COGNITION AND INSTRUCTION, *23*(4), 467–502

Relations Between Adolescents' Text Processing and Reasoning

Michael B. W. Wolfe
Department of Psychology
Grand Valley State University

Susan R. Goldman
Department of Psychology
University of Illinois at Chicago

This research examines adolescents' learning about a historical issue from multiple information sources. Adolescents read 2 contradictory texts explaining the Fall of Rome and thought out loud after each sentence. After reading, a series of questions probed their understanding and ability to reason with the information. Think-aloud protocols were coded for the type of processing they reflected, as well as the content that was utilized (e.g., prior knowledge, other text information). Paraphrases and elaborations were the most common types of processing activity. Elaborations involved connections to prior knowledge as well as other text information (both within and across texts) and often-represented self-explanations. The complexity of reasoning about the historical event was predicted by think-aloud comments that increased the coherence of the texts: self-explanations that used prior knowledge or previously processed text information as well as surface text connections. Results are discussed in terms of theories of text processing from single and multiple texts, adolescents' competencies when processing them, and implications of the research for providing students with opportunities to learn to think in discipline-based ways.

Much of the work in text processing research has examined processing of information from single texts. Readers employ a range of strategies while processing text information, and these strategies have different consequences for students' ability

Requests for reprints should be sent to Michael B. W. Wolfe, Department of Psychology, 2224 ASH, Grand Valley State University, Allendale, MI 49401. E-mail: wolfem@gvsu.edu; or Susan R. Goldman, Department of Psychology M/C 285, 1007 W. Harrison Street, University of Illinois at Chicago, Chicago, IL 60607. E-mail: sgoldman@uic.edu

to understand and reason with the information they read (Chi, de Leeuw, Chiu, & LaVancher, 1994; Coté & Goldman, 1999; Coté, Goldman, & Saul, 1998). Learning about topics in school more typically takes place over time, however, and involves students obtaining information from multiple texts that all pertain to the same general issue (Bloome et al., 1999). Prior research has shown that, when learning from multiple texts, students rarely integrate information across texts (Foltz, Britt, & Perfetti, 1996; Greene, 1994; VanSledright, 2002a, 2002b; VanSledright & Kelly, 1998). In this study, we examine adolescent students' processing and reasoning from multiple texts when cross-textual connections are highlighted by presenting conflicting accounts of the same historical event, the Fall of Rome. Integration across texts is particularly relevant for developing historical accounts, and thus the work has the potential to inform history instruction.

One goal of the work was to examine in detail the processing strategies adolescents used to make sense of these conflicting accounts of a historical event. Our second goal was to relate students' processing to subsequent efforts to construct their own explanations for the Fall of Rome. The conduct of this study was informed by theories and methods developed to account for how people learn from text.

TEXT PROCESSING, REPRESENTATION, AND REASONING

The consensual view in text-processing research is that readers construct multilayered mental representations of what they read. These layers include the surface text (the specific words, sentences, layout of the text), the meaning of the text itself (*textbase*), and the interpretation or model of the world referred to by the text (mental or *situation model*; cf. Goldman & Rakestraw, 2000; Graesser, Gernsbacher, & Goldman, 1997; Kintsch, 1998; Perfetti, 1989; van Dijk & Kintsch, 1983). The textbase captures the referential and intra- and intersentential relations among the words in the text. The situation model reflects the integration of prior knowledge with the information explicitly "in" the text. Different processing activities lead to the construction of textbase as compared to situation model layers. Readers whose processing of text information tends to stick to the information in the sentence they are currently reading (by paraphrasing the sentence to themselves, for example) may create accurate but "close to the text" representations. These support a reader being able to report back what the text said but tend not to support reasoning with the information or applying what was learned in novel situations (Coté et al., 1998). Paraphrase processing also does not generate connections to prior knowledge that would allow the text information to be integrated across sentences and with other information the reader knows.

In contrast to paraphrasing, readers also process text by elaborating on what is in the text, making connections within and across texts and to prior knowledge. Although many kinds of elaborations and connections fill in "gaps" in the text and contribute to more coherent mental representations, some types are less likely to do so. For example, personal associations to particular sentences add information from prior knowledge, but the information may not contribute to understanding the meaning of the text as a whole or how that particular sentence fits with the larger meaning of the text. For example, in reading a text about ancient Roman civilization, an elaboration that we would categorize as a personal association would be "My Aunt visited Rome last year." This idea does not contribute to greater coherence of a representation about ancient Roman civilization, at least not without further elaboration. We distinguish these from personal associations that *do appear to be* relevant to meaning construction. For example, had the learner also said, "She told me that the ancient Romans used to fight in the Coliseum and brought me a picture of it," the elaboration would be coded as a relevant association. Similarly, connections that focus on surface similarity (e.g., "They used the same word in this sentence as in the sentence back at the beginning") are ambiguous regarding their contribution to a meaning-based representation of the text. Other forms of elaborations are more explicit regarding making connections among individual ideas from different sentences in the text or integrating text information with the reader's prior knowledge, thereby contributing to the formation of a situation model (Coté et al., 1998; Kintsch, 1994; McNamara & Kintsch, 1996). Connections that explain why and how various ideas in the text are related are referred to as "self-explanation" inferences (Chi, 2000). The generation of self-explanations during comprehension results in a representation of the text that allows for good performance on memory tasks as well as better performance on learning tasks compared to paraphrase-based representations of the text information (Chi et al., 1994; Coté & Goldman, 1999; Coté et al., 1998; Graesser, Singer, & Trabasso, 1994; McNamara & Kintsch, 1996). The facilitative effect of self-explanation inferences has been observed in children as young as 8 years of age and in situations where individuals were reading to learn from informational, expository texts (Coté & Goldman, 1999; Coté et al., 1998).

