This article was downloaded by: [North Carolina State University]
On: 25 August 2013, At: 05:40
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Cognition and Instruction
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hcgi20

Reading Like a Historian: A Document-Based History Curriculum Intervention in Urban High Schools

Avishag Reisman [a]
[a] University of California, Los Angeles
Published online: 12 Jan 2012.

To cite this article: Avishag Reisman (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

To link to this article: http://dx.doi.org/10.1080/07370008.2011.634081

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

COGNITION AND INSTRUCTION, *30*(1), 86–112, 2012
Copyright © Taylor & Francis Group, LLC
ISSN: 0737-0008 print / 1532-690X online
DOI: 10.1080/07370008.2011.634081

Routledge
Taylor & Francis Group

Reading Like a Historian: A Document-Based History Curriculum Intervention in Urban High Schools

Avishag Reisman
University of California, Los Angeles

Downloaded by [North Carolina State University] at 05:40 25 August 2013

Enthusiasm about the instructional potential of primary sources dates to the late nineteenth century and has been echoed recently in the work of literacy experts, historians, and educational psychologists. Yet, no extended intervention study has been undertaken to test the effectiveness of primary source instruction in real history classrooms. This study, with 236 11th-grade students in five San Francisco high schools, represented the first extended curriculum intervention in disciplinary reading in an urban district. The Reading Like a Historian (RLH) curriculum constituted a radical departure from traditional textbook-driven instruction by using a new activity structure, the "Document-Based Lesson," in which students used background knowledge to interrogate, and then reconcile, historical accounts from multiple texts. A quasi-experiment control design measured the effects of a 6-month intervention on four dimensions: (a) students' historical thinking; (b) their ability to transfer historical thinking strategies to contemporary issues; (c) their mastery of factual knowledge; and (d) their growth in general reading comprehension. MANCOVA analysis yielded significant main effects for the treatment condition on all four outcome measures. This study has implications for both adolescent literacy instruction and history teaching at the middle- and high-school levels.

Enthusiasm about the instructional potential of primary sources dates to the late nineteenth century, when history first became a standard subject in the school curriculum. In 1899, the American Historical Association's Committee of Seven maintained that sources could show "the nature of the historical process, and at the same time may make the people and events of bygone times more real" (p. 104). Three years later, the New England Teachers' Association insisted that engaging with primary sources was "not as plain as reading; inquisitiveness, patience, imagination, must all be invoked" (Hazen, Bourne, Dean, Farrand, & Hart, 1902, p. 12). Yet, neither this early enthusiasm nor later efforts to design history curriculum around "inquiry"—for example, the New Social Studies movement of the 1960s—translated into enduring curricular reform (Scheurman & Reynolds, 2010). History teachers, then and now, have relied predominantly on recitation and lecture (Cuban, 1993; Goodlad, 1984; Osbourne, 2003; Stodolsky, 1998).

Today, scholars have been advocating again for the inclusion of primary sources. Facing a growing crisis in adolescent literacy, a spate of national reports recommended that literacy instruction in middle and secondary schools be tied to content and involve domain-specific

Correspondence should be addressed to Avishag Reisman, UCLA/CRESST, GSE & IS Building, Room 308, 300 Charles E. Young Drive North, Los Angeles, CA 90095-1522. E-mail: reisman@cse.ucla.edu

Downloaded by [North Carolina State University] at 05:40 25 August 2013

reading skills (Biancarosa & Snow, 2006; Carnegie Council on Advancing Adolescent Literacy, 2010; Heller & Greenleaf, 2007; National Institute for Literacy, 2007). The use of sources has received support from research that suggests students need exposure to multiple genres beyond the fictional texts they encounter in language arts classes (Moje, 2007; Shanahan & Shanahan, 2008) to achieve advanced literacy. Of course, effective instruction involves far more than mere exposure. Indeed, the recently published Common Core State Standards nod to the "particular challenges of reading, writing, speaking, listening, and language" in the content areas (Common Core, 2010, p. 3). These standards beg the question of what constitutes effective instruction with historical texts.

For some time, research on domain-specific historical reading has laid the groundwork for classroom instruction with primary sources (Voss & Wiley, 2006). Wineburg (1991a, 1991b) identified in historians an epistemological orientation toward texts that regards them as human constructions, whose probity can and should be interrogated. He distilled three discrete heuristics that historians applied while reading historical texts: *sourcing* (considering the document's source and purpose), *contextualization* (placing the document in a temporal and spatial context), and *corroboration* (comparing the accounts of multiple sources against each other). Subsequent research on disciplinary "historical thinking" (e.g., Hynd, 1999; Rouet, Britt, Mason, & Perfetti, 1996; Stahl, Hynd, Britton, McNish, & Bosquet, 1996) has found that certain instructional techniques that reveal the structure of the discipline can prompt students to reason historically. For example, researchers found effects for writing prompts that ask for an argument (Wiley & Voss, 1999), for the insertion of an author's voice in an otherwise passive textbook (Paxton, 2002), for the explicit request for sourcing information (Britt & Aglinskas, 2002), and for the juxtaposition of two contrasting arguments about an historical event (Wolfe & Goldman, 2005). These studies contribute to how we might think about history instruction with primary documents; however, their generalizability is limited because most were conducted with proficient readers and none were conducted in real classroom settings.

Two experimental classroom interventions, De La Paz (2005) and Nokes, Dole, and Hacker (2007), took the important step of embedding historical reading instruction in actual history classrooms. De La Paz's study, with 70 eighth graders in 12 lessons, found growth in student persuasive writing; however, effects on students' historical reasoning were mixed and not evident in student writing. Nokes et al.'s intervention, with 200 eleventh-grade students over 15 days of instruction, found that instruction with multiple documents had a significant effect on students' content retention and sourcing. However, student historical reasoning proved difficult to capture. Nokes et al. counted how many times students used certain strategies (i.e., sourcing and corroboration) in their final essays but did not report whether the strategies helped students write plausible and compelling historical arguments or whether these strategies led to deeper understanding of the nature of historical reasoning. By hoping to capture student historical reasoning in *writing*, both studies may have missed opportunities to capture growth in student historical reading.

The present study represented the first extended curriculum intervention in disciplinary reading in an urban district. The study employed a quasi-experimental control design with 236 eleventh-grade students in five public high schools and measured the effects of a 6-month documents-based history curriculum on (a) students' historical thinking, (b) students' ability to transfer their historical thinking strategies to contemporary problems, (c) students' retention of factual knowledge about history, and (d) growth in students' general reading comprehension skills. Like prior experimental interventions, the curriculum focused on teaching students the strategies of

disciplinary historical reading. However, the intervention differed in several ways. First, it took place in five urban high schools that represent a wide range of learning contexts and student populations. Second, this study tested a fully developed document-based curriculum, which consisted of 83 stand-alone lessons addressing the range of historical topics typically covered in an eleventh-grade history curriculum. Third, this was the first study to measure whether instruction in domain-specific reading improved students' disciplinary reading and transferred to growth in their general reading comprehension ability.

CONCEPTUAL FRAMEWORK

A school of research emerged in the 1980s that viewed subject matter as the central lens through which to understand classroom teaching and learning (cf. Shulman & Quinlan, 1996). The literature began with the premise that disciplines were distinct forms of knowledge with particular modes of inquiry (cf. Schwab, 1978; Hirst, 1965; Bruner, 1960) and that experts in each field had normative definitions of domain-specific knowledge and understanding (e.g., Chi, Feltovich, & Glaser, 1981; diSessa, 1985; Schoenfeld, 1985; Wineburg, 1991a, 1991b). From his work with expert historians, Wineburg (1994) concluded that the extant text processing models for generic reading comprehension (e.g., Kintsch & van Dijk, 1978; Kintsch, 1986) did not sufficiently represent the cognitive processes of historical reading. For example, Kintsch (1986) distinguished between a *textbase model*, which the reader builds by parsing sentences, clauses, and ultimately propositions in the text, and a *situation model,* which refers to a mental map of the situation in the text that the reader builds by connecting the content to prior knowledge. Despite its usefulness in describing the cognitive processes guiding the comprehension of solitary texts, Kintsch's model did not account for the intertextual aspect of disciplinary historical reading and the constructed nature of historical accounts.

Wineburg (1994) and Perfetti, Rouet, and Britt (1999) expanded Kintsch's model. Rather than a single situation model, Wineburg proposed that the historian constructs three representations: the representation of the text (rT), which involves the parsing and propositional integration that Kintsch suggests, but also includes an understanding of language as slippery and historically contextualized; the representation of the event (rE), which reflects the text's presentation of the event, including its historical actors and their motivations; and the representation of the subtext (rSB), which allows the reader to make judgments about the author's biases, purposes, and intentions. The larger Event Model grows and shifts as the reader encounters additional documents and modifies his or her synoptic judgment of the event. Perfetti et al. (1999) offered a similar model of intertextual reading that they called the Documents Model, composed of an interconnected Intertext Model and the Situation Model. Their Intertext Model included "intertext predicates" that represented the relationship between the documents (e.g., supports, opposes, agrees with, contradicts, comes before/after) and that demonstrated the reader's efforts to synthesize and reconcile the documents' accounts.

