thinking of eighth graders? Aided by answers to these questions, it is the aim of this study to provide teachers with a more nuanced understanding of the common errors in students' understandings of history and resources to nurture students' ability to analyze and use historical evidence in their writing. Further, this study should help teachers to design improved assessments and rubrics to diagnose students' particular struggles in historical writing.

2 Methods of Inquiry

2.1 Participants

Participants in this study were the 427 eighth-grade students being taught by eleven U.S. History teachers located in a suburban school district in the Intermountain Western United States. Some teachers taught a single U.S. History class and others taught multiple classes. Teachers taught at seven schools that differed in size, racial make-up, and socio-economic status, with two comprised of racially and economically diverse students and five predominantly White and middle class (see Table 1). Enrollment in the schools ranged between 564 and 707 students. Teachers were recruited during a mandatory professional development on fostering historical thinking. They were trained at the all-day professional development on methods for teaching cognitive strategies for reading documents, including sourcing and using documents as supporting evidence in historical argumentation. Teacher participation in the study was optional but encouraged by the district's social studies specialist. Six of the eleven participating teachers, seven of the sixteen classrooms, and 206 of the 427 participating students came from the district's most diverse schools.

The 427 eighth graders who participated in this study represented 35.7% of the 1,196 eighth graders in the school district. They were 12 or 13 years old. Participating

students represented a cross section of ability levels, reading levels, language background, and ethnicity common within a suburban school district in the Intermountain West. Percentages of students reading at a proficient level, as measured on the Student Assessment of Growth and Excellence (SAGE) exam, ranged from 40% to 56% across the schools, compared with state averages hovering around 41% (see Table 1). Thus the classes that participated in this study ranged from significantly above average in reading ability to slightly below average, even though the range of abilities within classes was much greater. The representativeness of this sample will be considered below.

Table 1. Demographic Information of Schools Participating in Study

School (enrollment)	Number of teachers	Number of classes	Racial make-up of eighth-grade class	English Language Learners	Percentage economically disadvantaged	Percentage Proficiency on ELA SAGE test
School 1 (707)	2	2	54.1% White 31.2% Latino	9%	53.7%	40%
School 2 (564)	1	2	82.3% White 10.5% Latino	fewer than 1%	27.0%	54%
School 3 (640)	2	3	44.9% White 39.8% Latino	22%	57.5%	42%
School 4 (607)	1	3	86.2% White 6.6% Latino	fewer than 1%	14.4%	55%
School 5 (595)	2	3	78.9% White 12.4% Latino	2%	29.4%	56%
School 6 (475)	1	1	84.6% White 5.4% Latino	Fewer than 1%	17.3%	55%
School 7 (610)	2	2	72.6% White 17.8% Latino	8%	42.8%	44%

2.2 Instruments

At the start of the school year participating teachers administered a series of assessments to their students, one of which served as the data source for this study. Although the teachers had been trained on methods for nurturing historical thinking, the test was administered at the start of the school year, before any instruction had been given on historians' strategies, as a pretest associated with a larger, ongoing study.

Teachers were instructed to read the assessment instructions aloud with their classes and answer technical questions (i.e. how many words does our answer need to be?) but to avoid providing answers to conceptual questions (i.e. how are we supposed to use the documents in our essays?). The purpose in avoiding such discussions was to assess students in the absence of instruction on historian's writing.

The assessment (see Appendix A) consisted of a series of writing prompts associated with the texts and activity Wineburg (1991) used in his pioneering study on historians' heuristics. Students were given 7 short accounts of the Battle of Lexington from reliable and unreliable sources, with one painting of the battle. Accounts included primary and secondary sources from American and British perspectives, an American textbook, a historical novel, and a historic newspaper account. Accounts had been edited for Wineburg's study to eliminate information that was not central to the problem at hand. The passages ranged in length from 129 to 290 words and ranged on the Flesch-Kincaid assessment of reading level from 6th to 12th grade, though the results of this assessment may have been skewed by the fact that the passages measured at the 12th grade level consisted of run-on sentences common in 18th century prose (see Table 2). It should be noted that the source of each passage was provided at the bottom of the page as in Wineburg's study. The purpose for positioning the source information there was to observe whether students would intentionally seek out information about the source of the document. It was assumed that students would be more likely to pay attention to and write about the source if it was provided above the passage. The assessment instructions asked students to produce a written evaluation of the accuracy of the painting after "analyzing the documents as a historian would." The writing prompt explicitly asked students to "write about the documents" and to "try to write as a historian would."

Table 2. Source, Word Length, and Flesch Kincaid Reading Level of Assessment Passages

Passage characteristics	Doc. 1	Doc. 2	Doc. 3	Doc. 4	Doc. 5	Doc. 6	Doc. 7
Source	Patriot deposition	U.S. historical novel	British officer journal	London newspaper	Yale President diary	U.S. textbook	British soldier memoir
Word length	248	169	232	222	290	129	153
Flesch-Kincaid level	12	6	8.5	12	7.5	8.9	10.9
Words per sentence	123	15.1	57.5	55	26.1	15.8	30.2

After writing about the painting, students were asked to evaluate in a sentence or two the relative reliability and usefulness of the documents. The purpose for having students write in this sequence was to see whether students would evaluate the sources of

documents without being prompted to when first analyzing the painting. It was assumed that asking students to evaluate the relative usefulness and reliability of the evidence prior to having them write about the painting might have promoted more critical analyses. The intent of this study was to investigate students' writing without such prompting. The assessment instrument had been piloted with 11th and 12th grade students and assessed for reliability and validity in an earlier study that involved over 200 students and multiple reviewers (Nokes, et al, 2007).

2.3 Data Analysis

Students' writing, both their analysis of the painting and their evaluation of the documents, was analyzed, with a record kept of incidents of sourcing. Sourcing was operationally defined as any writing about the source of a text when source information was used to try to understand or critique the contents of a document or to bolster a claim, even when the analysis demonstrated flawed background knowledge (see Appendix B for the coding guidelines that were followed). Evaluators proceeded sentence by sentence through the students' writing, highlighting those sentences that included students' use of sourcing. Merely mentioning the source was not sufficient to count as sourcing. Instead, students had to evaluate a passage based upon a) the author's position, motivation, participation, or background, b) the date the text was produced relative to the event it described, or c) the type of text the passage was taken from (i.e. a journal, a sworn deposition, or a novel).