The same active processes that have been shown to be important with single-text processing are important in processing multiple texts. The case of multiple text comprehension also places additional demands on the reader. Readers must not only comprehend the information within a text, but they must make connections to information across texts to develop a general understanding of the situation being described by the texts. Perfetti, Rouet, and Britt (1999) proposed that, when learning from multiple texts, readers need to construct a *documents model*, consisting of two components, a *situations model* and an *intertext model*. The situations model is a representation of the general situation described across multiple texts. To construct a situations model, readers must integrate the content of different doc-

uments into a single understanding of an event or series of events. In history, multiple texts about the same incident will frequently refer to some of the same events as well as events unique to each document. Readers need to fit the two sets of events together into a coherent whole. Vague or often contradictory associations and details can complicate understanding these cross-textual relations. The intertext model reflects information about the source, purpose, potential bias of a document, and sometimes, global relations between texts (e.g., they "agree" or "disagree"). The information about the source is "tagged" to the content of the document in the situations model. One important aspect of considering the source is that information may be evaluated as untrustworthy. If it is, the information could be comprehended, but rejected as part of a reader's representation of the events being described across documents. The reader could then take this into account in subsequent interpretive or reasoning tasks.

PROCESSING AND REASONING FROM MULTIPLE TEXTS IN HISTORY

The study of history is ideal as a domain in which to examine students' processing and reasoning from multiple texts. For one thing, learning about history typically involves gathering information from different sources that all relate to a particular event or time period. Thus, it is an inherently intertextual practice. VanSledright (2002b) described the challenge of history in terms of an investigation for which the evidence is incomplete but from which we must "reconstruct and recreate" (p. 4) the past. In addition, it is an authentic educational task whose cognitive underpinnings can be subjected to the sort of scrutiny afforded by text-processing methods such as think-aloud protocol analysis. Finally, reasoning with historical documents is a practice that is undertaken quite differently by expert historians compared to either adolescents or college students (Perfetti, Rouet, & Britt, 1999; VanSledright & Kelly, 1998; Wineburg, 1991, 1998).

Expert historians employ a set of heuristics that embody a critical perspective on the documents they read. As described by Wineburg (1991, 1994), these heuristics are *corroboration* across sources (Where do the evidence and the interpretations coincide? Where do they differ?), *contextualization* of texts in their temporal and spatial setting, and *sourcing,* or author identification, especially regarding biases and rhetorical intent (Who wrote the account? What was their purpose in writing it? What sets of beliefs and political perspectives did they hold?). The heuristics that historians use when interpreting historical documents imply that they process not only what the text says but also what it means and its interpretation. In doing so, they bring to the document prior knowledge of other documents about the same event and domain knowledge of the context of the document and its author. They also bring to the task criteria for judging the quality of historical evi-

dence and ways to look for consistency both within a document and across documents. These kinds of processing activities result in the creation of a documents model about a historical event (Perfetti, Britt, & Georgi, 1995). As well, they are precisely the sorts of processing activities that text-processing researchers find associated with better performance on tasks that require readers to use information (learning tasks) rather than just repeat it back (memorization tasks; e.g., Chi, 2000; Coté & Goldman, 1999; Kintsch et al., 1993; McNamara & Kintsch, 1996).

Research to date suggests that readers who are not historians rarely construct documents models. For example, although undergraduate college students learn history by reading multiple sources of different types (e.g., primary sources, participants' and historians' accounts, and textbooks), they tend to approach the different texts separately and uncritically (Foltz et al., 1996; Greene 1993, 1994; Rouet, Britt, Mason, & Perfetti, 1996). College students do tend to make use and keep track of the source information associated with specific documents (Britt, Perfetti, Sandak, & Rouet, 1999; Rouet et al., 1996), but they frequently do not differentiate facts related in historical novels from those related in primary source documents such as treaties (Britt & Aglinksas, 2002). Furthermore, college students tend to stay away from the kind of successive modification and restructuring that historians engage in as they iterate through multiple documents (Perfetti et al., 1995).

Data on adolescents' comprehension of information from multiple texts in classroom settings suggests that their level of intertextual processing is typically low. VanSledright and Kelly (1998) made detailed observations of a fifth-grade classroom in which students were working in groups on tasks that involved learning about a historical event. Students gathered information from textbooks, as well as other kinds of books such as historical fiction novels and history books written in a more narrative style than is typical in textbooks. The observational data on the class, as a whole, suggested that students approached the task as one of accumulating as much information as possible, without engaging in any sort of evaluation of the information or effort to integrate the information across sources. Students did not question the content of the texts nor examine the sources of the texts to make judgments about the validity or potential bias of the information. This information gathering approach was quite reasonable given the teacher's instructions to "learn as much about <topic>" as possible.

VanSledright and Kelly (1998) also interviewed a subset of students to collect more detailed information on the use of multiple documents. The interview data indicated that the students recognized that it was reasonable for authors of historical texts to hold differing views on historical events because different people have access to different information, or because authors may possess a bias toward a particular side of a controversy (see also Lee, 2001). The potential for differing views led these students to conclude that gathering information from multiple sources was beneficial because then students would be exposed to differing opin-

ions and could decide which one made the most sense. They leaned toward plausibility as a criterion for resolving contradictions among accounts rather than examining the evidence base for the interpretations. VanSledright and Kelly (1998) concluded that these adolescents possessed the beginnings of the ability to be critical readers of historical documents. One important aspect of these students' reasoning that differed from historians was that they did not discuss the relation between evidence and opinion but relied on their own judgments of plausibility in deciding among contradictory accounts, replicating a similar finding for high school students (Stahl, Hynd, Britton, McNish, & Bosquet, 1996). The adolescents' reasoning patterns may be due to the lack of a distinction between explanation and information gathering (Ashby, Lee, & Dickinson, 1997), ideas about the relation between the past and accounts of the past (Lee, 2001), or both.