Both Wineburg's and Perfetti et al.'s intertextual models demand that the reader possess a fundamentally different epistemological understanding of history than that which is promoted in typical textbook instruction. Readers who construct such complex cognitive representations view history as open to interpretation and approach historical sources as pieces of evidence that must be interrogated and corroborated. Textbook instruction, by contrast, presents historical

Downloaded by [North Carolina State University] at 05:40 25 August 2013

knowledge as fixed and easily retrievable, so long as students can comprehend the text. The intervention curriculum sought to shift students' orientation toward historical knowledge by explicitly teaching them the strategies of disciplinary reading. Explicit strategy instruction was identified as the instructional mechanism that would redefine the reading process. Rather than extract fixed historical knowledge from the text, students would learn to construct historical knowledge through the process of evaluating and reconciling competing truth claims about the past.

Explicit strategy instruction maintains that for novices to be "apprenticed" into expert practices of domain-specific reading, they must see the strategies named and explicitly modeled (cf. Collins, Brown, & Newman, 1989; Collins, Brown, & Holum, 1991). Cognitive modeling is the first step in explicit strategy instruction, which has been implemented widely in literacy instruction, and consists of a three-part structure: (1) explicit instruction, including naming and modeling the strategy; (2) guided practice, including group work; and (3) independent practice (cf. Duke & Pearson, 2002; Harris & Graham, 1996; Nokes & Dole, 2004; Pearson & Gallagher, 1983). The model emphasizes a gradual shift of cognitive responsibility from teacher to students.

The current study departed from previous implementations of explicit strategy instruction by modeling the disciplinary strategies of *historical* reading, rather than generic reading comprehension strategies. In addition to *sourcing, contextualization,* and *corroboration,* the intervention curriculum trained students in the strategy of *close reading*—carefully considering an author's use of language and word choice. Although close reading has not appeared in the research on historical thinking—nor is it unique to historical reading—it has been identified as a key strategy that readers apply when assessing an author's probity and constructing an intertextual Event Model (Martin & Wineburg, 2008). Together, these domain-specific reading strategies represent cognitive tools that allow the reader to interpret historical texts.

METHOD

Curriculum

The Reading Like a Historian (RLH) curriculum consisted of 83 lesson plans on topics ranging from the early settlement of the New World to the Vietnam War. Topically, the lessons were designed to correspond with the eleventh-grade U.S. survey as it is covered in many states. Epistemologically, however, the lessons inverted the typical relationship between students and historical knowledge and between students and texts. Whereas in traditional history classrooms students are expected to accept and memorize an established historical narrative from a single text (typically, the classroom textbook), in RLH lessons, students were expected to use background knowledge to interrogate, and then reconcile, the historical accounts in multiple texts.

At the core of the RLH approach was the invention of a new "activity structure" (cf. Gump, 1967; Stodolsky, 1998), which I refer to as the "Document-Based Lesson" (Reisman, 2012). The Document-Based Lesson consisted of four distinct lesson segments: (1) Background knowledge; (2) Central historical question; (3) Historical documents; and (4) Discussion. Students first reviewed relevant historical background information that prepared them to engage with the lesson's documents. Materials providing background knowledge ranged from lecture notes, to PowerPoints, to detailed timelines, to clips from historical documentaries. Second, students

Downloaded by [North Carolina State University] at 05:40 25 August 2013

were presented with a historical question that required documentary investigation. Students read between 2 to 5 primary documents that shed light on the central historical question from several perspectives. All of the lessons included explicit strategy instruction, whether in the form of cognitive modeling, guided practice, or independent practice. Lessons with cognitive modeling included scripts for teachers to follow. Generally, students worked in small groups to answer questions that guided their review and interpretation of the primary documents. Finally, students engaged in whole-class discussion about the central historical question, using evidence from the documents to substantiate their claims. Written lesson plans provided suggested questions for teachers to ask during discussion. Whole-class discussion simulated the knowledge-construction processes of the discipline and provided students with opportunities to defend their claims about the past in the face of peer critique. See Figure 1 for an outline of a Document-Based Lesson.

The Document-Based Lessons directly addressed three enduring classroom realities that had eluded prior curricular interventions focused on historical inquiry: (a) classroom teachers' lack of experience and knowledge about historical inquiry; (b) the lack of classroom-ready materials that clearly distilled the processes of historical inquiry; and (c) the failure of the materials to provide or accommodate a chronological historical narrative (Scheurman & Reynolds, 2010). The RLH curriculum tackled the challenge of teacher knowledge by embedding explicit strategy instruction in the lessons, effectively teaching both students and teachers the disciplinary reading strategies used in historical inquiry. The intervention curriculum also included detailed lesson plans and classroom-ready materials—resources that have been tied to effective instructional reform and that address the enormous time demands placed on public school teachers (cf. Brown, 1996; Rowan & Miller, 2007). In addition to providing relevant primary sources, guiding questions, and graphic organizers, the lessons included historical documents that had been modified, both lexically and syntactically, and made visually and cognitively accessible to students reading below grade level (cf. Wineburg & Martin, 2009). Although originals were available to all students, these adaptations allowed struggling readers to engage, substantively, in the process of historical inquiry. Finally, whereas prior inquiry-based history interventions relied on "post-holing," or preparing in-depth units to be used intermittently at the teacher's discretion, the stand-alone Document-Based Lessons acknowledged teachers' need to cover the chronological narrative required by state curricula. Rather than rehash the classic dilemma between curricular coverage and depth, the RLH curriculum attempted to resolve it: In the Document-Based Lesson, factual historical knowledge *enabled* disciplinary historical inquiry. Table 1 compares the features of the RLH approach with traditional textbook instruction. By inviting students to participate in the process of historical knowledge construction, rather than defer to the textbook's authoritative account, the RLH curriculum represented a radical reorientation in students' relationship to historical knowledge (cf. Paxton 1999; Ramirez, Stearns, & Wineburg, 2007; Ravitch & Finn, 1987; Wineburg, 2007).

Participants

Participants were 236 eleventh graders from five public high schools in San Francisco Unified School District, enrolled in U.S. History. Each school contributed one treatment and one control classroom. Treatment teachers attended a four-day training the previous summer; teachers who

Downloaded by [North Carolina State University] at 05:40 25 August 2013

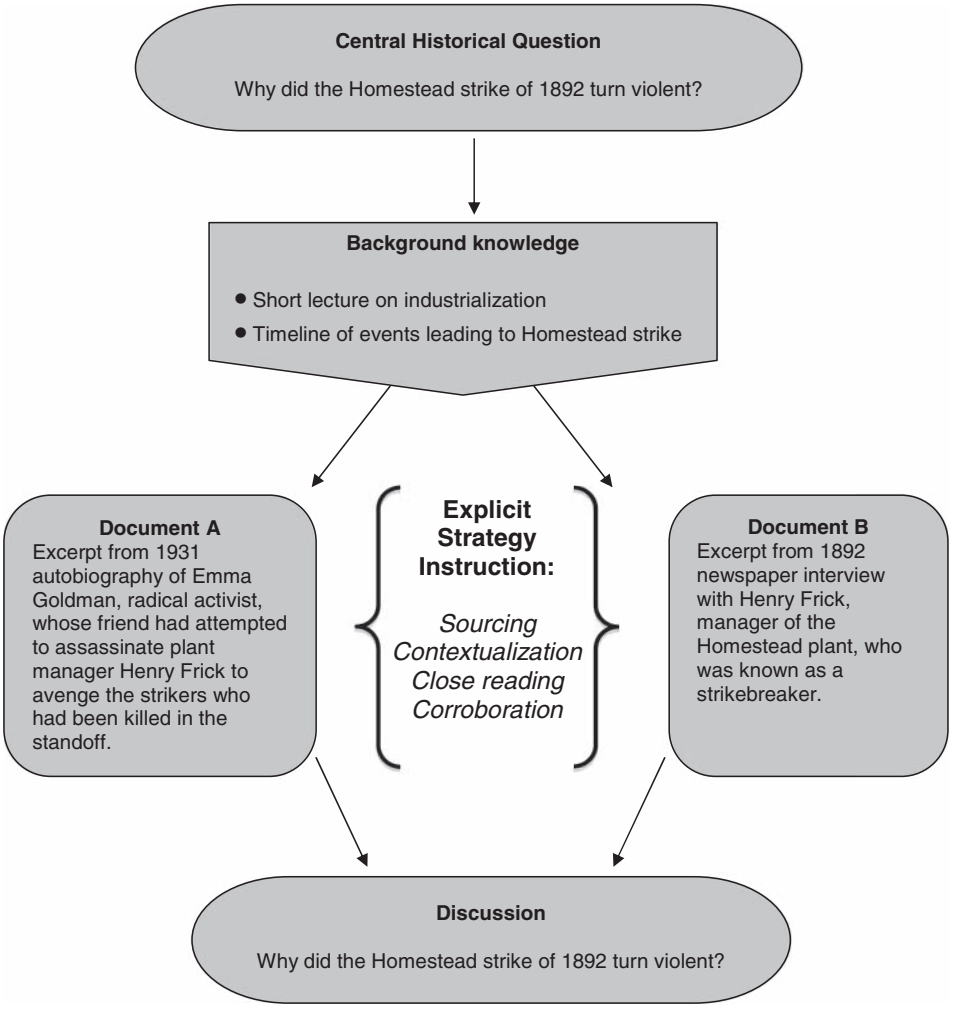

FIGURE 1 Outline of a sample "Document-Based Lesson."