Two reviewers coded 75 of the students' written assessments to establish interrater reliability. Using SPSS the reviewers' sentence by sentence coding was compared. Cohen's kappa proved to be significant, indicating a strong agreement between reviewers ($k = .815, p < .001$). Additionally, points of disagreement were discussed until agreement was reached. Once this high level of reliability was reached, one of the two reviewers scored each of the remaining assessments.

Once incidents of sourcing had been identified, through a process of open coding (Strauss & Corbin, 1990), two reviewers explored patterns in the nature of the sourcing relative to the analysis of the painting, with common characteristics of students' processing of the sources identified inductively. Both reviewers coded 25 essays with the purpose of noticing patterns in students' writing. Reviewers came up with five main patterns: a) students mentioning nothing about the sources of documents in their writing, b) students using ahistorical thinking to evaluate a source, c) students listing a source without critique, d) students providing a historical review of sources without using the review to strengthen their analysis of the painting, and e) students critiquing the sources of documents and using the critique to bolster their interpretation of the painting. Once the categories had been agreed upon, both reviewers evaluated 75 students' essays and placed students in one of the five categories. Their placement of the students was compared using SPSS, and Cohen's kappa indicated a strong agreement between reviewers ($k = .841, p < .001, n = 75$). A single reviewer used the students' assessments to place the remaining students into one of the five categories.

Only two students, identified through discussion between both reviewers, did not fit into one of the five categories and are described below.

Considering the results of the open coding, the reviewers developed a progressive spectrum of levels of development (see Figure 1) through a process of axial coding. The reviewers agreed that the most sophisticated use of sourcing was identified by students' tendency to critically evaluate documents (level 4) and to use that evaluation in their written critique of the painting (level 5). Additionally, both reviewers agreed that the failure to make any mention of the source represented the least sophisticated historical thinking (level 1). The reviewers had a more difficult time distinguishing between the level of sophistication of students who noticed the source but did not critique the source and students who used ahistorical thinking to critique a source. After continued discussion and by returning to the students' writing, the reviewers determined that students who attempted to critique a source but used ahistorical thinking in the process appeared to have a better view of the nature of history than those who offered no critique. Through this process levels 2 and 3 were added to the spectrum (see Figure 1).

Next, the reviewers returned to the students' writing to find the correlation between more or less sophisticated use of sourcing and other evidence of historical writing. Using a Spearman's rank correlation test, students' position on the spectrum was compared with the frequency of students' writing about sources of documents. This analysis was conducted to make certain that more sophisticated levels of sourcing were positively correlated with another indicator of historical thinking, namely the frequency of reference to sources. The Spearman two-tailed test showed a strong correlation between students' identified level and the frequency of their use of sourcing ($\rho = .864$, $p < .001$, $n = 427$). A second Spearman test was conducted with levels 2 and 3 reversed to assess whether level 3 indeed represented more sophisticated thinking. This second test showed a nearly identical correlation ($\rho = .861$, $p < .001$, $n = 427$), and we chose to keep our original placement of the second and third levels.

In addition, the assessment instrument was studied by evaluating patterns in the location of students' sourcing in their writing. For example, were there certain writing prompts that tended to elicit sourcing more than others. In order to investigate this question, students' sourcing score was broken down into occasions of sourcing in their analysis of the painting, and occasions of sourcing in the follow-up questions. Descriptive statistics were used to compare the relative value of various sections of the assessment instrument in drawing out students' sourcing.

3 Findings

The purpose of this study is to investigate (a) patterns in eighth graders use and/or misuse of sourcing in their argumentative writing, (b) the range of quality in eighth graders use of sourcing and argumentative writing, and (c) how written assessments can help teachers diagnose strengths and weaknesses in students' historical thinking.

3.1 Patterns in Eighth Graders Use of Sourcing

When it came to sourcing, students exhibited a wide range of strategy use in their writing. On one hand, over 40% of students failed to write critically about sources, even when prompted to consider the reliability of the documents. On the other hand, some students peppered their writing with as many as 15 critical comments about the sources of documents. The mean number of critical references to sources was 2.63 (SD = 2.736). Not only was there great variation in the quantity of sourcing, there were also stark differences in the quality. For example, Blaine (all names are pseudonyms) was critical of the novel, writing “it is a secondary source and it is from a novel. A novel could have fiction anywhere in it, and you wouldn’t know.” In contrast, Emily wrote about the passage from a textbook, explaining that she trusted document 6 because, “it is a textbook so it should teach the correct history.” In both instances students took into consideration the source of a document to critically evaluate its reliability, although Blaine’s sourcing demonstrated a nuanced understanding of genre that Emily’s lacked.

3.2 The Range of Sourcing Quality

As mentioned, through the qualitative analysis of data, coders inductively constructed a five-level spectrum of sourcing that captured patterns in the quality of students’ use of sourcing in their writing.

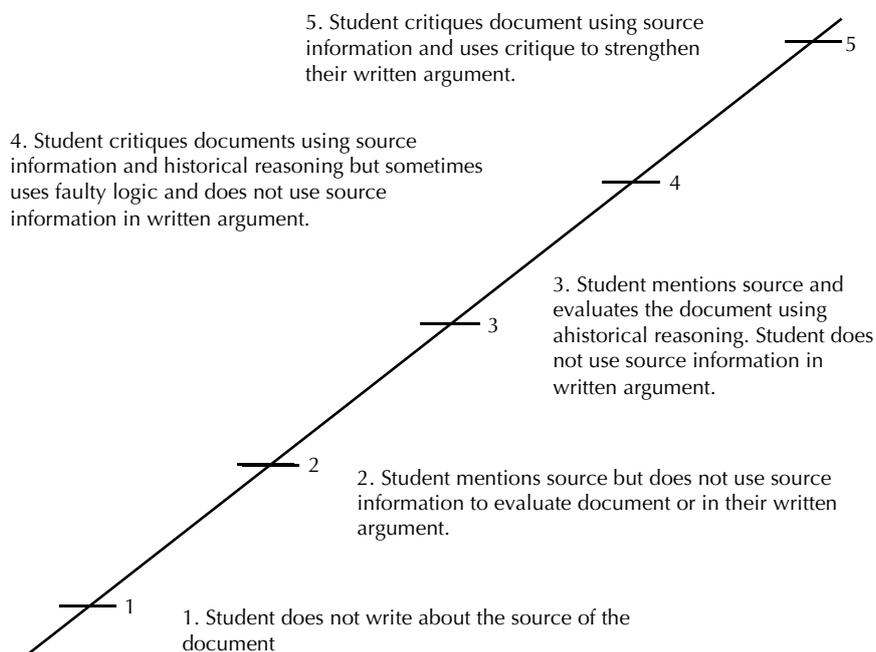


Figure 1: Sourcing spectrum with defining characteristics of levels 1-5.