The research findings regarding processing of multiple history documents by undergraduates, high school students, and adolescents' should not be surprising given that history and social studies curricula rarely go beyond the presentation, memorization, and recitation of other people's facts (Holt, 1990; Smith & Niemi, 2001; VanSledright, 1995). This tendency is more pronounced the younger the age of the students (Barton, 1996; Cuban, 1991; Seixas, 1999). Well into early adolescence, students learn lists of facts and acquire minimal understanding of important causal principles and relations among various aspects of social systems (Beck & McKeown, 1992, 1994; Levstik & Barton, 1997). VanSledright's (2002a, 2002b) research suggested that changing the texts and the tasks provided to students is necessary but not sufficient for altering how adolescents process historical documents and reason about historical events. During a 4-month teaching experiment, VanSledright attempted to encourage reading like a historian through the tasks and documents that he made available to students in a fifth-grade classroom. The tasks and documents were designed to encourage sourcing, corroboration, and contextualization. Think-aloud document-processing data were collected from eight case-study students in the classroom. At the beginning of his teaching, VanSledright found virtually no evidence of processing across documents. By the end of the 4 months, there were substantial increases in the incidence of "cross-text" vocalizations. These focused on comparisons of the authors of the documents and details of events but not on the construction of an overall situations model (Perfetti et al., 1999). Furthermore, processing within a single text continued to dominate the think-aloud comments.

THIS RESEARCH

Our research builds on VanSledright's (2002a, 2002b) recent work by using think-aloud methodology to understand how adolescents approach the task of processing historical information from multiple documents. In the research reported

here, we were interested in the relation between the strategies used in processing multiple documents about a single historical event and the complexity of subsequent reasoning about that event. To examine processing and subsequent reasoning, we intentionally designed the processing task to support and encourage cross-document connections, especially comparison and contrast of claims and evidence. Our effort to structure the task to support cross-document processing parallels Palincsar and Magnusson's (2001) efforts to create materials (i.e., Lesley's Science notebook) that provided students with developmentally appropriate representations of data and the process of thinking about data in the context of science investigations. Palincsar and Magnusson measured content learning and reasoning when students used the new materials as compared to students who used a traditional text. They found significant advantages for the students using the new materials. We thought it possible that infrequent cross-document processing in prior research might be due to the comprehension demands of a single document, leaving little "room" for students to think across documents. Second, in the case of single text processing, students who engage in explanatory elaborations and connections within single texts enjoy comprehension and learning advantages over those who engage in other forms of processing (e.g., paraphrasing and associating). We wanted to determine whether similar processing–learning relations could be demonstrated in the multiple text case.

To optimize conditions for cross-document processing, we constructed two historical accounts of the Fall of Rome written by two historians (fictional). Some of the content was shared across the texts, and some of it was unique to each text. In addition, the two texts each stated an explanation for the Fall of Rome and provided evidence to support the explanation. Each text directly contradicted the explanation offered in the other text. Thus, the texts have a degree of referential connection to each other that would not normally be found among documents that were collected from more disparate sources. With these texts, the research is designed to address three primary questions:

1. *How do students read and process multiple accounts of an historical event?* This question is addressed with think-aloud data that are designed to make the adolescents' thinking visible during processing.
2. *Having processed such accounts, how do they reason about them?* This question is addressed with a series of postreading probes to assess comprehension of the information, students' evaluations of the argument presented in each document, and their judgments about the relative merit and convincingness of these arguments.
3. *What is the relation between processing and subsequent explanations of the historical event?* This question is addressed through correlational analyses of the complexity of the postreading explanations students provided for the Fall of Rome and their document-processing strategies.

To preview the materials and procedure, students first completed a prior knowledge assessment and then read the two texts, one after the other, on the Fall of Rome. Students thought out loud as they processed each text. After processing the texts, students answered several kinds of questions relating to the two texts, and also read some additional information that did not necessarily fit well with the explanations provided in the texts. The experiment culminated with the participants providing their own explanation for why Rome fell. Reasoning data are based on this explanation task.

METHOD

Participants

Participants were 44 sixth-grade students (25 girls) who were attending five different public schools in an urban school district in the southeast United States. Although taught by five different teachers, the teachers were covering the same district-mandated content on similar schedules. At the same time, the five teachers were participating in professional development designed to encourage an inquiry orientation to social studies. The professional development was ongoing throughout the academic year (once a month) and engaged teachers in designing inquiry projects that required students to gather information from multiple sources. However, due to district mandates regarding curriculum sequencing and timing, teachers implemented rather limited versions of their inquiry units. As a result, all participants had completed a 6- to 8-week unit on the Roman Empire approximately 2 months prior to participating in this study. The unit involved the use of traditional textbook sources, as well as inquiry projects that required the students to use a variety of library resources to address questions about the relation of various facets of Roman society (e.g., geography, economy, government, and arts and architecture).

Materials

Materials consisted of two "historical documents" that were written by the researchers (Appendix A), a map (taken from the student's textbook) of the Roman Empire with an accompanying time line of major events in the rise and fall of the Empire, and a list of facts about the Roman Empire that were "additional" to the information in the other materials (Appendix B).