Downloaded by [North Carolina State University] at 05:40 25 August 2013

were interested in the project, but could not attend the summer training, served as controls. We strove for equivalence in teachers' backgrounds in history, as measured by the average number of history classes in undergraduate and graduate careers (treatment $M = 11.4$; control $M = 13.4$). Treatment teachers agreed to use RLH lessons during at least 50% of their instructional time (or 2–3 lessons per week), but they were free to teach any of the lessons in any order. All control teachers were promised the full curriculum and training at the end of the study. None of the treatment teachers had prior familiarity with our materials or approach. Participating schools constituted a representational cross-section of the city's schools. Table 2 lists the demographic characteristics of each school.

TABLE 1
Differences Between Traditional Textbook Instruction and Reading Like a Historian Approach

	Textbook Instruction	*Reading Like a Historian*
Goals	Present students with a chronology of historical events, concepts and people.	Engage students in legitimate historical inquiry.
Role of facts	Students memorize facts in order to perform on standardized tests that emphasize recall.	Background knowledge helps students contextualize and make meaning of primary documents.
Role of texts	Students read a single authoritative text.	Students interrogate historical accounts in multiple primary sources.
Epistemology	Historical knowledge is the accumulation of discrete facts about the past.	Students construct historical knowledge by reading and reconciling accounts in multiple historical documents.

Downloaded by [North Carolina State University] at 05:40 25 August 2013

As is often the case in research in large public school districts (cf. Weinstock, Tseng, Humphrey, Gillespie, & Yee, 2011), we were unable to assign students or teachers to condition randomly. A chi-square test of goodness-of-fit found no significant difference between treatment and control groups in racial/ethnic composition, X^2 (1, $N = 203$) = 6.33, $p = .28$, number of English Language Learners, X^2 (1, $N = 235$) = .19, $p = .67$, and number of students reading below the 25th percentile, X^2 (1, $N = 228$) = .001, $p = .98$. Treatment and control groups did differ by gender X^2 (1, $N = 236$) = 4.98, $p = .03$, with the treatment group having significantly more females. Table 3 compares demographic characteristics of students in treatment and control classrooms.

Pretest Measures

We included three pretests in the design, which we used as covariates:

1. *Factual knowledge*: The tenth-grade history California Standards Test (CST), a standardized end-of-year exam in world history. The 60-item multiple-choice test was scored on a scale of 1–5.
2. *Historical thinking*: We designed and validated a 30-item (22 multiple-choice and 8 constructed response) Historical Thinking Test that we administered both as a pretest and posttest. The items were designed to capture students' historical thinking in general and their ability to practice the historical reading strategies that were taught in the curriculum. The measure included questions about historical documents that were greatly modified so as not to confound student reading comprehension with historical thinking.

   In two pilot administrations with different groups of eleventh graders ($N = 21$ and $N = 13$), the Historical Thinking Test had a reliability of .81 and .91 (Cronbach's α), respectively. The improved reliability on the second pilot reflected the revision of certain items. When we administered the measure to students in the study, we found a reliability of .79 (Cronbach's α) for the pretest administration ($N = 227$) and .82 (Cronbach's α) on the posttest administration ($N = 232$).

Downloaded by [North Carolina State University] at 05:40 25 August 2013

TABLE 2
School Demographics

		Percent Racial/Ethnic Representation (Above 2%)								
School	*Total Student Enrollment*	*Latino*	*White*	*African American*	*Chinese*	*Filipino*	*Other Non-White*	*Per-cent Free/ Re-duced Lunch*	*Percent English Language Learners*	*Percent 11th Graders at or above Proficient on 10th-Grade ELA*
1	939	26.7	—	16	27.6	16.4	11.2	53.2	25.7	29.9
2	2500	16.5	4.8	6.7	49.8	6.6	11	47.5	23.6	42.1
3	666	74	2	11.1	3	4.7	3.9	62.0	44.0	11.5
4	637	21	5.7	18.4	33.8	4.2	12.1	57.4	17.1	35.2
5	2400	9.7	8.9	6.1	51.8	4.1	14.8	44.8	16.1	43.9

TABLE 3
Percent Student Representation in Treatment and Control Classrooms

			Race/Ethnicity						*Reading*	
	Gender									
Percent	*M*	*F*	*Asian*	*Latino*	*African American*	*White*	*Filipino/ Pac. Isl.*	*Native American*	*ELL*	*Struggling Readers*
RLH	44	56	39	27	16	5	12	2	14	27
Control	58	42	49	23	10	9	10	0	12	27

Downloaded by [North Carolina State University] at 05:40 25 August 2013

Each of the eight constructed response questions was scored on a scale of 0 or 1. The rubric developed for each question narrowly defined the answers that could receive a 1. We chose to score conservatively to minimize the potential for Type I error. A second rater scored 10% of the tests, achieving .94 interrater reliability (Cohen's *k*). See Appendix A and B for examples of multiple choice and constructed response items on the Historical Thinking measure.

3. *Reading comprehension*: The Gates-MacGinitie Reading Test (GMRT) (4th ed.), a widely used standardized reading test, was used to measure students' reading comprehension (MacGinitie, MacGinitie, Maria, & Dryer, 2000). We administered the 10th/12th-grade version of the test, using only the comprehension section, which includes 48 multiple-choice questions about 11 distinct reading passages. Two alternative forms (S and T) were used as pre- and posttests.

Posttest Measures

Four posttests were administered at the end of the 6-month intervention. The four measures were as follows:

1. *Historical thinking*: Same measure as pretest.
2. *Reading comprehension*: Parallel form of Gates-MacGinitie Reading Test—Fourth Edition (GMRT).
3. *Transfer of historical thinking*: We designed and validated a 20-item multiple choice transfer measure. Items concerned contemporary topics, such as the presidential campaign, a popular referendum on farming, a controversial movie, and similar issues, but in each item we attempted to embed a discipline-specific reading strategy (cf. Perkins & Salomon, 1989). For example, in an item about global warming, we sought to measure students' attention to sourcing. The measure was piloted with a different group of 11th graders ($N = 31$), with a reliability of .79 (Cronbach's α). When we administered the measure to students in the study ($N = 221$), we found a reliability of .67 for the posttest administration (Cronbach's α).
4. *Factual knowledge:* We used a 30-item multiple-choice test that measured students' retention of factual knowledge. The posttest was comprised of released multiple-choice

items from the New York State Regents Exam in U.S. History and the Grade 11 California Standards Test (CST) in U.S. History. Selected items corresponded to *California History-Social Science Content Standards* (California State Department of Education, 1998) and addressed a range of topics, from events (e.g., Spanish–American War, World War I, Scopes Trial), to social movements (e.g., First Great Awakening, Populist Party, Americanization movement), to constitutional issues and Supreme Court cases (e.g., passage of the 14th Amendment, *Plessy v. Ferguson, Schenck v. United States).* The posttest administration of the Factual Knowledge test ($N = 222$) had a Kuder-Richardson Formula 20 reliability coefficient of .81.

Downloaded by [North Carolina State University] at 05:40 25 August 2013

Teacher Training

Treatment teachers participated in a 4-day training prior to the school year, and two 3-hour follow-up workshops during the 6-month intervention. The summer training introduced teachers to the process of historical inquiry and to the structure of the Document-Based Lesson. Teachers worked through six sample lessons that ranged from comparisons of two opposing viewpoints to lessons with sophisticated central historical questions that required substantial background knowledge and included up to seven primary sources. Practice lessons were followed by debriefing sessions during which teachers anticipated and discussed their students' potential questions and concerns. Considerable emphasis was also placed on explaining and practicing explicit strategy instruction for each of the discipline-specific reading skills. During the summer trainings, teachers received half the lessons in the curriculum and posters to hang in their classrooms listing all the questions a student should ask when practicing each of the historical reading strategies (see column 1 of Appendix C). Follow-up workshops during the school year used video footage of classroom observations to revisit and discuss effective examples of cognitive modeling and whole-class discussion.

Observations

The intervention began in the second week of September and concluded in the first week of March. Excluding holidays and professional development days, the study included 105 days of instruction (approximately 21 weeks). Treatment teachers were observed twice weekly on days they used the RLH curriculum. For the remaining days, teachers reported whether or not they taught a lesson from the curriculum, and if so, which one.

A rubric was developed to gauge teacher fidelity to treatment on a score of 0–3. To earn a base score of 2, teachers needed to implement the lesson as a coherent whole that engaged students in historical inquiry. This meant fulfilling three criteria: (a) indicate to students that history is open to interpretation by clearly posing the central historical question; (b) allow sufficient time for students to engage with at least two documents; and (c) engage in explicit strategy instruction of historical reading strategies either individually, in small groups, or with the teacher modeling, with some degree of teacher feedback (often achieved through recitation). If teachers also successfully facilitated a whole-class discussion about students' interpretations of the documents, they earned a 3. A half score of 2.5 was assigned to teachers who offered summative statements that tied back

to the lesson's central question, or who reserved less than 2 minutes for student discussion of the central question.