The lowest level was indicated by students remaining oblivious to the source and accepting information in the document at face value. This was common among the eighth-grade students, with 41.3% failing to make any mention of the source of any of the documents they wrote about. When prompted to identify the documents that they thought were most reliable, a representative response from the first level was given by Rachel who trusted document 1 [a deposition] because “it tells more info,” and Ben who rated document 2, a novel, reliable because “it was specific and I understood it clearly.” One student rated two texts as equally reliable because, as he explained, “I could picture what they were talking about really well in my mind.”

At this lowest level, texts were appreciated for their detail, or criticized for their complexity. “It confused me,” one student wrote, explaining why he had labeled a text unreliable. Some students trusted texts without question: “they state exactly what happened,” Jasmine wrote to explain why she found certain documents to be reliable. Spencer appreciated two texts because they contained “the cold, hard, truth.” Another contended that he valued texts that “seem like they don’t have any bias and just info.” To summarize, students who were evaluated at the lowest stage of sourcing judged texts based upon their detail, the amount of information they contained, their seemingly unbiased content, and their clarity, without any mention of the source. Their argumentative writing lacked the leverage that can be gained when sources of documents are critically evaluated and that evaluation is explained in writing. There was a rare exception to this trend, which will be discussed below.

The second level was reached by students who noticed and mentioned the source of a document but did not critique its content. For example, when asked which documents she found least useful, Tiffany explained, “Document 6 because [it] is from a textbook.” Few students, only 5.9%, were classified in this second level because most students who wrote about the source of text also evaluated it. The students in this category made comments like Heather’s, who trusted document 1 because “it distinctly states what happened in the colonists’ point of view.” She had looked at the source and thought about the perspective it represented but did not use that information to critically review its content. She thought the text was reliable because “it distinctly [stated] what happened,” not because it came from eye-witnesses. Of note was the pronoun that she used—“it distinctly states.” She apparently viewed the words as having come from a deposition rather than from the colonists who watched the event unfold and subsequently produced the deposition.

The third level was identified by students noticing the source of a document but using ahistorical thinking to evaluate it. As with the second level, there were few students, 7.8%, who fell into this category. For instance, some students recognized that two of the accounts of the Battle of Lexington came from eye-witnesses, but they discounted them because of the potential bias of people involved in the battle. One student, Angie, ranked document 3, a diary entry from a British officer, as one of the least reliable and least useful resources. In her explanation she suggests that the writer seemed uncertain in some of the details. To Angie, document 6, the textbook account,

was much more certain of its content and was thus more reliable and useful. Students in the third sourcing level sometimes trusted the textbook, as one student wrote, “because it’s in a textbook so it won’t [favor] one side too much.” Other students, like James, became highly critical of sources that admitted uncertainty. He criticized documents 2 [a novel] and 7 [a memoir] because “both writers continue to say ‘to the best of my remembrance’ or ‘to my recollection.’ These statements tell me that they do not remember as well as others would therefore it may be incorrect.” While memory certainly was an issue in the memoir, it was not a problem for the author of the novel, who used the fictional character’s admissions of uncertainty, ironically, to create a feeling of authenticity within the account. Echoing James’ concern, Trevor wrote, “I don’t trust Doc 5 [a diary entry] because he said he wasn’t sure and if you don’t know [then] you are sort of already a liar.”

Documents were trusted or mistrusted for a variety of reasons. Amber didn’t trust some texts because they “only show one side” of the issue. Conner labeled two texts as most reliable because “they have more thought behind them since they were written by a college professor and a textbook.” Echoing this idea, Matt liked the novel “because it was written from an expert’s point of view instead of someone who might have a bias against one other side.” To summarize, at the third stage, students noticed the source of the document and critiqued the content based on the source, but did so in a manner that represented ahistorical thinking. They consistently feared the bias of the eye-witnesses who produced a deposition and journal entries, favored the textbook and the novel because of their distance from the event, and got hung up by admissions of uncertainty in the memoir. It should be noted that when students included both appropriate and inappropriate uses of source information, and there were many students who did, the coders placed them at the level, either level three or level four, that represented the preponderance of their written critiques.

Students at the fourth level, which included 39.2% of them, noticed the source of a document, used sound historical thinking to evaluate it, but did not use the source information to support their claims about the painting, their main writing task. For example, when Mallory was prompted to write about which texts she viewed as most reliable she wrote, “I trust 3 the most because 3 was written by an officer in the British army, so he probably knows more about it because he was involved.” But Mallory had previously argued that the painting was accurate without employing document 3 to support her claim. Gayle was insightful enough to note that she didn’t find document 2 useful “because in novels like #2 people will make the story more or less dramatic.” However, Gayle judged the painting accurate because it showed the British firing first and “in the second document it was stating that the British actually fired first.” So, in spite of her subsequent reservations about its trustworthiness or usefulness, Gayle had previously cited document 2 just as she did other documents to support her evaluation of the painting.

It should be noted that students at the fourth sourcing level often employed sound historical thinking in their analysis of documents. For instance, many students valued

primary sources over secondary sources, many even using those terms. Primary sources were viewed as reliable because they were “told from first person” or “from a 1st source aka primary source.” Paul went against his patriotic sentiments when writing about the diary of a British officer, admitting, “I trust it the most even though it is the British writing.” When asked which documents he thought were most reliable, Parker wrote,

I would say documents 1 [deposition] and 3 [journal entry], because they are diaries from those who witnessed the events and saw w/ their own eyes the truth. I trust those the most because they’re not trying to convince someone that their story is more valid than the rest. It’s just their recollection of what happened and they didn’t think, ‘Hey, future generations will read this, I better make it sound like we’re innocent.’ No, it was just journal entries.

Parker’s comment is representative of a common pattern among students at the fourth level whose writing juxtaposed flashes of brilliance with naïve, unsophisticated worldviews, and indefensible claims. Just as Parker recognizes document 1 is from eye-witnesses, a source he values, he fails to acknowledge that it was a sworn deposition, a very public document, created to be shared with current, if not future, generations. Further, he fails to understand that journal entries are sometimes created with the intent that they can be shared in order to publicly defend an action. Further, he celebrates what he perceives as a lack of bias in these very biased accounts, failing to see that both accounts are indeed reliable and useful, not because they lack bias, but in spite of the bias they represent.