The historical documents. These were two texts written by the researchers to appear as if they were historical interpretations developed by two different historians, but no information was provided about when the historians had written the accounts. Participants read both texts. Each provided a different perspective on the

Fall of Rome (see Appendix A). In creating the texts, we consulted with a historian about different explanations for the Fall of Rome and how we could reasonably present them in a short text that would be understandable by sixth-grade students of average reading skill. We decided to focus on the Barbarian invasion of 476 A.D. and two different reasons why the Roman Empire could not defend itself against this invasion, resulting in the Fall of the Roman Empire. We selected hedonism as one reason and, on the basis of our pilot data, instantiated this idea in terms of laziness.[1] The other reason was the sheer size of the empire and its implications for defensibility. Thus in the "lazy" text, (fictional) historian Michael Hatcher argued that the Roman Empire could not defend itself because the citizens had become lazy. In the "size" text, (fictional) historian Rebecca Walsh argued that the Roman Empire could not defend itself because the large size of the Empire had caused troops to be spread too thinly to mount an adequate defense.

We intentionally constructed the texts to be structurally isomorphic, to agree on some information, including the Barbarian invasion, but to disagree as to the underlying cause of why the empire could not defend itself. As can be seen in Appendix A, the first four sentences of the texts are structurally and semantically aligned with each other. Each text began with an explicit statement of disagreement with the theory presented in the other text. This was followed by a sentence indicating that the author had collected evidence to support his or her theory, followed by an explicit statement of his or her theory. Each text then developed a causal sequence of events that culminated in the Empire either consisting of lazy citizens who could not effectively defend themselves, or of the Roman armies being too spread out to be able to defend against an attack on Rome. Each text cited a primary source in support of its argument (a painting or a diary). The last two sentences of each text stated that the Barbarians attacked and defeated the Roman Empire and reiterated the specific theory about the cause of the defeat. The texts were of comparable length and readability. The lazy text was 12 sentences long, contained 209 words, and had a Flesch-Kincaid grade level score of 5.1. The size text was 15 sentences long, contained 203 words, and had a Flesch-Kincaid grade level score of 5.5.

Similarity in the structure and content of the two texts served several purposes. First, we wanted to maximize the likelihood that differences in the explanations for the Fall of Rome would be noticed. Pilot testing indicated that, without this isomorphism, students were less likely to attend to the different causal explana-

[1]Three texts were pilot tested with the goal of facilitating comprehension across texts as much as possible. The third text presented the theory that Rome fell because high turnover among the emperors led to weak leadership, and the lazy text originally stated that the citizens were selfish and hedonistic. Participants thought out loud as they comprehended the texts. The researchers noted comprehension problems and tested different ways of expressing the ideas with which participants were having problems. Due to pragmatic constraints on the length of an experimental session, we could only use two of three texts. The two were selected and modified based on the think-aloud pilot data.

tions. Similarly, we included the explicit disagreement statement at the beginning of the text because, without it, several students indicated that the historians did not really disagree. We wanted it to be clear that the historians knew of each other's theory and disagreed with it, so that if participants decided that the two causes were compatible it would be in light of the historians' disagreement. Finally, pilot testing suggested that the isomorphism in structure and similarity of content type made processing the texts somewhat easier, potentially facilitating comparisons and connections across the two texts.

Time line and map. Participants were also provided with a time line and a map that were available during reading of the texts. The time line listed the dates of five events that were significant in the Roman Empire, including the Barbarian defeat of Rome in 476 A.D. The statements listed on the time line were designed to present the participant with a modest amount of background information, but were not specifically designed to be internally coherent, or to have a particular relation to either of the texts. The map showed the area encompassed by the Roman Empire, with different colors indicating the size of the Empire at different points in time. The time line and map were printed on the same sheet of paper. Participants were told that historians generally agreed on the information in the time line and map.

Fact list. A list of five additional facts, actually events in the life of the Roman Empire, was provided to participants after they had read the historical texts (see Appendix B). Participants were told that historians were in agreement about the occurrence and dates of these events. Our historian consultant suggested inclusion of this information. He indicated that often a problem with historical accounts is the data they leave out even though they provide a coherent account of other data. He chose the events to represent information accepted by historians but that might not fit with the arguments made by the historian authors of the lazy and size texts. This additional information represented an attempt to provide students with some content that might challenge the theories in the two historical accounts and stimulate them to engage more deeply with the information in the texts.

In addition to these materials there was a structured interview that we describe in the procedure section.

Procedure

There were seven phases in the experimental session: prior knowledge elicitation, think-aloud instruction and refamiliarization with basic content on the Roman Empire (through the time line and map), reading the historical accounts with accompanying think-aloud, postreading interview about similarities and differences between the historical accounts, reading the fact list that contained additional information about the Roman Empire, generation of questions to ask each histo-

rian, and explaining why Rome fell. Participants' individual sessions ranged in length from 20 to 45 min. All sessions were audio taped for later transcription.

Phase 1. Prior knowledge elicitation. At the outset of the study, each participant was told that we were interested in talking with them about the Roman Empire, and we asked them if they remembered studying it in class earlier in the school year. (Participants were all members of classrooms in which a 6-week unit on the Roman Empire had been completed sometime in the previous 2 months.) We asked them to tell us what they remembered about the Roman Empire, encouraging them to recall anything they could, and occasionally probing to provide more information or elaborate on a statement. When they had no more to say, we specifically reminded them that the Roman Empire fell and asked them to recall whatever they could about this specific aspect of Roman history.