A score of 1 reflected a lesson where the teacher attempted to follow the lesson plan but the lesson felt disjointed and did not cohere for one or more of the following reasons: (a) the teacher did not clearly establish the purpose, context, or question guiding the lesson; (b) the teacher did not allow time for students to engage with the documents; or (c) the teacher assigned the documents but did not provide feedback on student reading. A half score of 1.5 meant that the teacher established the purpose of the lesson and students read, answered questions, and received feedback on only one document. Reminding students to "source" the document before reading did not qualify as sufficient feedback to earn a score of 2 if it was not accompanied by further post-reading feedback. A zero meant that the lesson did not draw on RLH materials. A second observer independently rated six lessons and a perfect interrater reliability was achieved on this observation measure.

Each of the five control classrooms was observed four times across 6 months to assess whether the control classrooms served as sufficient contrasts to RLH classrooms. Field notes were taken and all materials from each control lesson were collected. Following each observation, control teachers responded to three questions: (a) Why did you choose the materials you did? (b) What did you expect students would learn? and (c) How typical is this lesson in your teaching?

Downloaded by [North Carolina State University] at 05:40 25 August 2013

RESULTS

Curriculum Fidelity

We observed variability in implementation. Five treatment teachers used RLH materials anywhere from 42–72% ($M = 58.3$, $SD = 11.64$) of their instructional time. We also found variability in teacher pacing. In general, teachers required approximately 1.5 class periods to complete an RLH lesson. However, pacing had as much to do with a teacher's style (i.e., whether unread documents were assigned for homework) and students' reading levels. Teachers' average curriculum fidelity scores approached, but fell below, the baseline score of 2, with the exception of Teacher 2. Table 4 summarizes teachers' fidelity to treatment. These scores suggest that teachers used RLH materials on observation days, that they prompted students to source documents before reading,

TABLE 4
Teacher Fidelity to Treatment

	Percent of Class Periods Devoted to RLH During Intervention	*Total Number of RLH Lessons Taught*	*Average Curriculum Fidelity*
Teacher 1	63.8	42	1.74
Teacher 2	51.4	40	2.17
Teacher 3	71.5	50	1.95
Teacher 4	41.9	36	1.64
Teacher 5	62.9	47	1.83

but they occasionally failed to allow time for students to engage with at least two documents or did not provide any post-reading feedback on student use of the historical thinking strategies. Furthermore, only three of the five teachers (Teachers 2, 3, and 5 in Table 4) ever engaged in the pedagogically complex practice of facilitating whole-class text-based discussion, and these discussions were extremely rare. However, these seemingly low fidelity scores belie the fact that when teachers had students read documents and gave them interpretive guiding questions, they earned a 1 or 1.5. Moreover, teachers referred to and/or modeled the skill of sourcing—which preceded the act of reading—anytime they introduced a document.

The range in teacher fidelity scores generally indicates the degree to which the intervention materials departed from the teacher's normal instruction. For teachers 1 and 4, the learning curve was steepest. Prior to the intervention, teacher 1's classroom instruction consisted almost exclusively of lecturing from an overhead transparency revealing, sentence by sentence, hand-written notes copied from the textbook. Student desks in this class were arranged in rows for the bulk of the intervention, and transitions to group work or discussion constituted a major pedagogical variation. Teacher 4 also had less experience orchestrating group work or whole-class discussion; typical instruction in her class generally involved students reading textbook chapters and answering accompanying questions. By contrast, teachers 2, 3, and 5 all had professional development experiences with curricular packages that promote student interaction and group work (e.g., Teachers' Curriculum Institute). Although the focus of these approaches differed from RLH, the instructional supports (e.g., graphic organizers) and some of the activity structures (e.g., group work) resembled those in the intervention materials. Teacher 2 demonstrated the greatest facility with implementing the intervention materials. She designed PowerPoint lectures replete with images and videos to deliver the lessons' background information. She color-coordinated the laminated historical thinking strategies posters with matching handouts that students were instructed to keep in their notebooks. Of all the teachers, only she extended explicit strategy instruction to include students thinking aloud while reading documents in front of the whole class. In general, teachers 2, 3, and 5 demonstrated greater facility moving between the various activity structures (i.e., lecture, small group work, whole-class discussion) than teachers 1 and 4. Yet, while teachers' growth curves differed, they all moved in same direction. In all cases, curriculum fidelity scores improved over time as teachers became more familiar with the approach.

Despite the variability among teachers, we determined that the participating classrooms demonstrated sufficient fidelity to treatment to be included in the study. In classrooms where the primary mode of instruction, prior to the intervention, was lecture from the textbook, the introduction of a single document—even another textbook account—was revolutionary. It effectively shifted the locus of knowledge away from the textbook to the evidentiary record. Sourcing, in this context, was the tool with which students assessed the reliability and trustworthiness of an account. Even in the absence of explicit feedback on the close reading, contextualization, and corroboration, we considered the practice of sourcing a single document to be a significant departure from traditional history instruction, in that it positioned students as arbiters of others' truth claims.

Each control teacher was observed four times over the course of the study to assess whether their classrooms served as sufficient contrasts to RLH classrooms. In three of the five control classrooms, teachers relied almost exclusively on the classroom textbook for both instruction and assignments. In the remaining two classrooms, instruction and classroom materials varied, and teachers occasionally included primary sources, but these sources were not used in the

Downloaded by [North Carolina State University] at 05:40 25 August 2013

service of historical inquiry. For example, in one instance, primary sources provided students with a particular perspective to use in a role-play, and in another instance they were presented to illustrate a point made in the textbook. At no point were students observed evaluating the trustworthiness of an historical account. Based on the control teachers' responses to questions about their lessons and the materials they used, we concluded that their classrooms provided effective, albeit varied, contrasts to the RLH classrooms.

Test Outcomes

Downloaded by [North Carolina State University] at 05:40 25 August 2013

The study was a 2 × 5 quasi-experiment with a set or vector of outcome variables. The effects of treatment condition and schools were examined with multivariate analysis of covariance (MANCOVA). Because the three pretest measures (Grade 10 History CST, Gates-MacGinitie Pre-Test, and Historical Thinking Pre-Test) were highly correlated, we conducted a principal component analysis. A single component with eigenvalues greater than 1 was extracted, explaining 77.6% of the variance in the measures. We used this composite measure as the covariate, with the four outcome measures as a vector: Historical Thinking Post-Test; Transfer of Historical Thinking Test; Factual Knowledge Test; and Gates-MacGinitie Reading Post-Test. Little's MCAR test was conducted to examine missing student data on any of the pre- or posttests; the result was not significant, indicating that data were missing completely at random, $X^2 = 112.701$ ($df = 101$, $p = .200$). Students with missing data were excluded listwise from the analysis, reducing the total N from 236 to 182 students.

One-way analysis of variance was used to test for differences between treatment and control groups in the reduced data set on the three pretests and on the composite covariate. No significant differences were found between treatment and control groups on the 10th-grade history test (CST), $F(1, 180) = .041$, $p = .84$, on the Historical Thinking Pre-Test, $F(1, 180) = 3.6$, $p = .06$, nor on the Gates-MacGinite Reading Test, $F(1, 180) = .018$, $p = .89$. Nor were significant differences found between treatment and control groups on the composite measure used as the covariate, $F(1,180) = .372$, $p = .54$. Tables 5a and 5b list the means in the full and reduced data sets for treatment and control groups on three pretest measures.

MANCOVA analysis showed a significant overall effect on all outcome measures for both independent variables: treatment, $F(4,168) = 6.889$, $p < .001$, $\eta_p^2 = .141$, and school, $F(16,$

TABLE 5a
Means and Standard Deviations for Treatment and Control on Three Pretest Measures in Full Data Set

	Control		*RLH*	
Measure	*M*	*SD*	*M*	*SD*
Tenth Grade CST	3.20	(1.2)	2.90	(1.3)
	$n = 99$		$n = 125$	
Historical Thinking	13.89	(5.1)	15.03	(5.5)
	$n = 104$		$n = 123$	
Reading Comprehension	25.63	(10.4)	25.00	(9.9)
	$n = 105$		$n = 123$	

TABLE 5b
Means and Standard Deviations for Treatment and Control on Three Pretest Measures in Reduced Data Set ($N = 182$)

	Control ($n = 79$)		*RLH ($n = 103$)*	
Measure	*M*	*SD*	*M*	*SD*
Tenth Grade CST	3.16	(1.2)	3.13	(1.3)
Historical Thinking	14.06	(5.3)	15.59	(5.5)
Reading Comprehension	26.33	(10.1)	26.13	(10)

Downloaded by [North Carolina State University] at 05:40 25 August 2013

684) = 4.565, $p < .001$, $\eta_p^2 = .096$. There was no school X treatment interaction effect, $F(16, 684) = 1.226$, $p = .242$. Follow up univariate ANCOVA analysis found a significant effect for school on three of the outcome measures: Historical Thinking, $F(4,171) = 3.997$, $p = .004$, $\eta_p^2 = .085$, Factual Knowledge, $F(4,171) = 13.15$, $p < .001$, $\eta_p^2 = .235$, and Reading Comprehension, $F(4,171) = 2.65$, $p = .035$, $\eta_p^2 = .058$, but not for Transfer. These findings suggest that school context predicts student achievement, regardless of treatment condition. However, treatment condition was found to have a main effect on all four of the outcome measures: Historical Thinking, $F(1,171) = 17.37$, $p < .001$, $\eta_p^2 = .09$, Transfer of Historical Thinking, $F(1,171) = 14.95$, $p < .001$, $\eta_p^2 = .08$, Factual Knowledge, $F(1,171) = 5.65$, $p = .019$, $\eta_p^2 = .03$, and Reading Comprehension, $F(1,171) = 8.70$, $p = .004$, $\eta_p^2 = .05$. Table 6 lists observed and adjusted means for treatment and control groups on all four outcome measures.