Students at the fourth level also noticed the timing of the creation of the documents. Brianna thought some texts were reliable because, “they wrote them about the date of the battle.” In one of the more curious analyses, Candice integrated faulty information about the source with accurate content of the document in her assessment of its reliability: “Document 4 [British produced newspaper] seems like it’s a reliable source because it was written by patriots and it states facts that could [go] against them, showing honesty.” If indeed the newspaper had been written by patriots her assessment would have been extremely insightful. However, with her misread of the source information it left the reviewers a bit puzzled about her historical thinking.

Students at the fourth sourcing level were sometimes critical of texts for good reasons. Many doubted the reliability of secondary sources. For instance, Scott questioned the reliability of two texts “because there written by someone else, not the people that actually experienced it,” or Trevan who explained “they were written in future times by people who weren’t even there.” Brandy wondered about the reliability of secondary sources because “I don’t know if they are guessing like me or if they really know.” The timing of the production of the text was also an issue for some students. Curtis explained that he didn’t trust “document 2 [novel] or document 6 [textbook] because it wasn’t written on the date around the date it happened.” In summary, students at the fourth stage used some sound reasoning to critique documents but failed

to apply their criticism to their task of analyzing the painting. Most students simply failed to provide any evidence from the documents in their written analysis of the painting.

Some students in the fourth level made a feeble attempt to use the documents as evidence in the defense of their assessment of the painting. However, students who did so made no distinction between those documents they found reliable or unreliable in their subsequent assessment of the documents. Other students wrote generic phrases about the documents, such as Andrew's argument, "the documents all show that the British fired first," suggesting a generic cherry-picking of evidence. Although some documents do make the claim that the British fired first, there are other compelling documents in the collection that make the opposite argument. Many of the students in the fourth level made no distinction between the subtle but significant differences between texts, citing "all of the documents" or "most of the documents" to support their claims even though the passages contained conflicting content. No student at the fourth level supported their claims about the painting's accuracy by using the specific texts that they found to be most reliable.

Finally, students at the fifth level integrated into their assessment of the painting, not only information from the documents, but a critical review of the sources of the documents. Only 5.9% of the eighth graders engaged in this level of historical thinking. Students at this fifth level wrote with a degree of sophistication that their peers lacked. For instance, Jamie explained why she trusted document 3 with the following explanation. "Document 3 is a journal entry of Lieutenant John Barker, who personally experienced the battle, and put it in his own words of what took place April 19th." In her analysis of the painting she quotes document 3, showing how it supports what the artist has drawn. Similarly, Carter explains that "I am led to believe [the British fired first] because the most reliable sources state so." He goes on to explain why he finds some of the sources reliable. Megan reported in her analysis of the painting that, "the documents that I thought were true describe it like that..." Joey based his analysis of the painting on its portrayal of the British "firing first." The documents were then employed to identify whether the British did indeed fire first. He concludes,

I believe that the British were the first ones to fire. Nathaniel Mulliken, Philip Russell, and 32 other men said [that they did]. Also in the diary of the Lieutenant John Barker he said [his men fired]. I feel I can trust these, they seem like good resources. In the end I believe the British soldiers fired first.

Darrin, Carter, Megan and Joey, however, represented a small minority of students who both critically considered the sources of documents and employed sourcing to write persuasively about their interpretation of the painting. Their writing differed substantially from that of students at other levels. To summarize, five levels of sourcing, representing different levels of historical thinking, emerged from the analysis of students' writing and are represented in Table 3.

Table 3. Summary of Descriptions, Data, and Indicators of the Five Levels of Sourcing

Level	Description	Percentage of Students	Primary Indicators
1	Student does not write about the source of the documents.	41.3%	Texts critiqued based on clarity, amount of information, possession of the truth, lack of bias.
2	Student mentions source but does not use source information to evaluate documents or in their written argument.	5.9%	Source or point of view is stated without explanation or critique.
3	Student mentions source and evaluates the document using ahistorical reasoning. Student does not use source information in written argument.	7.8%	Eye-witnesses accounts discounted because of bias. Admissions of uncertainty cause alarm. Expertise misplaced.
4	Student critiques documents using source information and historical reasoning but sometimes uses faulty logic and does not use source information in written argument.	39.2%	Students favor eye-witnesses accounts produced shortly after an event but engage in some misreading and fail to use trusted sources rather than less reliable sources to support claims.
5	Student critiques document using source information and uses their critique to strengthen their written argument.	5.9%	Students favor eye-witnesses accounts produced shortly after an event and use them to write persuasive arguments.

As mentioned, the written response of two students failed to fit into this spectrum. These students demonstrated some skill in using the documents as evidence to support their assessment of the painting without addressing the source of the documents. They may have sensed which sources to trust but, when asked to relate why they viewed these sources as reliable, they appeared unable to articulate an answer in terms that would be recognized as historical thinking. With so many participants, it is possible that some happened to choose by coincidence primary sources that strengthened the defense of their interpretations. Or it might have been that they intuitively sensed that these texts were preferred, but did not understand why. In their use of documents to

bolster their claims, these two students' historical thinking matched their peers at the fifth level. However, in their analysis of sources their thinking mirrored students in the first level. These two students did not fit into the spectrum for which over 99.6% of the eighth graders did.

Additionally, the writing of Diane hinted at a possible sixth level that no other eighth grader reached. She wrote, "In document one, 34 men swore our [American] side did not fire, but they must've not known. Other documents, mainly from the British side, said they [the British] were shot at." Here Diane includes evidence that goes against her interpretation, but contests it by claiming, "they must've not known." Citing the "other documents, mainly from the British side," she shows that the accounts she trusts tell a different story. The sophistication of Diane's analysis was unique among the students who participated in this study.

3.3 Assessments of Historical Thinking

In addition to exploring patterns in students' historical writing, this study also yielded information about how assessments can help teachers diagnose strengths and weaknesses in students' historical thinking. As described above, the assessment consisted of two parts. First, students wrote a critical analysis of a painting. In doing so they were instructed to "imagine that you are a historian and analyze the documents as a historian would," "use the documents," "read the documents as a historian would," "be sure to write about the documents," and "try to write as a historian would." In spite of these guidelines, only 5.9% of students included information about the source of any document to bolster their interpretation of the painting. Of the 1,135 total incidents of sourcing, only 27, or 2.4% were written in response to the first writing prompt. In contrast, on the second part of the assessment, during which students were asked to rank the reliability and usefulness of the documents and to justify their answers, 57.2% of students included an evaluation of at least one source in their writing. In fact, 1,108 of the 1,135 incidents of sourcing, or 97.6% were written in response to the second part of the assessment. These patterns in students' responses provide insights into the design of assessment instruments intended to identify students' ability to engage in historical reading, thinking, and writing, a topic that will be discussed below.