Phase 2. Think-aloud instruction, time line, and map. Participants were told that they would be reading two historians' explanations for why the Romans could not defend themselves against the Barbarian invasion that ended the Roman Empire. To get ready for that they were going to first read some background information. Participants were then instructed about thinking out loud while they read: Participants were told that, after reading each sentence, they should say whatever the sentence made them think about. These thoughts could be whatever came to mind: what the sentence meant to them, if it reminded them of something, if it made them think about something else they had read or already knew, or if they understood or did not understand the sentence. Participants were also told that the experimenter would be reminding them to think out loud after each sentence. The prompt, "What does that sentence make you think about?" was given if the participant made no comment after 5 sec. Participants then read the time line one sentence at a time and practiced thinking out loud. When reading the first two sentences, participants who appeared to have difficulty thinking out loud had the procedure modeled by the experimenter. Some examples of modeling statements were "509 B.C. Wow, that's a long time ago" and "I've never heard of the Punic Wars. I wonder if there will be more information about them." After reading the time line, participants were asked to look at the map and indicate the area shown on the map that represented the Roman Empire when it was at its largest size. This task was designed to make sure participants took some notice of the map and the information contained on it, and to make sure that participants understood the extent of the Roman Empire.

Phase 3. Reading the historical accounts. Participants read the two historical accounts on the Fall of Rome, with 25 participants reading the lazy text first, and 19 the size text. Participants thought out loud after reading each sentence, with prompts given after 5-sec periods of silence. To ensure that participants were only

looking at one sentence at a time, a moving window procedure was used: Each sentence of the texts was printed on a separate line with a space between each line. The experimenter positioned a cover sheet that had a window cut in it over the text and moved it to reveal each of the sentences, such that only one sentence at a time was visible. Participants were told they could go back to any previous sentence by asking the experimenter to move the window.

Phase 4. Postreading interview: Comprehension of historical accounts. Participants answered questions that were designed to probe their understanding of the information in each of the historical accounts, as well as their integration of information across accounts. Participants discussed similarities and differences between the texts, and rated the strength of each individual argument presented. For each question, participants explained their answers. The interview questions are provided in Appendix C, specifically sections I and II.

Phase 5. Postreading interview: Additional facts. Participants were handed the list of additional facts about Rome and told that these were some events in the history of the Roman Empire on which all historians agreed. They were asked to read the information. During the readings, participants could think out loud if they wished to, but they were not specifically instructed to do so. We positioned this task just prior to asking participants to generate questions for the authors of the texts because the additional information might cue participants to events that the authors had not considered in their accounts. We wanted to see if participants would utilize these additional events in generating questions for the authors that challenged their original claims as being incomplete or ignoring agreed-on facts about Rome.

Phase 6. Postreading interview: Questioning the historians. Participants generated questions they would want to ask each of the historians about their arguments. The goal of these questions was to prompt students to think about what information or evidence might have been left out of the historians' texts. This task was modeled after the "Questioning the author" technique (McKeown, Beck, & Sandora 1996; Sandora, Beck, & McKeown, 1999). However, it differed in that the current task was designed to prompt students to construct meaning from text through a process of challenging the claims and evidence presented by the text's author rather than the goals or purpose of the author. Students were prompted with the question, "If the authors of each of the two texts were here right now and you could ask them some questions, what would you ask them about the arguments they made?"

Phase 7. Postreading interview: Explanation of the Fall of Rome. Finally, participants constructed their own explanations for the Fall of the Roman Empire. We asked them the question, "If someone were to ask you why the Roman Empire could not defend itself against the Barbarian invasion, what would you say to that person?"

Coding of the Data

We used fairly elaborate coding and scoring schemes to analyze the think-aloud data and Phase 7 responses. These are described in this section. Scoring and coding of other responses are described in conjunction with the results.

Think-Aloud Comments

The data providing information on how students processed the two historical accounts are contained in the "think-alouds." The transcriptions of the think-alouds were first divided into comments. A comment is the speech burst following the reading of a text sentence and corresponds to the coarse grain size used by Chi et al. (1994) in coding verbal protocols.

Events within comments. Think-aloud comments could reflect one or more different kinds of thoughts about the text. Each comment was further parsed into events following procedures used by Coté et al. (1998) and corresponding to Chi's use of the idea unit (Chi et al., 1994). Events reflect different kinds of processing or ways of thinking about the text. For example, consider two students' think-aloud comments in response to the focal sentence from the Hatcher account: "I disagree that the empire covered too much territory." One student said, "I guess he disagrees with the other person, because she thought that the army was too spread out. So he doesn't think they covered too much territory." This comment contains two events. The first event (first sentence) connects the content of this sentence to the other text, and the second event (second sentence) is a paraphrase of the content of the focal sentence. Connecting and paraphrasing are different kinds of processing, as elaborated later. Another student said, "He's thinking that the territory wasn't too big. He didn't think that it was too much land for the Romans." This comment contains a single event, a paraphrase of the focal sentence. Think-aloud data for all participants were coded by two independent coders, with disagreements resolved through discussion. Interrater agreement on number of events was 90%, and all disagreements were resolved. Across the 27 sentences of both texts, there was an average of 1.20 events per think-aloud comment.