Reading Comprehension

Even though we did not explicitly target general reading strategies, we hypothesized that there might be effects on general comprehension from sustained practice in close reading, as well as from increased "time-on-task" on literacy (cf. Rosenshine & Berliner, 1978) in the RLH curriculum. Indeed, a main effect for treatment was found on students' reading comprehension in the univariate ANCOVA. If we compare observed scores of treatment and control students

TABLE 6
Observed and Adjusted Means for Treatment and Control on Four Outcome Measures ($N = 182$)

	Observed				*Adjusted*			
	Control		*RLH*		*Control*		*RLH*	
Measure	*M*	*SD*	*M*	*SD*	*M*	*SE*	*M*	*SE*
Historical Thinking	16.01	(5.4)	18.63	(5.4)	16.14	(.387)	18.27	(.335)***
Transfer of HT	10.37	(3.2)	12.24	(3.5)	10.52	(9.95)	11.99	(.25)***
Factual Knowledge	17.46	(5.7)	19.09	(5.6)	17.24	(.41)	18.52	(.35)*
Reading Comprehension	29.56	(9.1)	32.22	(8.8)	29.18	(.66)	31.76	(.574)**

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

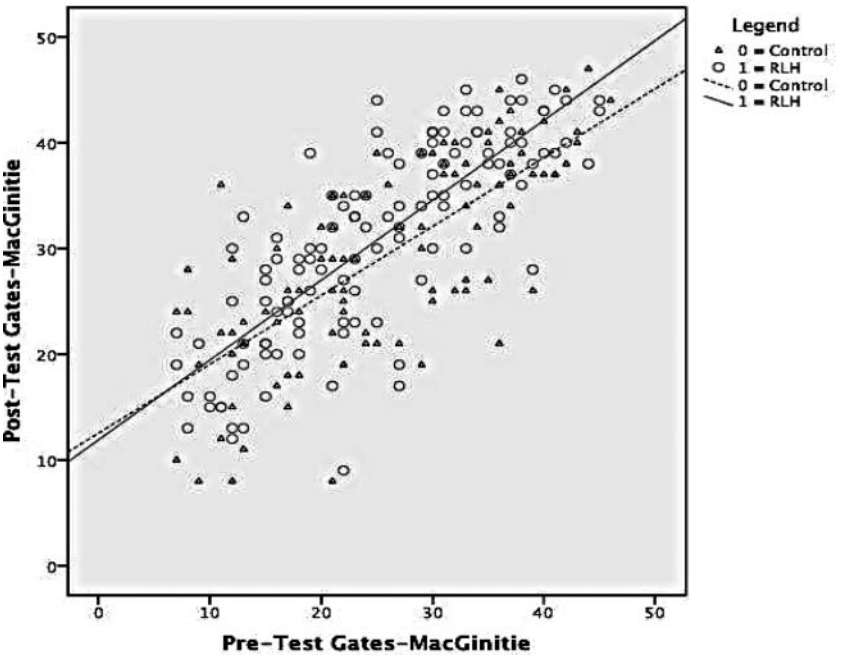

FIGURE 2 Scatterplot of treatment and control scores on pre– and post–Gates MacGinitie Reading Test.

who completed both the pre– and post–Gates MacGinitie test ($N = 218$), we see that treatment students outperformed control on reading comprehension, despite having had lower scores at the outset of the intervention (see Figure 2).

Historical Reading Strategies

Follow-up analysis was conducted on the Historical Thinking and Transfer of Historical Thinking Test to determine whether treatment effects could be traced to any particular historical thinking construct. Items from the two tests (50 items total) were clustered into categories representing each of the four historical reading strategies: sourcing (25 items), contextualization (10 items), close reading (7 items), and corroboration (8 items). Using the principal component as the covariate, we conducted multivariate analysis of covariance with the four categories as a vector. MANCOVA analysis showed a significant overall effect for both independent variables: treatment, $F(4,178) = 6.75, p < .001, \eta_p^2 = .13$, and school, $F(16, 724) = 2.47, p < .001, \eta_p^2 = .05$. There was no school × treatment interaction effect, $F(16, 724) = 0.82, p = .667$. Follow-up univariate ANCOVA analysis found a significant effect for the treatment condition on two of the strategies: sourcing, $F(1, 181) = 15.89, p < .001, \eta_p^2 = .08$, and close reading, $F(1, 181) = 9.62, p = .002, \eta_p^2 = .05$, but not for contextualization or corroboration. Significant effects were found for school on three of the strategies: sourcing, $F(4, 181) = 5.83, p < .001, \eta_p^2 = .114$, contextualization,

Downloaded by [North Carolina State University] at 05:40 25 August 2013

$F(4, 181) = 4.09$, $p = .003$, $\eta_p^2 = .08$, and corroboration, $F(4, 181) = 7.28$, $p < .001$, $\eta_p^2 = .14$, but not for close reading.

Student Demographics

Further analysis was conducted to determine whether the effect for treatment interacted with student characteristics, namely gender, race/ethnicity, English language proficiency, and incoming reading ability. The effects of these variables were analyzed one by one in conjunction with treatment and school, using multivariate analysis of covariance, with the four outcome measures as a vector. Gender was of particular interest because we had found a significantly higher percentage of female students in the treatment condition. However, gender was not found to be significant, $F(4, 158) = 1.072$, $p = .372$, nor was there a significant gender X treatment effect, $F(4, 158) = .071$, $p = .991$. We did not find a significant main effect for student race and/or ethnicity, $F(20, 456) = 1.114$, $p = .331$, nor for the interaction of race/ethnicity X treatment, $F(16, 456) = 1.00$, $p = .454$.

The effect of students' incoming language proficiency was analyzed through two dummy variables: English Language Learner (ELL) designation (RLH $n = 9$; control $n = 12$) and struggling readers, defined as those who scored in the bottom 25th percentile on the Gates-MacGinitie pretest (RLH $n = 25$; control $n = 17$). ELL designation was not found to have a significant main effect, $F(4, 161) = 1.91$, $p = .111$, nor was there a significant effect for ELL X treatment, $F(4, 161) = 1.24$), $p = .294$. Incoming reading ability was found to have a significant overall effect, $F(4,158) = 3.082$, $p = .018$, but the interaction between incoming reading ability and treatment was not significant, $F(4, 158) = 2.315$, $p = .06$. Furthermore, between-subjects univariate analysis showed an effect for incoming reading only on Reading Comprehension, $F(1, 161) = 5.2$, $p = .024$, and not on the other three outcome measures. In other words, students' incoming reading ability predicted their growth in reading comprehension, as one might expect, but had no significant relationship to their performance on the three outcome measures that captured historical thinking, transfer, and factual knowledge.

To explore the possibility that the trend toward an interaction effect between incoming reading ability and treatment suggested that strong readers were responsible for the main treatment effect, we conducted further analysis to determine the effect of treatment on the subgroup of struggling readers. Using MANCOVA on the split data, we found a main effect for treatment for struggling readers, $F(4,28) = 3.07$, $p = .032$, $\eta_p^2 = .31$. Univariate analysis showed that struggling readers in the treatment condition performed significantly better than their counterparts on the Historical Thinking Post-Test, $F(1, 31) = 6.132$, $p = .019$, $\eta_p^2 = .17$, and the Factual Knowledge Test, $F(1,31) = 8.53$, $p = .006$, $\eta_p^2 = .22$, and their adjusted means were higher on Reading Comprehension, but not on Transfer of Historical Thinking (see Table 7).

DISCUSSION

This study was the first to test the effectiveness of a history curriculum focused on disciplinary reading with primary sources. The curriculum included daily lesson plans, with classroom-ready materials and adapted primary sources that recognized and addressed the crisis in adolescent

Downloaded by [North Carolina State University] at 05:40 25 August 2013

TABLE 7
Observed and Adjusted Means of Struggling Readers in Treatment and Control on Four Outcome Measures ($N = 42$)

	Observed				*Adjusted*			
	Control		*RLH*		*Control*		*RLH*	
Measure	*M*	*SD*	*M*	*SD*	*M*	*SE*	*M*	*SE*
Historical Thinking	12.0	(4.3)	13.64	(4.7)	10.41	(1.1)	13.51	(.66)*
Transfer of HT	8.71	(2.6)	9.16	(2.6)	9.41	(.79)	9.10	(.49)
Factual Knowledge	11.65	(4.1)	14.48	(5.3)	11.31	(1.1)	15.03	(.68)**
Reading Comprehension	20.88	(7.5)	21.56	(6.0)	20.52	(2.2)	21.70	(1.4)

*$p < .05$. **$p < .01$.

literacy in urban classrooms (Grigg, Donahue, & Dion, 2007; National Center for Education Statistics, 2009). The results exceeded our expectations—not only did students in treatment classrooms outperform their counterparts on two historical thinking measures that tested students' ability to apply strategies of disciplinary reading, but they also scored significantly higher on measures of factual knowledge and reading comprehension. These findings held true across widely varying school contexts and student demographics. We found no interaction effects for treatment by school, race/ethnicity, gender, English language proficiency, or incoming reading ability. In fact, closer analysis revealed that struggling readers in the treatment condition scored significantly higher on historical thinking and factual knowledge, and comparably higher on reading comprehension. These findings suggest that the RLH curriculum speaks to the urgent call to embed literacy instruction in the content areas to prepare students for the academic tasks that characterize tertiary education (Carnegie Council on Advancing Adolescent Literacy, 2010; Heller & Greenleaf, 2007; National Institute for Literacy, 2007).