4 Discussion

This study reports the results of research on patterns in the writing of 427 eighth-grade students with a focus on their use of the strategy of sourcing to analyze evidence and substantiate their claims in a written historical argument. This discussion includes (a) insights on students' use of sourcing in historical writing, (b) insights on the development of assessments of historical thinking, (c) limitations of the study, and (d) concluding thoughts.

4.1 Students' Use of Sourcing in Historical Writing

Findings of this study show that students varied widely in their ability to write about the sources of documents and to use source information to defend a historical interpretation. With only 5.9% of the students integrating source information into their historical writing, it appears that many students may not understand the role of argumentation in historical writing (Nokes, 2014; VanSledright, 2014). To explain, students who view history as the past, known or unknown, factual, and learned through memorization, report events as a collection of facts rather than as an evidence-based interpretation (VanSledright, 2002; Lee, 2005). Without an understanding of history as interpretive and evidence-based, students have little reason to evaluate evidence or use it to attempt to persuade a reader that their ideas have merit. Additionally, the findings of this study suggest that other barriers might interfere with students' historical writing.

For example, students who did not mention sourcing in their writing appeared to be satisfied with literal comprehension and did not seek the in-depth, discipline-specific comprehension associated with the critical analysis of evidence. Texts that were clear, easily comprehended at a surface level, full of detail, and rich in imagery were considered by many to be "more reliable," regardless of the source of the passage, a conclusion also reached by students in Wineburg's (1991) pioneering research. Many students in both studies accepted the information that documents contained without critical thought. As VanSledright (2002) argued, perhaps the focus on literal comprehension in the early grades creates, in the minds of students, the notion that deeper, critical comprehension is unnecessary. For them, literal comprehension becomes the goal of reading and, by extension, demonstrating literal comprehension is the ultimate end of writing. For these students, surface level literacy strategies such as monitoring comprehension, rereading difficult passages, or creating mental images (Alexander, 2003) appeared to have been the cognitive strategies employed during reading, for they praised, as reliable, those texts that were easier to comprehend. Missing were the deep level processing strategies associated with the Model of Domain Learning (Alexander, 2003). Just as Wineburg had discovered in 1991, for some students in this study, the textbook's factual content was above criticism. Students at the first level were not critical unless a text was difficult to understand, in which case their criticism revolved around its lack of clarity rather than its source or content. And they more frequently referred in their writing to the texts that they found to be straightforward.

Additionally, a handful of students, primarily those at the third level, wrote about texts' apparent level of omniscience and objectivity. Admissions of uncertainty, fairly common among eye-witnesses (including some reliable documents in this study), were a major cause of distrust for these students. This tendency to doubt texts that admit uncertainty is likely caused by students' familiarity with textbooks and other authoritative sources that use a voiceless, omniscient tone (Paxton, 1999). One difference in this study than earlier studies, such as Wineburg's (1991), is that a minority (only 7.8%), rather than a majority, of students in this study expressed a

preference for the textbook account over primary sources. Wineburg found the opposite to be true: that even high achieving high school students placed greater trust in the veracity of the textbook. Could it be that the increasing emphasis on primary sources and disciplinary reading and writing, found in the Common Core State Standards (CCSS, 2010), the C3 Framework (NCSS, 2013), history pedagogy texts (such as Nokes, 2013; Lesh, 2011), and in a growing body of research (De La Paz, et al., 2012; Reisman, 2012) is leading to greater exposure to primary sources in earlier grades? The common use of the terms “primary source” and “secondary source” in students’ writing suggest that some students appear to have received such instruction in earlier grades. If young students are being exposed to some elements of historical reading, thinking, and writing, leading to distrust of the textbook, this current research makes clear that such instruction is falling short in nurturing students who can write critically about documents and who use source information to support an interpretive claim in their writing. Future research might look into students’ evolving view of history textbooks and, if a trend is discovered, how this might translate into improved argumentative historical writing.

In spite of most students’ failure to use the documents in persuasive writing, many of the students (45.1% if stages four and five are combined) demonstrated some skill in sourcing in their writing, including a handful who did so with a high degree of sophistication. This tendency echoes a finding of Lee and Ashby’s (2000) that students sometimes learned skills at a quicker rate than they were able to understand the purposes of historical reading, thinking, and writing, for which those skills were necessary. In other words, a teacher might train students to pay attention to a source, without helping them fully grasp the vital role of sourcing in analyzing documents, differentiating between solid and spurious evidence, or writing a persuasive historical argument. In the case of the current study, many students engaged in sourcing, though few appeared to know how sourcing could bolster their written claim. Although the students’ use of sourcing is a cause of modest optimism, the results of this study show that teaching skills for analyzing evidence is not enough. Students must understand the nature of disciplinary writing if they are to employ those skills in a manner that meets disciplinary norms. As Lee and Ashby explain, history “is not a set of generic ‘skills’ that can be improved by practice, but a complex of multitrack understandings” (2000, 216). Discipline-focused writing instruction, as Monte-Sano (2008), Martinez and her colleagues (2015), and others suggest, is crucial in helping students go beyond the analysis of evidence in order to produce an argumentative historical essay that synthesizes students’ analytical reading (De La Paz, et al., 2012).

4.2 Designing Assessments of Historical Thinking

The results of this study provide insights for the design of assessments of historical thinking. The vast majority of students in this study, including the 94.1% in the first four stages, did not seem to understand the writing prompt that included the instructions: “try to imagine that you are a historian,” “analyze the documents as a historian would,”

“read the documents as a historian would,” “try to write as a historian would,” “use the documents,” and “be sure to write about the documents.” Why, with explicit directions like this, would so few students use the documents to write persuasively about whether the accompanying painting was accurate? The simple answer is, most likely, that students do not understand how historians analyze, read, or use documents in their writing. Such was the case of 5th grade students who, when asked how historians spent their time, had little inkling of historians’ work (Nokes, 2014). As Klein (1999) and Penrose (1992) suggest, students will not engage in the targeted cognitive activities when they do not know how to write within a particular genre. In short, in order for historical writing and historical thinking to occur, students must understand the nature of the task they are being asked to do. As a result, any effort to help students write historically must include instruction on the nature of history; how historians construct historical understanding; the role of evidence in historical reading, thinking, and writing; the nature of historical interpretations; historiography and the ever-evolving understanding of historical events; and the role of persuasion in historians’ writing. Without an understanding of the work of historians, including the purpose of their writing, it seems unlikely that students would be able to engage in historical writing at more than a superficial level, a notion also argued by Monte-Sano (2014). Thus, a simple explanation for the phenomena observed in this study is that most eighth-grade students simply did not know what the writing prompt meant when they were asked to read, think, and write like a historian. As Klein (1999) suggested, students cannot engage in a writing task they do not understand. Or, as Shemilt (1983) has been arguing for over three decades, the study of history in secondary schools continues to favor content knowledge over disciplinary knowledge. This study corroborates the notion that most eighth graders continue to misunderstand the nature of history as a discipline.