Types of think-aloud events. All think-aloud events were classified into one of five categories reflecting different kinds of processing of, or thinking about, the text: paraphrases, evaluations, comprehension problems, comprehension successes, and elaborations. These categories were adapted from prior research (Coté et al., 1998) to accommodate the multiple text aspect of this task. Two independent raters agreed on the classification of 92% of the think-aloud events. Disagreements were resolved in discussion.[2]

- Paraphrases: These are think-aloud events that repeat the gist of a sentence in the text without adding additional information. Paraphrasing a text sentence creates a statement that is semantically similar to it. A paraphrase preserves the meaning of the original text sentence. Paraphrases, in contrast to elaborations, add no additional content information to the comprehension situation. For example, a paraphrase for the focal sentence, "These places [Egypt, England, and Syria] were very far away from Rome, the center of the empire," is "Rome was the center of the empire, and these places were far away from them."
- Evaluations: These are think-aloud events that convey a positive or negative judgment about some aspect of the text. Evaluations typically address either the content of the text (e.g., "That's a lot of information"), or the author of the text (e.g., "He's not adding anything new, and not showing enough evidence").
- Comprehension problems: These are think-aloud events that indicate a lack of understanding. Participants reported problems at the word, phrase, or sentence level; in relating the focal sentence to another part of the same text, the previous text, the map or time line information; or in understanding the point of the entire text. For the focal sentence used in the paraphrase example, the think-aloud comment, "They [these places] were very far away from Rome – what places?" indicates a comprehension problem.
- Comprehension successes: In this category are participants' think-aloud comments indicating that they understood what they were reading (e.g., "That makes sense").
- Elaborations: These are think-aloud events that add meaning to, or embellish, the information in the focal sentence or in the text as a whole. For the focal sentence used in the paraphrase example, an elaboration is "So they were afraid that something would happen because it was so far away or something." The text does not mention fear so the participant is making an inference about how the soldiers felt, presumably based on knowledge of when people experience fear. Elaborations were of many different kinds. To capture that variation, we coded each elaboration event with respect to two orthogonal dimensions: the source of information and the kind of elaborative processing.

[2]A small number of think-aloud comments (3%) could not be categorized (e.g., "I don't think anything about that").

We distinguished among four sources of information: prior knowledge; information from earlier in the text (same text); information from the previous text read, the map, or the time line (previous text); or information from a previous think-aloud comment. To determine the information source, we looked at what information was mentioned in the think-aloud event and matched it to the information in the same and previous text materials or in the previous think-aloud comments. If there was no match with one of these sources, we attributed the information to the participant's prior knowledge.

The four types of elaborative processing are self-explanations, surface text connections, irrelevant associations, and predictions.

- *Self-explanations* create new knowledge that serves the purpose of improving understanding of the text information. This new knowledge can involve meaningful integration of text content with prior knowledge or generation of a connection between text information that was not explicitly stated in the text. As such, self-explanations build situation model level understanding. Any of the four information sources could be used in self-explanations. All self-explanations were also coded as to the type of relation they conveyed: causal, comparative, associative, evaluative, analogy, example, or alternative. The example provided earlier, "So they were afraid that something would happen because it was so far away or something," is a causal self-explanation. It added a causal connection between the focal sentence from the text ("It was so far away") and an emotional state (being afraid), presumably based on prior knowledge of circumstances that generate fear. Thus, the self-explanation used prior knowledge to make a causal inference.
- *Surface text connections* relate the focal sentence to another sentence on the basis of surface features ("That talked about distance too"); were not explicit about the similarity ("That's like this one"); or combined information from the focal sentence and other text sentences, without creating an explanation. In combining, participants usually made a compound or complex sentence out of two separate text sentences.
- *Irrelevant associations* used prior knowledge that did not deepen or add to the meaning of the text (e.g., "My family visited Rome last year").
- *Predictions* were comments about information that the reader expected. For example, following the focal sentence from the Hatcher text, "I disagree that the empire covered too much territory," a participant said "He's gonna tell his theory."

In Table 1 we illustrate the combined application of the information source and elaborative processing types to four examples of think-aloud comments from the corpus. The first column of Table 1 indicates the focal sentence that elicited the think-aloud comment. The second column provides the think-aloud comment. It is divided into events according to the event parsing process described earlier. These four examples show each of the four information sources. In the third column, we

TABLE 1
Examples of Coding of Think-Aloud Comments

Focal Sentence (Sentence Being Read)	*Think Aloud Comment*	*Annotation and Coding*
1. These places were very far away from Rome, the center of the empire.	So they were afraid that something would happen because it was so far away or something.	Annotation: Participant brings to the text prior knowledge of situations that make people afraid. The assertion is a causal explanation: The soldiers are afraid because they are far away from their homes in Rome. Number of think-aloud events: One. Inference of a causal relation between being far from home and being afraid. Information source: Prior Knowledge. Fear is not mentioned in presented materials or participant's previous think-aloud comments. Processing type: Self-explanation–causal. Distance from home causes fear.
2. Distances made it difficult to put together a large army quickly.	Egypt, Syria, and England were far from Rome so they could not put an army together as quickly as they could have if it was closer to Rome.	Annotation: Information from two different sentences in the same text are connected. The think-aloud event expresses in one sentence the relation that was expressed in two sentences in the text. Number of think-aloud events: One. Connecting two sentences within same text. Information source: Earlier sentence in same text: (These places [Egypt, England, and Syria] were very far away from Rome, the center of the empire). Processing type: Surface text connection between focal sentence and previous sentence from the same text.

3. My theory is that the reason Rome fell is because the empire got too big.	Event 1: Instead of being that the people got too lazy and worried about themselves, she thought that the empire fell because it had too much land. Event 2: and not enough people to support it.	Annotation: Event 1: Compares thesis statements of the two historical accounts (laziness vs. too much land). Event 2: Adds to the too much land argument. Large size is a problem because there were not enough people to populate it. This connection is based on prior knowledge rather than information explicit in the texts. Number of events: Two, as described. Information Source: Event 1: Sentence from previous text (sentence from the other historical account: My theory is that the Roman Empire fell because the Roman people became lazy). Event 2: Prior knowledge. This idea was not stated explicitly in any of the materials. Processing type Event 1: Comparative self-explanation. States why the second historian disagreed with the first. Event 2: Causal self-explanation. Adds an additional reason that large size was a problem.
4. The Romans' laziness allowed the barbarians to defeat the Roman army and ended the power of the Roman Empire.	Since the army was just there instead of fighting and training, they just fell apart.	Annotation: In a previous think-aloud event, participant indicated that the army was "just there." The current event makes a causal connection between falling apart and being "just there," instead of fighting and training. Number of events: One. Causal inference (falling apart and not being ready to fight). Information source: Previous think-aloud comment (The army wasn't able to support the empire because they no longer had the advantage of being good. They were just there). Processing type: Causal self-explanation. Not being prepared to fight caused the army to fall apart.

provide an annotation that captures the gist of our thinking about the think-aloud comment and the specific coding decisions we made based on this analysis. Interrater agreement on the annotation and coding of the elaboration events was 86%.