Treatment teachers implemented between 36–50 Document-Based Lessons over the course of 6 months of instruction. Yet, it is important to note that teacher fidelity scores were relatively low, hovering below the baseline requirement to (a) present history as open to interpretation by clearly posing the central historical question; (b) allow sufficient time for students to engage with at least two documents; and (c) engage in explicit strategy instruction of historical reading strategies with some degree of teacher feedback. While some teachers came closer than others, none of the teachers implemented the curriculum to its full potential. In a laboratory setting, such low fidelity scores might be grounds to dismiss a study; in an instructional setting, however, poor implementation places the burden of the intervention on the strength of the materials. In this case, the materials were sufficiently flexible to open a corridor of participation for teachers at all points along the pedagogical spectrum. For some teachers, implementation involved sustaining student inquiry about a particular historical question over the course of several activity structures (e.g., individual seatwork, small group discussion, recitation, whole-group discussion). For other teachers, implementation simply consisted of assigning documents as individual seatwork and reviewing student answers to the guiding questions, activities that nonetheless engaged students in disciplinary reading. The findings of a main effect for treatment and the absence of a significant interaction effect for treatment × school suggest that the curriculum did not need to be implemented to its full potential to be beneficial.

Downloaded by [North Carolina State University] at 05:40 25 August 2013

Downloaded by [North Carolina State University] at 05:40 25 August 2013

Certain limitations should be addressed before going further. To begin, neither students nor teachers were randomly assigned to condition, and even though we sought equivalent controls, a quasi-experiment remains less powerful and susceptible to error. Second, a hierarchical linear model (cf. Raudenbush & Bryk, 2002), with students nested in classrooms and schools, would have been preferable but was not possible, given our small sample of teachers. Third, our measures for capturing historical thinking remain inchoate. Although our Historical Thinking findings suggest that students in treatment classes learned to apply disciplinary reading strategies, this measure did not capture students' understanding of the discipline's epistemology, nor any affective or attitudinal shifts that students may have experienced. A closer look at student comments in whole-class discussion would have yielded more information about students' epistemological understanding, but as the fidelity scores indicated, few whole-class discussions were observed in any of the five treatment classrooms.

A fourth limitation concerns the extent to which teaching played a role in the intervention. Curricular materials were necessarily confounded with teacher implementation, which varied widely across schools. Although we might wish to conclude, given the absence of a school × treatment effect, that the materials worked equally well across classrooms, even a casual observer would note that the curriculum was more effective in the hands of skilled teachers who had a firm grasp of subject matter. It is possible that the study's measures were not sufficiently sensitive to detect learning differences between treatment classes. It is equally possible that the materials required a certain threshold of competence that had been met by all teachers in the study. Further research will be needed to examine the role of teaching in implementation and, in particular, whether certain pedagogical techniques (e.g., lecture, recitation, whole-class discussion) were essential to successful implementation.

Finally, the study's design confounded teacher training and teacher observation with the curricular intervention. Treatment teachers attended four days of summer training and two additional workshops, and they were observed twice weekly by a researcher who was available to answer questions about the curriculum. My presence ensured that teachers taught the curriculum on days they were observed, even if they would have preferred to screen a movie or assign seatwork. In practice, I occupied a delicate role as the designer of the intervention materials, and would occasionally be asked by the teacher to clarify certain parts of the lesson plan. On several occasions in the earlier months of the intervention, I stepped in to assist the two teachers who struggled with classroom management. It is possible that my assistance helped clarify the lessons for students; it is equally plausible that I confused students and/or created unnecessary dissonance. We cannot know. What is true is that control classrooms did not experience the regular presence of an outsider over the course of the intervention. The bigger question is whether my presence contributed to the overall effect of the intervention curriculum on student learning. It is possible that learning gains in treatment classrooms were produced by the combined effect of explicit curricular materials and additional coaching. Such a finding would be consistent with research on effective instructional interventions (cf. Correnti & Rowan, 2007), but also calls for further study.

In sum, all the questions raised above call for further study of the intervention materials in an effort to identify the instructional mechanisms that were most responsible for student learning gains. Nevertheless, given that the results were robust across a range of school and classroom contexts, as well as a range of different measures, a discussion of the effectiveness of the materials is warranted. In the following section, I offer possible explanations for student outcomes on each of these measures.

Historical Thinking and Transfer of Historical Thinking

Students in treatment classes outperformed counterparts on two measures of historical thinking. Follow-up analysis revealed that the most robust differences were on items related to *sourcing* and *close reading*. Sourcing questions asked students to consider an author's perspective and purpose, or to evaluate the trustworthiness of a particular account. For example, one multiple choice question asked students why someone might question the trustworthiness of an account written in 1922 about the Battle of Little Bighorn in 1876. Another question asked why someone might trust what a soldier on the battlefield wrote in a letter to his beloved wife. *Close reading* questions asked students to assess a document's argument, word choice, or tone. For example, one close reading question asked why a Mississippi senator would claim to speak for "all mothers and fathers of the South" when writing to President Eisenhower about the integration of Little Rock. Another question, from the Transfer of Historical Thinking measure, asked students to evaluate the tone of a headline from the *New York Post*. Both sourcing and close reading stand in contrast to the passive exercises that constitute reading in textbook-driven history instruction (cf. Schoenbach, Greenleaf, Cziko, & Hurwitz, 1999).

Several possible explanations exist for why there were treatment effects for sourcing and close reading, but not for contextualization or corroboration. Sourcing became habitual in all of the treatment classrooms. In fact, students often referred to the entire approach as "the sourcing thing." Before reading a document, teachers would ask, "What's the first thing we do when we read a document?" to which students would respond in unison, "Source!"[1] No familiar chant introduced the practice of close reading, but the strategy was made visible each time the teacher modeled highlighting, underlining, taking notes, and re-reading challenging sections. The extent to which teachers encouraged or expected students to annotate and highlight documents varied widely between and within classes; however, even in classrooms where teachers did not explicitly require that students do so, such practices were nonetheless modeled. In other words, sourcing and close reading were demonstrated by discrete, concrete actions, such as immediately bringing one's eyes to the source note at the bottom of a page, or underlining vivid language. Moreover, these strategies could be practiced on a single document. Contextualization and corroboration, by contrast, are intertextual strategies that require one to draw connections to prior knowledge or between multiple texts. Such processes are more difficult to model with discrete, concrete behaviors.

These distinctions between sourcing and close reading, on the one hand, and contextualization and corroboration, on the other, are ripe for further exploration. It remains unclear whether contextualization and corroboration are more sophisticated strategies than sourcing and close reading, whether they depend on a deeper epistemological understanding of the discipline, rather than mastery of discrete behaviors. Regardless of the relative complexity of contextualization and corroboration, the results suggest that explicit strategy instruction may not be the appropriate method for teaching these strategies. It is possible that whole-class discussion would have afforded students multiple opportunities to practice and internalize these strategies. Because this intervention ultimately did not include sufficient whole-class discussion, the question of how to teach contextualization and corroboration effectively remains unanswered.

[1]See Reading Like a Historian video: http://www.youtube.com/watch?v=wWz08mVUIt8

Downloaded by [North Carolina State University] at 05:40 25 August 2013

Historical Content

Students in treatment classes scored significantly higher on factual knowledge than their counterparts. These results replicate those of Nokes and colleagues (2007), who found that instruction with multiple documents increased students' content retention. At first glance, this would seem counterintuitive, considering that control classrooms, in general, placed greater emphasis on memorization and factual recall, and RLH classrooms devoted less instructional time to learning conventional historical content (i.e., names and dates). However, the RLH curriculum may have provided meaningful activities and schematic frameworks for students to organize and retain otherwise disparate facts. That treatment students scored high on a measure comprised of multiple choice items from standardized tests suggests that the opportunity to read primary sources and argue about the past may have not only compensated for the "lost" instructional time, but may, in fact, have been more effective in helping students commit facts to memory.

Reading Comprehension

Our conceptual framework suggests that the processes of disciplinary and general reading occur simultaneously, as the reader constructs a "situation model" of a particular historical event. Obviously any act of disciplinary reading engages processes of general reading comprehension. At the same time, even without instruction in disciplinary reading strategies, students in RLH classrooms spent more "time-on-task" (cf. Rosenshine & Berliner, 1978) engaging with print than controls because of the activity structure of the Documents-Based Lesson. Normal social studies classrooms have a wide latitude for how time is spent, including recitation, contests or games, poster-making, role-plays, student reports, films, and so on (cf. Stodolsky, 1998). However, in RLH classrooms, students read daily. On average, each Document-Based Lesson included nearly three primary documents (ranging between 150–250 words). This means that even students exposed to the fewest lessons (36 lessons in Teacher 4's class) read dozens of primary sources. The effect, at the very least, was that students regularly engaged in the act of reading, and it is not surprising that this practice had an effect on general reading comprehension.