Of particular note is the fact that the second part of the assessment elicited sourcing at a much higher rate than did the first part. In fact, 97.6% of the incidents of sourcing were written in response to the prompts in the second part of the assessment. Had either part of the assessment been administered alone the findings on students’ sourcing would have yielded widely different results. Administered together students’ widely different scores on the two parts demonstrate the influence of the writing prompt in drawing out certain kinds of thinking and writing. In some ways, the results of this study say as much about the assessment instrument as they do about the students. It is possible that if the order of prompts were reversed, with students writing a critique of the documents prior to writing their analysis of the painting, they would have taken that critique of the source into consideration in their argumentative writing. However, without being prompted to first critique the evidence, only a handful included sourcing in their argumentative writing. The implications for assessing students’ historical thinking are evident. A teacher must design assessments carefully and tailor them to meet his/her specific instructional objectives and the level of the students. It might be that students who are in the early stages of learning to read, think, and write historically might be better evaluated by developing an assessment that provides more support—

perhaps asking for a written analysis of the sources before asking for an argument. However, for students who are more advanced, the sequence employed in the current study might provide greater evidence of students' sophisticated historical writing abilities. Given the sequence here, students were less likely to stumble onto the strategies they were not already fluent in using.

The nature of the documents used in this assessment may have also contributed to the writing produced by students, including their use of sourcing. Had other documents been used to study a different historical event, students may have been more or less likely to engage in sourcing, employ other heuristics, or to use documents as evidence to bolster a written argument. As mentioned, a different writing prompt might have elicited different elements of historical thinking. For example, it would have been unexpected to observe historical consciousness in students' writing given the writing prompt used in this study. In short, the findings of this study may not be generalizable across historical writing tasks, and comparisons across studies that used different assessment instruments must be made with caution.

The spectrum described in this paper might provide insights for the design of assessments of historical thinking. The identification of patterns in the strengths and weaknesses of students' writing can help teachers and researchers prepare rubrics to assess students' historical writing, with clear descriptions of stronger and weaker examples of sourcing. With knowledge of this spectrum of writing about sources, teachers could better identify elements of students' historical thinking and score students' writing. It should be noted that this spectrum can best be considered a continuum of sophistication rather than steps or stages of development. Numerous students, like Parker, who included both sophisticated and underdeveloped notions in his analysis of the deposition and journal entries, would best be placed in the gray areas between levels.

Further, the notion of a spectrum of sourcing might be useful in the development of multi-year curricula that include historical skill development, with students in the early grades being taught to notice the source of documents, recognize that different people might view the same event differently, and understand how historians spend their time. Students in middle grades might be taught how to evaluate sources, and to create a list of evidence that supports a claim. With that preparation in the early and middle grades, teachers can eventually teach students to use source information within formal argumentative writing. Lee and his colleagues have established a model for the development of historical thinking (1993). Perhaps the current study and other investigations of patterns of historical strategy-use and writing might inform the models that they have previously established. This study suggests that students who do not fully comprehend the purpose of persuasion in historical writing might still be able to incorporate elements of perspective in their critique of texts.

4.3 Limitations

In addition to concerns about the ability to transfer results across studies that use different assessment instruments, mentioned above, questions exist about the representativeness of the participants in this study. How well do the eighth graders in this study represent students in other contexts, both within the United States and globally? Replication of the study in diverse settings would provide evidence on whether the patterns observed in the writing of students in the suburban Intermountain Western United States are similar to the tendencies of students elsewhere. Certainly the students in this study represented a wide range of reading abilities, even if they did not represent the ethnic and linguistic diversity common in many schools.

In addition, questions exist about the bipolar nature of the placement of students within levels of sourcing, with 41.3% in the first level and 39.2% in the fourth. Does this placement indicate flaws in the assessment instrument or the qualitative analysis of the data? If not, how can this trend be explained? It could be that epistemological factors might channel students into certain ways of thinking about texts, resulting in their uncritical acceptance of texts at face value or their tendency to begin to question text content. Perhaps when students notice the source of a document they instinctively begin to think more critically about it, which would explain why students were funneled into the first or fourth levels. Or perhaps students' reading ability might have played a factor—with struggling readers focusing on the literal comprehension, which would have channeled them into the first level, and more skillful readers beginning to adopt a more critical comprehensive stance, pushing them into the fourth level. Unfortunately, individual reading scores were not available for use in the data analysis. Future studies, like that of De La Paz and her colleagues (2005, 2010, 2012) could do more to consider the role of reading fluency in students' ability to write with sources.

4.4 Conclusions

In spite of these limitations, this study illuminates the process of fostering the writing component of historical literacies. With the admission that any classification system oversimplifies students' complex cognitive world, perhaps the spectrum could help teachers as they observe flaws in students' writing and diagnose the underlying misconceptions or missing conceptions that contribute to these errors. For example, it is likely that students in the first three levels fail to understand the nature of history and historical texts. Perhaps explicit instruction on the work of historians, examples of conflicting interpretations of the same event, or other instructional activities designed to promote a more sophisticated epistemic stance might help students engage in a deeper analysis of evidence and improved argumentative writing (VanSledright, 2011).

History instruction is undergoing a shift in the United States that is bringing it into greater alignment with global trends. For decades, European and Canadian educators have promoted a balance of history content instruction and historical process instruction (Eliasson, et al., 2015; Körber, 2011; Lee & Ashby, 2000; Seixas, Gibson, & Ercikan, 2015; Seixas & Morton, 2012; Van Drie & Van Boxtel, 2008; Waldis, et al.,

2015) In an era when the C3 Framework (NCSS 2013) and the Common Core State Standards (CCSS, 2010) encourage the integration of historical reading and writing instruction into the history curriculum of the United States, and when innovative assessments are making it easier to assess historical thinking (Ercikan & Seixas, 2015; Smith & Breakstone, 2012; VanSledright, 2014), it makes sense that more attention be paid to the links between historical reading, thinking, and writing. Further, as increasing attention is focused on nurturing historical consciousness, greater effort must be made to help students use historical evidence to orient themselves within a historical context (Eliasson, et al., 2015; Körber, 2011). This study extends the important discussions on “writing like a historian” and assessing historical thinking, vital elements of modern history instruction around the world.