Explanations for the Fall of Rome

We coded participants' explanations for why the Romans could not defend themselves against the Barbarian invasion along three dimensions.

Number of causes generated. These could be one or both of the causes from the texts (laziness or large size), taken from the additional information that was provided during Phase 5, or generated from prior knowledge (e.g., "The Barbarians were too strong"). Participants received 1 point for each cause they generated. They received an additional point if, during a prior phase of the interview, they had explicitly rejected one of the two causes from the historical accounts. The rationale behind this scoring decision was based on the pragmatics of the interview situation. If participants had previously discussed and rejected a cause there would be no reason for them to repeat this information during Phase 7 of the interview. Yet it could well be part of their thinking about the cause that they did propose and discuss during Phase 7. Thus, participants who discussed both causes during Phase 7 received 2 points, and participants who discussed and rejected one of the causes prior to Phase 7 and then proposed the other cause during Phase 7 received 2 points.

Complexity of the students' reasoning. There was variation in how students talked about causes they listed. Some participants listed a cause or causes and said nothing else about them. Other students provided additional information. We gave 2 points if the additional information explained why they were providing that cause or causes. We gave 1 point if students elaborated on one or more of their stated causes, but without explaining why they had stated it. If the response contained multiple forms of reasoning, participants received points according to the maximum possible given the multiple forms. For example, if a participant listed one cause and elaborated a second, 1 point was awarded for complexity. If a participant stated and elaborated one cause and stated and explained a second cause, the participant received 2 points.

Integration of causes. Many historical events are the result of multiple causes interacting and converging at a particular point in time. Some of the participants' responses integrated multiple causes by explaining the relation between (or among) causes they had stated. These integrated responses received an additional point.

The points on these three dimensions were summed to give a score reflecting the complexity of the student's reasoning about the fall of the Roman Empire. Note that the complexity score had a maximum of 2 and the integration score a maximum of 1. The number of causes had no maximum, although in practice the number of causes ranged from one to three. Ten of the explanations were coded by two raters, and agreement was 90% or better on each of the three dimensions. Differences were resolved through discussion. The remaining explanations were then coded by one of the raters.

Examples illustrate the coding scheme and character of the responses. In the first example, the student provided one reason for the Barbarian defeat (size) and explained that reason (2 points for reasoning) for a total of 3 points:

> I'd say that the empire was so big it was too far spread out. All of the troops were spread out all around it so nobody could come in, but they had so little troops that it really didn't do a good job of protecting them.

In the second example, the student listed two reasons (lazy and size), but there was no elaboration or explanation of either reason, and there was no integration of the reasons, for a total of 2 points:

> Either because they're lazy or because the troops separated because of the empire getting bigger.

In the third example, the student provided two reasons, and explained them by integrating them, saying that laziness would not have been a problem if the empire had not been so big. This response received 5 points.

> I think that those reasons are good reasons why. It could be a combination, too. If they were too lazy and didn't want to go to the very edge of the Roman Empire, that would be a combination of those two. The Empire was too big and they were too lazy. It was a combination. They were too lazy and the empire was too big. The distances between the armies could have been a long time apart, but if they moved quickly like a couple however miles a day, they could get there quicker and it wouldn't be that hard.

RESULTS AND DISCUSSION

A primary goal of this work is to examine in detail the way students process historical accounts of an event and to determine the relation between their processing and subsequent explanations for the event. We first present the data from the prior knowledge assessment. It was unrelated to processing or reasoning and is not dis-

cussed further. We then present the data on processing by examining the think-aloud protocols and we compare the results obtained in this study to prior research on the processing of single texts. Data from the other parts of the study provide additional insights into adolescents' comprehension of the texts and evaluation of the historians' arguments. Finally, we examine the results from the explanation task (Phase 7) and the relation between these explanations and how participants had processed the historical accounts.

Prior Knowledge Assessment

Participants provided an indication of their prior knowledge of the Roman Empire by telling what they knew about the topic during Phase 1. We were interested in variations in processing and reasoning that might be related to the amount of prior knowledge participants possessed. Prior knowledge was scored as the number of correct statements ("prior knowledge clauses") provided about the Roman Empire. One point was given for each correct clause and one-half point was given for mention of a name or city without further information. Some example comments: "They copied their agriculture from the Greeks," "If you were a woman you have very few rights," "They had emperors instead of kings. They control soldiers and send them out to war," "It fell," and "They would watch chariot racing in the Coliseum."

There was 1 participant with a prior knowledge clause score that was 3.69 standard deviations above the mean. This participant reliably altered the calculations of the relation between prior knowledge, processing, and reasoning, and thus was excluded from the analyses reported here. The remaining 43 adolescents stated an average of 4.16 prior knowledge clauses, with a standard deviation of 3.27 and a range of 0 to 14 (16% of the participants received scores of 0). There were no significant correlations between the prior knowledge clauses and any of the processing measures, nor was there a significant relation between prior knowledge clauses and the scores on the explanations for the Fall of Rome.[3] Thus, the prior knowledge measure used here, the number of clauses generated to an open-ended question, did not predict the processing or reasoning participants exhibited in this study. The lack of a relation should not be surprising given that the level of knowledge about the Roman Empire was relatively low.