IMPLICATIONS

This study unfolded during a unique moment in the history of history education. Since 2001, the federal government has spent nearly $1 billion on an effort "to develop, implement and strengthen programs to teach American history (not social studies) as a separate subject within the school curricula" (Byrd, S.AMDT.3731, 2000). Although the most recent budget sharply cut allocations to the Teaching American History (TAH) program, the past decade saw funding swell to approximately $119 million per year, more than double the original $50 million proposed by Senator Byrd (D–Va). Grants have supported professional development opportunities for teachers that range from summer visits to archives and museums, to lecture series by prominent historians, workshops on designing history instruction using technology such as Google maps, and often all of the above. Most of these professional development programs have used teacher self reports and items from Advanced Placement (AP) tests to gauge increases in teacher knowledge, but few

Downloaded by [North Carolina State University] at 05:40 25 August 2013

studies—not just in history—have tried to link professional development to student achievement (cf. Weinstock et al., 2011; Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). In other words, we do not know whether any of the $1 billion has led to meaningful gains in student achievement (cf. Wineburg, 2009).

The materials and measures in this study offer a different vision of how to effect meaningful instructional reform in history. Yet, they represent only a first step. Additional research is needed to replicate the findings and determine which elements of the treatment were responsible for gains in student reading comprehension and historical thinking. The Historical Thinking measures need to be fine-tuned and expanded to capture changes in students' epistemology. Future studies must examine how to best support history teachers as they implement new materials. Teachers' responses to RLH materials indicate that they are hungry for materials that encourage students to read and think, yet hundreds of thousands of digitized online primary sources remain unusable in classrooms with struggling readers. Rather than sending teachers on field trips to historical archives, this study suggests that classroom instruction might be best supported by classroom-ready materials. Indeed, by providing concrete, classroom-ready materials and lessons focused on explicit instruction of discipline-specific strategies, the study may provide a model that can be leveraged to address literacy demands across the content areas.

Downloaded by [North Carolina State University] at 05:40 25 August 2013

ACKNOWLEDGMENTS

This research is part of the Reading Like a Historian Project, funded by the Stanford K–12 Initiative, Sam Wineburg, Principal Investigator. I was also supported by the National Academy of Education Predoctoral Fellowship in Adolescent Literacy. I thank the students and teachers who so generously gave of their time in the project, as well as members of my dissertation committee—Sam Wineburg, Pam Grossman, and Dan Schwartz—for helpful comments on earlier versions. I am indebted to Rich Shavelson for his astute statistical advice, as well as to Brad Fogo for his ongoing collaboration. This support is gratefully acknowledged; however, I alone am responsible for the contents of this article.

REFERENCES

American Historical Association. (1899). *The study of history in schools: A report to the American Historical Association by the Committee of Seven*. New York, NY: Macmillan.

Biancarosa, C., & Snow, C. E. (2006). *Reading next—A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York* (2nd ed.). Washington, DC: Alliance for Excellent Education.

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, *20*(4), 485–522.

Brown, R. H. (1996). Learning how to learn: The Amherst Project and history education in the schools. *Social Studies*, *87*, 267–273.

Bruner, J. S. (1960). *The process of education*. Cambridge, MA: Harvard University Press.

Byrd Amendment, S. AMDT.3731, H.R.4577, 106th Cong. (2000).

California State Department of Education. (1998). *History-Social Science content standards for California public schools*. Sacramento, CA: California State Department of Education.

Carnegie Council on Advancing Adolescent Literacy. (2010). *Time to act: An agenda for advancing adolescent literacy for college and career success*. New York, NY: Carnegie Corporation of New York.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121–152.

Collins, A., Brown, J., & Newman, S. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing and mathematics. In L. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453–493). Hillsdale, NJ: Erlbaum.

Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator* (Winter), 6–91.

Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects. (2010). National Governors Association Center for Best Practices and Council of Chief State School Officers.

Correnti, R., & Rowan, B. (2007). Opening up the black box: Literacy instruction in schools participating in three comprehensive school reform program. *American Educational Research Journal*, *44*(2), 298–338.

Cuban, L. (1993). *How teachers taught*. New York, NY: Teachers College Press.

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*(2), 139–156.

diSessa, A. A. (1985). Learning about knowing. In E. L. Klein (Ed.), *Children and cognition* (pp. 97–124). San Francisco, CA: Jossey-Bass.

Duke, N. K., & Pearson, D. (2002). Effective practices for developing reading comprehension. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 205–242). Newark, DE: International Reading Association.

Goodlad, J. I. (1984). *A place called school*. New York, NY: McGraw-Hill.

Grigg, W., Donahue, P., & Dion, G. (2007). *The nation's report card: 12th-grade reading and mathematics 2005* (NCES 2007-468). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Gump, P. V. (1967). *The classroom behavior setting: Its nature and relation to student behavior*. Final report, U.S. Office of Education (Project No. 2453). Lawrence, KS: University of Kansas Press.

Harris, K. R., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation* (2nd ed.). Cambridge, MA: Brookline Books.

Hazen, C., Bourne E. G., Dean, S. M., Farrand, M., & Hart, A. B. (1902). *Historical sources in schools: Report to the New England History Teachers' Association by a Select Committee*. New York, NY: Macmillan.

Heller, R., & Greenleaf, C. (2007). *Literacy instruction in the content areas: Getting to the core of middle and high school improvement*. Washington, DC: Alliance for Excellent Education.

Hirst, P. (1965). Liberal education and the nature of knowledge. In R. S. Peters (Ed.), *Philosophy of education* (pp. 113–138). London: Routledge.

Hynd, C. (1999). Teaching students to think critically using multiple texts in history. *Journal of Adolescent and Adult Literacy*, *42*(6), 428–436.

Kintsch, W. (1986). Learning from text. *Cognition and Instruction*, *3*(2), 87–108.

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

MacGinitie, W., MacGinitie, R., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie Reading Test* (4th ed.). Itasca, IL: Riverside Publishing Company.

Martin, D., & Wineburg, S. (2008). Seeing thinking on the web [Electronic version]. *The History Teacher*, *41*(3), 305–320.

Moje, E. B. (2007). Developing socially just subject-matter instruction: A review of the literature on disciplinary literacy teaching. *Review of Research in Education*, *31*, 1–44.

National Center for Education Statistics. (2009). *The nation's report card: Reading 2009* (NCES 2010-458). Institute of Education Sciences, U.S. Department of Education, Washington, DC: U.S. Government Printing Office.

National Institute for Literacy. (2007). *What content-area teachers should know about adolescent literacy*. Washington, DC: National Institute for Literacy.

Nokes, J. D., & Dole, J. A. (2004). Helping adolescent readers through explicit strategy instruction. In T. Jetton & J. A. Dole (Eds.), *Adolescent literacy research and practice* (pp. 162–182). New York, NY: Guilford Press.

Nokes, J. D., Dole, J., & Hacker, D. J. (2007). Teaching high school students to use heuristics while reading historical texts. *Journal of Educational Psychology*, *99*(3), 492–504.

Osborne, K. (2003). Voices from the past. *Canadian Social Studies*, *37*(2). Retrieved from http://www.quasar.ualberta.ca/css/Css_37_2/CLvoices_from_the_past.htm

Downloaded by [North Carolina State University] at 05:40 25 August 2013

Paxton, R. J. (1999). A deafening silence: History textbooks and the students who read them. *Review of Educational Research*, *69*(3), 315–339.

Paxton, R. J. (2002). The influence of author visibility on high school students solving a historical problem. *Cognition and Instruction*, *20*(2), 197–248.

Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, *8*, 317–344.

Perfetti, C. A., Rouet, J., & Britt, M. A. (1999). Toward a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 99–122). Mahwah, NJ: Lawrence Erlbaum Associates.

Perkins, D., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher*, *18*(1), 16–25.

Ramirez, S., Stearns, P. N., & Wineburg, S. (2007). *Human legacy* (10th grade world history textbook). Austin, TX: Holt, Rinehart, & Winston.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Newbury Park, CA: Sage Publications.

Ravitch, D., & Finn, C. (1987). *What do our 17-year-olds know?* New York, NY: Harper & Row.

Reisman, A. (2012). The "document-based lesson": Bringing disciplinary inquiry into high school history classrooms with adolescent struggling readers. *Journal of Curriculum Studies*, *44*, forthcoming. doi: 10.1080/00220272.2011.591436.

Rosenshine, B. V., & Berliner, D. C. (1978). Academic engaged time. *Journal of Education for Teaching*, *4*(1), 3–16.

Rouet, J., Britt, A. M., Mason, R. A., & Perfetti, C. (1996). Using multiple sources of evidence to reason about history. *Journal of Educational Psychology*, *88*(3), 478–493.

Rowan, B., & Miller, K. (2007). Organizational strategies for promoting instructional change: Implementation dynamics in schools working with comprehensive school reform providers. *American Educational Research Journal*, *44*, 252–297.