References

- Alexander, P. A. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Researcher*, 32(8), 10-14. doi: 10.3102/0013189X032008010
- Alexander, P. A., & Jetton, T. L. (2000). Learning from text: A multidimensional and developmental perspective. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research*, (Vol. 3, pp. 285-310). Mahwah, NJ: Erlbaum.
- Ashby, R. A., Lee, P. J., & Shemilt, D. (2005). Putting principles into practice: Teaching and planning. In M. S. Donovan & J. D. Bransford (Eds.), *How students learn: History, mathematics, and science in the classroom* (pp. 79-178). Washington, DC: National Academies Press, 2005.
- Berninger, V. W., & Swanson, H. L. (1994). Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writing. In J. S. Carlson & E. C. Butterfield (Eds.), *Advances in cognition and educational practice, Vol.2: Children's writing: Toward a process theory of the development of skilled writing* (pp. 57-81). Greenwich, CN: JAI Press.
- Britt, M. A., Perfetti, C. A., Van Dyke, J., & Gabrys, G. (2000). The Sourcer's Apprentice: A tool for document-supported history instruction. In P. Stearns, P. Siexas, & S. Wineburg (Eds.), *Knowing, teaching, and learning in history: National and international perspectives* (pp. 437-470). New York: NYU Press.
- Common Core State Standards (CCSS). (2010). Common core state standards. Retrieved February 19, 2011, from <http://www.corestandards.org/>
- De La Paz, S. (2005). Effects of historical reasoning instruction and writing strategy mastery in culturally and academically diverse middle school classrooms. *Journal of Educational Psychology*, 97, 139-156. doi: 10.1037/0022-0663.97.2.139
- De La Paz, S., & Felton, M. K. (2010). Reading and writing from multiple source documents in history: Effects of strategy instruction with low to average high school writers. *Contemporary Educational Psychology*, 35(3), 174-192. doi: 10.1016/j.cedpsych.2010.03.001
- De La Paz, S., Ferretti, R., Wissinger, D., Yee, L., & MacArthur, C. (2012). Adolescents' disciplinary use of evidence, argumentative strategies, and organizational structure in writing about historical controversies. *Written Communication*, 29(4), 412-454. doi: 10.1177/0741088312461591
- Draper, R. J., Broomhead, P., Jensen, A. P., Nokes, J. D., & Siebert, D. (Eds.). (2010). *(Re)imagining content-area literacy instruction*. New York: Teachers College Press.
- Duke, N., Coughlan, S., Juzwik, M., & Martin, N. (2012). Teaching genre with purpose. *Educational Leadership*, 69(6), 34-39.
- Eliasson, P., Alvéén, F., Yngvéus C. A., & Rosenlund, D. (2015). Historical consciousness and historical thinking reflected in large-scale assessment in Sweden. In K. Ercikan & P. Seixas (Eds.) *New directions in assessing historical thinking* (pp. 171-182). New York: Routledge.

- Ercikan, K., & Seixas, P. (Eds). (2015). *New Directions in Assessing Historical Thinking*. New York: Routledge.
- Graff, G., Birkenstein, C., & Durst, R. (2015). *They say, I say*. New York: Norton.
- Grafton, A. (1997). *The footnote: A curious history*. Cambridge, MA: Harvard University Press.
- Halldén, O. (2001). On reasoning in history. In J. F. Voss & M. Carretero (Eds.) *Learning and reasoning in history: International review of history education* (pp. 272-280). New York: Routledge.
- Klein, P. D. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review*, 11, 203-269. doi: 10.1023/A:1021913217147
- Körber, A. (2011). German history didactics: From historical consciousness to historical competencies—and beyond? In H. Bjerg, C. Lenz, & E. Thorstensen (Eds.). *Historicizing the uses of the past: Scandinavian perspectives on history culture, historical consciousness and didactics of history related to World War II* (pp. 145-164). Bielefeld: Verlag. doi: 10.14361/transcript.9783839413258.145
- Lee, P. J. (2005). Putting principles into practice: Understanding history." In M. S. Donovan & J. D. Bransford (Eds.), *How students learn: History, mathematics, and science in the classroom* (pp. 31-77). Washington, DC: National Academies Press.
- Lee, P. & Ashby, R. (2000). Progression in historical understanding among students ages 7-14. In P. Stearns, P. Siexas, & S. Wineburg (Eds.), *Knowing, teaching, and learning in history: National and international perspectives* (pp. 199-222). New York: NYU Press.
- Lee, P. L., Ashby, R., & Dickinson, A. (1993). Progression in children's ideas about history: Project CHATA Concepts of History and Teaching Approaches: 7-14. Paper presented at the Annual Conference of the British Educational Research Association. Liverpool, September 11, 1993.
- Lesh, B. A. (2011). *Why won't you just tell us the answer? Teaching historical thinking in grades 7-12*. Portland, ME: Stenhouse.
- Limón, M., & Carretero, M. (1998). Evidence evaluation and reasoning abilities in the domain of history: An empirical study. In J. F. Voss & M. Carretero (Eds.) *Learning and reasoning in history: International review of history education* (pp. 252-271). New York: Routledge.
- Martínez, I., Mateos, M. D. M., Martín, E., & Rijlaarsdam, G. (2015). Learning history by composing synthesis texts: effects of an instructional programme on learning, reading and writing processes, and text quality. *Journal of Writing Research*, 7(2), 275-302. doi: 10.17239/jowr-2015.07.02.03
- Mateos, M., Solé, I., Martín, E., Cuevas, I., Miras, M., & Castells, N. (2014). Writing a synthesis from multiple sources as a learning activity. In P. Klein, P. Boscolo, L. Kirkpatrick, & C. Gelati (Eds.), *Writing as a Learning Activity* (pp. 169-190). Leiden, Netherlands: Brill. doi: 10.1163/9789004265011_009
- Moje, E. B. (2008). Foregrounding the disciplines in secondary literacy teaching and learning: A call for change. *Journal of Adolescent and Adult Literacy*, 52, 96-107. doi: 10.1598/JAAL.52.2.1
- Moje, E. B., Stockdill, D., Kim, K., & Kim, H. J. (2011). The role of text in disciplinary learning. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerbach (Eds.), *Handbook of reading research*, Vol 4, pp. 453-486). New York: Routledge.
- Monte-Sano, C. (2008). Qualities of historical writing instruction: A comparative case study of two teachers' practices. *American Educational Research Journal*, 45, 1045-1079. doi: 10.3102/0002831208319733
- Monte-Sano, C. (2010). Disciplinary literacy in history: An exploration of the historical nature of adolescents' writing. *Journal of Learning Sciences*, 19, 4, 539-568. doi: 10.1080/10508406.2010.481014
- Monte-Sano, C., De La Paz, S., & Felton, M. (2014). *Reading, thinking, and writing about history: Teaching argument writing to diverse learners in the common core classroom, grades 6-12*. New York: Teachers College Press.

- National Council for the Social Studies (NCSS) (2013). *The college, career, and civic life (C3) framework for social studies state standards: Guidance for enhancing the rigor of K-12 civics, economics, geography, and history*. Silver Spring, MD: NCSS.
- Nokes, J. D. (2013). *Building students' historical literacies: Learning to read and reason with historical texts and evidence*. New York: Routledge.
- Nokes, J. D. (2014). Elementary students' roles and epistemic stances during document-based history lessons. *Theory and Research in Social Education*, 42(3), 375-413. doi: 10.1080/00933104.2014.937546
- Nokes, J. D., & De La Paz, S. (in press). Writing and argumentation in history education. In S. Metzger & L. Harris (Eds.) *The Wiley International Handbook of History Teaching and Learning*. Hoboken, NJ: Wiley.
- Nokes, J. D., Dole, J. A., & Hacker, D. J. (2007). Teaching high school students to use heuristics while reading historical texts. *Journal of Educational Psychology*, 99, 492-504. doi: 10.1037/0022-0663.99.3.492
- Paxton, R. J. (1999). A deafening silence: History textbooks and the students who read them. *Review of Educational Research*, 69, 315-337. doi: 10.3102/00346543069003315
- Penrose, A. W. (1992). To write or not to write: Effects of task and task interpretation on learning through writing. *Written Communication* 9: 189-210. doi: 10.1177/0741088392009004002
- Reisman, A. (2012). Reading like a historian: A document-based history curriculum intervention in urban high schools. *Cognition and Instruction*, 30, 1, 86-112. doi: 10.1080/07370008.2011.634081
- Seixas, P., Gibson, L., & Ercikan, K. (2015). A design process for assessing historical thinking: The case of the one-hour test. In K. Ercikan & P. Seixas (Eds.) *New directions in assessing historical thinking* (pp.102-116). New York: Routledge.
- Seixas, P., & Morton, T. (2013). *The Big Six Historical Thinking Concepts*. Toronto: Nelson.
- Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review*, 78(1), 40-59. doi: 10.17763/haer.78.1.v62444321p602101
- Shemilt, D. (1983). The devil's locomotive. *History and theory*, 22(4), 1-18. doi: 10.2307/2505213
- Smith, M., & Breakstone, J. (2012). New history assessments for the 21st Century. Paper presented at the annual conference of the College and University Faculty Assembly. Seattle, WA: November 15, 2012.
- Stanford History Education Group (2012). Charting the future of teaching the past. Downloaded February 27, 2012. Found at <http://sheg.stanford.edu/>
- Van Drie, J., & Van Boxtel, C. (2008). Historical reasoning: Towards a framework for analyzing students' reasoning about the past. *Educational Psychology Review*, 20, 87-110. doi: 10.1007/s10648-007-9056-1
- VanSledright, B. (2002). *In search of America's past: Learning to read history in elementary school*. New York: Teachers College Press.
- VanSledright, B. (2011). *The challenge of rethinking history education*. New York: Routledge
- VanSledright, B. (2014). *Assessing historical thinking and understanding: Innovative designs for new standards*. New York: Routledge.
- Waldis, M., Hodel, J., Thünemann, H., Zülsdorf-Kersting, M., & Ziegler, B. (2015). Material-based and open-ended writing tasks for assessing narrative competence among students. In K. Ercikan & P. Seixas (Eds.) *New directions in assessing historical thinking* (pp. 117-131). New York: Routledge.
- Wiley, J., & Voss, J. F. (1996). The effects of 'playing historian' on learning in history. *Applied Cognitive Psychology*, 10, 63-72. doi: 10.1002/(SICI)1099-0720(199611)10:7<63::AID-ACP438>3.0.CO;2-5
- Wiley, J., & Voss, J. F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Research*, 91, 301-311. doi: 10.1037/0022-0663.91.2.301

- Wineburg, S. S. (1991). On the reading of historical texts: Notes on the breach between school and academy. *American Educational Research Journal*, 28, 495-519. doi: 10.3102/00028312028003495
- Wineburg, S., Martin, D. & Monte-Sano, C. (2011). *Reading like a historian: Teaching literacy in middle and high school classrooms*. New York: Teacher's College Press.

Appendix A: Assessment Instrument

United States History

Name _____

Period ____

Instructions

You will be given a number of documents that are related to the Battle of Lexington that was fought during the Revolutionary War. You will need to carefully consider the information in the documents. Try to imagine that you are a historian and analyze the documents as a historian would. Use the documents to try to get a good feel for what happened at Lexington. There are three parts to your assignment.

- First, read the documents as a historian would, trying to figure out what actually happened at Lexington. You may take notes on the documents, or highlight or underline or take notes on your own paper to help you figure out what happened.
- Second, in the space below write a 100-word response that answers the bold question below. Be sure to write about the documents. Try to write as a historian would.
- Third, answer the four follow-up questions on the back of this page.

ESSAY QUESTION: In your opinion is the picture that is included with the documents accurate in its portrayal of the Battle of Lexington? Why or why not?

Appendix B: Instructions Used to Identify Sourcing

SOURCING: An individual who uses sourcing looks at the source of a document before reading and keeps the source of the document in mind as he or she reads. The reader's understanding of the document is influenced by the document's source. Sourcing only occurs when the consideration of the source helps the individual make sense of the document. If the student analyzes two documents together, give two marks (For example, if the student was to write "Both Document 2 and Document 7 are biased because the authors of both documents wanted to blame the other side for the event"). When evaluating the students' essays, the following items will be viewed as evidence of the use of sourcing.

- "author's position": 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."
- "evaluation 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.”
- “**evaluation 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...”
- “other”: Any other reference to the source in an effort to understand what the document says or suggest that it is more or less reliable would qualify as sourcing.

NOT SOURCING: Students should **NOT** be given credit for sourcing if:

- they show an awareness of the type of text, but don’t tell why the type of text is important (unless it is obvious, such as in the case of an eye witness account). For example, if they wrote “Document 3 is most reliable because it is a historical novel” but they don’t explain why historical novels are reliable. However, if they give any legitimate reason, 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.”