Scheurman, G., & Reynolds, K. (2010). The "History Problem" in curricular reform: A warning to constructivists from the New Social Studies. In B. Slater Stern (Ed.), *The new social studies: People, projects, and perspectives* (pp. 341–360). Charlotte, NC: Information Age Publishing.

Schoenbach, R., Greenleaf, C., Cziko, C., & Hurwitz, L. (1999). *Reading for understanding*. San Francisco, CA: Jossey-Bass.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. San Diego, CA: Academic Press.

Schwab, J. J. (1978). Education and the structure of the disciplines. In I. Westbury & N. J. Wilkof (Eds.), *Science, curriculum, and liberal education* (pp. 229–272). Chicago, IL: University of Chicago Press.

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review*, *78*(1), 40–59.

Shulman, L. S., & Quinlan, K. M. (1996). The comparative psychology of school subjects. In R. C. Calfee & D. C. Berliner (Eds.), *Handbook of educational psychology* (pp. 399–422). New York, NY: Macmillan.

Stahl, S., Hynd, C., Britton, B., McNish, M., & Bosquet, D. (1996). What happens when students read multiple source documents in history? *Reading Research Quarterly*, *31*(4), 430–456.

Stodolsky, S. (1998). *The subject matters: Classroom activity in math and social studies.* Troy, NY: Educator's International Press.

Voss, J. F., & Wiley, J. (2006). Expertise in history. In K. A. Ericsson, N. Charness, P. Feltovich, & R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 569–584). Cambridge, UK: Cambridge University Press.

Weinstock, P., Tseng, F., Humphrey, D., Gillespie, M., & Yee, K. (2011). *Teaching American history evaluation: Final report*. U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service. Washington, D.C.

Wiley, J., & Voss, J. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology*, *91*(2), 301–311.

Wineburg, S. S. (1991a). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, *83*(1), 73–87.

Wineburg, S. (1991b). On the reading of historical texts: Notes on the breach between school and academy. *American Educational Research Journal*, *28*(3), 495–519.

Wineburg, S. (1994). The cognitive representation of historical texts. In G. Leinhardt, I. Beck, & C. Stainton (Eds.), *Teaching and learning in history* (pp. 85–135). Hillsdale, NJ: Erlbaum.

Wineburg, S. (2007, June 6). Opening up the textbook. *Education Week*, 36–37.

Wineburg, S. (2009). "Teaching American History" program: A venerable past and a challenging future. In K. Woestman & R. Ragland (Eds.), *Teaching American history state by state* (pp. 2–7). New York, NY: Routledge.

Downloaded by [North Carolina State University] at 05:40 25 August 2013

Wineburg, S., & Martin, D. (2009). Tampering with history: Adapting primary sources for struggling readers. *Social Education*, *73*(5), 212–216.

Wolfe, M., & Goldman, S. (2005). Relations between adolescents' text processing and reasoning. *Cognition and Instruction*, *23*(4), 467–502.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). Reviewing the evidence on how teacher professional development affects student achievement (Issues & Answers Report, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from http://ies.ed.gov/ncee/edlabs

Downloaded by [North Carolina State University] at 05:40 25 August 2013

APPENDIX A

Multiple Choice Items on Historical Thinking Test

	Document Prompt	*Item*
Sourcing	*Kate Bighead, a Cheyenne Indian, told the following story to Dr. Thomas Marquis in 1922. Dr. Marquis was a doctor and historian of the Battle of Little Bighorn, which occurred in 1876. He interviewed and photographed Cheyenne Indians in the 1920s.* "Little Big Horn was not the first meeting between the Cheyennes and General Custer. In 1868 Custer attacked our camp, destroying all our food and belongings. Then, Custer promised peace and moved the Cheyenne to a reservation. When gold was discovered, white people came and the Indians were moved again. My brothers and I joined Sitting Bull and the Sioux. As conditions on the reservations became worse more and more Indians moved west, joining our group. Six tribes lived peacefully for several months, hunting buffalo, curing the meat for the winter, and tanning buffalo hides. In the early summer, 1876, we set up camp near Little Big Horn River."	If you wanted to question this document's trustworthiness, you could say: a. Kate Bighead doesn't know what she's talking about. b. Kate Bighead can only speak for the women, not the men. c. Kate Bighead is telling this story forty-six years later, and probably doesn't remember all the details. d. Kate Bighead is lying to make the Indians look good.
Contextualization	N/A for this item.	In 1892, President Benjamin Harrison declared that the birthday of Italian explorer Christopher Columbus a national holiday. What information might explain the timing of Harrison's declaration? a. Harrison read books about Columbus's voyages as a young boy. b. Harrison was born in 1833 near Cincinnati. c. Harrison knew that Italian immigrants were among the largest group of immigrants to America in the early 1890s. d. Harrison thought that most Americans hadn't heard of Columbus

	Document Prompt	*Item*
Corroboration	*In September 1957, nine African-American students tried to enroll in a white high school in Little Rock, Arkansas. Arkansas's Governor ordered police to keep the students out. President Eisenhower sent in the U.S. Army to protect the nine African-American students. Please read the following two letters sent to President Eisenhower and answer the questions:* Document A: *Letter from Mississippi Senator to President Eisenhower, October 1, 1957* Mr. President: All mothers and fathers of the South are strongly against using federal troops to integrate the schools. This is true for both white and black parents. The use of soldiers will completely destroy the public schools. The innocent victims will be children of both races. It is our tradition to keep the races separate. We've passed this tradition from mother to daughter, from father to son. Document B: *Letter from the parents of the nine African-American students to President Eisenhower. October 1, 1957.* Mr. President, We, the parents of the nine children enrolled at Little Rock Central High School, want you to know that your actions in protecting our children's rights have strengthened our faith in democracy. We believe that freedom and equality can only be achieved if everyone has the same opportunities. You have shown that you also believe in freedom and equality. We are deeply grateful to you.	If read together, the two letters above provide support for the claim that: a. Eisenhower made a mistake by sending the U.S. Army to Arkansas. b. Eisenhower did a great thing by protecting the nine students. c. Not everyone in the South thought the same way about segregation. d. The use of the military in the schools always leads to violence.
Close Reading	Same as Corroboration prompt	Document A is written by someone who says that he speaks for "all the mothers and fathers of the South." From this statement, we might assume that: a. The Senator has spoken to equal numbers of black and white parents. b. The Senator thinks that his argument will be stronger if he uses this phrase. c. All white parents in the South agree with Mississippi Senator. d. All black parents support using federal troops to integrate the schools.

Downloaded by [North Carolina State University] at 05:40 25 August 2013

Downloaded by [North Carolina State University] at 05:40 25 August 2013

APPENDIX B

Constructed Response Item, Rubric, and Sample Student Answers on Historical Thinking Test

Question

A British soldier wrote the letter below to his wife during World War One. World War One began in August 1914 and lasted until 1919.

If you were a historian and you wanted to know how Bill really felt about the war, what is one reason why you would NOT trust what Bill says in a letter to his wife?

February 18, 1916

Darling,

I can't bear you to be unhappy about me. Think of the *cause,* the cause. It is England, England, England, always and all the time. The individual counts as nothing, the common cause counts as everything. Have faith, my dear, If only you will have faith in the ultimate victory of the good, the true, and the beautiful, you will not be unhappy even if I never return to you. Dear, I am here, and I shall either survive or not survive. In the meantime, I have never been truly happier. It's all one long blaze of glory!

All my love,
Bill

Rubric and Student Answers

Score	*Criteria*	*Sample Answers*
0	• Not believable that he would love England that much OR • No one is that happy in war	Example 1: *"I can't bear you to be unhappy about me . . . I have never been truly happier." It doesn't make sense.* Example 2: *Because war is not a happy thing even for soldiers.*
1	• Wouldn't want her to worry; wants to protect her OR • Military might have censored letters	Example 1: *Bill wants his wife to be happy, so he does not tell her that he is sad and upset. He tries to be optimistic. [sic]* Example 2: *He could've wrote [sic] the letter and been afraid of who might open and read it, so he put unsincere [sic] thoughts in it.*

APPENDIX C

Historical Reading Strategies Chart

Historical Reading Strategies	*Questions*	*Students should be able to . . .*
Sourcing (Before reading document)	• What is the author's point of view? • Why was it written? • When was it written? • Is this source believable? Why or why not?	• Identify author's position on historical event • Identify and evaluate author's purpose in producing document • Predict what author will say BEFORE reading document • Evaluate source's believability/ trustworthiness by considering genre, audience, and author's purpose.
Contextualization	• What else was going on at the time this was written? • What was it like to be alive at this time? • What things were different back then? What things were the same?	• Use context/background information to draw more meaning from document • Infer historical context from document(s) • Recognize that document reflects one moment in changing past • Understand that words must be understood in a larger context
Close Reading	• What claims does the author make? • What evidence does the author use to support those claims? • How does this document make me feel? • What words or phrases does the author use to convince me that he/she is right? • What information does the author *leave out?*	• Identify author's claims about event • Evaluate evidence/reasoning author uses to support claims • Evaluate author's word choice; understand that language is used deliberately
Corroboration	• What do other pieces of evidence say? • Am I finding different versions of the story? Why or why not? • What pieces of evidence are most believable?	• Establish what is true by comparing documents to each other • Recognize disparities between two accounts

Downloaded by [North Carolina State University] at 05:40 25 August 2013