---

[3]In addition to prior knowledge about the topic, we were interested in the relation between sixth-graders' scores on a standardized reading test and their performance on these historical reasoning tasks. Data from 21 students on the Terra Nova Reading Test could be obtained. Correlations were computed between these standardized scores and the processing of the arguments, as well as the students' explanations of the Fall of Rome. The only significant relations obtained was that students with higher reading scores generated more causal self-explanations during processing of the texts ($r = .58$, $p = .006$). No other processing frequencies, or the scores on the explanations, were related to the standardized reading scores.

Document-Processing Strategies

Table 2 shows the mean frequency of event types per participant, the relative proportions of different types of elaborations, and the processing reflected in elaborations. Elaborations constituted 58% of the events, paraphrase events 22%, and evaluative events 13%. The evaluative events were overwhelmingly about the content, with only 9% being evaluations of the author, for example, "He's not adding anything new, and not showing enough evidence."

Sources of information. Half of the information used in elaborations was prior knowledge, indicating that participants were attempting to relate information in the focal sentence to information they already knew. However, the prior knowledge information was overwhelmingly general world knowledge or personal experiences, not knowledge of the topic. This is not surprising given students' low levels of knowledge of the Roman Empire as measured by the prior knowledge assessment. The other elaborations were distributed about equally across information from the same text (16%), the prior text (18%), or a prior think-aloud comment (17%). What is of significance in this pattern is that 34% of the elaborations involved connections between the focal sentence and information that was directly traceable to other text materials that were processed during the session. This pattern suggests that students were attempting to compare and contrast text content, a conjecture born out by the processing data reported later.

Processing used in elaborations. The third portion of Table 2 shows the kind of processing that characterized the elaborations and addresses the question of how students used information to make sense of the text. Students' processing was divided about equally between associating (45%) and self-explaining (41%). As we indicated earlier, associative processing refers to associations that are not relevant to constructing a coherent understanding of the text (e.g., "My aunt visited Rome last year"). Within the self-explanations, the majority reflected causal relations (62%), with only 21% comparative relations. Surface-level text connections accounted for an additional12% of the elaborations (e.g., "They both talk about the army"). These were equally divided among connections within the same text and connections to the prior text, time line, and map.

How was information used in processing? Students used the different sources of information in different kinds of processing activities. This point is important because certain types of processing are more likely to contribute to the formation of coherent representations of the historical event. For example, although students used prior knowledge in about half of the elaboration events, 70% of these were irrelevant associations, a type of processing that does not contribute to coher-

TABLE 2
Event Type Data and Processing Data for Elaborations

	Mean Event	
	Frequency	*%*
Total event frequency	32.5	
Paraphrase	6.8	22
Evaluation	4.8	13
Comprehension success	0.9	3
Comprehension problem	1.0	3
Elaboration	19.1	58
Information used in elaborations		
Prior knowledge		49
Same text connection		16
Prior text–time line connection		18
Prior comment connection		17
Processing used in elaborations		
Self-explanation	7.8	41
Cause		62[a]
Comparison		21[a]
Other		17[a]
Association		45
Predictions		2
Surface text connections		12

[a]Percentages of self-explanation.

ence. Only 23% of the instances of prior knowledge use were in the context of causal self-explanations, a kind of processing that contributes to coherence.

The other sources of information also showed different distributions across types of processing. When the *previous* text or the time line or map was the source of information in an elaboration (overall $M = 3.36$), students generated self explanations for the connection 68% of the time ($M = 2.27$). Of these self-explanations, 46% were causal connections and 40% were comparative. When the *current* text was the source (overall $M = 3.07$), students generated self-explanations for the connection 59% of the time ($M = 1.82$). Of these self-explanations, 75% were causal and 15% were comparative. Thus, within the current text the reasoning was overwhelmingly causal, probably reflecting an attempt to create coherence among the various ideas within that text. However, when students connected across texts, they reasoned comparatively as well as causally. Causal reasoning may reflect efforts to establish coherence among the ideas across the texts, whereas comparative reasoning may contribute to understanding the similarities and differences between the arguments, especially given the degree of structural alignment between texts.

Thus, the within-text causal self-explanations indicate that students were attempting to create causally coherent accounts for each of the texts. The cross-text causal and comparative self-explanations suggest attempts to develop a coherent

account of the historical event based on the similarities and differences across all the information. That the efforts to understand each text were generally successful is attested to by the results of the postinterview questions (Phase 4) that assessed comprehension of historical accounts.

Comprehension of the Arguments in the Historical Accounts

We verified that participants had comprehended the basic arguments being presented in each of the texts by examining their responses on the similarities and differences questions (see Appendix C). We coded participants' open-ended responses to these questions for whether they mentioned the basic point that one text presented laziness as the cause of the Romans' inability to defend themselves, and the other text presented large size as the cause. In fact, participants understood the basic difference between the texts: 35 out of 44 (80%) participants indicated the difference in causes. A typical comment about the similarity was "They are kind of similar, but really different, because Hatcher thought that it was mostly the Roman people being lazy, and Walsh thought that it was because the Roman Empire was too big." Other comments pointed to global similarities, reflected in statements such as "They both talk about Rome." The evidence–strength ratings indicated that participants thought there was no difference between the arguments—mean ratings: Hatcher = 3.50, Walsh = 3.77, $F(1, 43) = 1.32$, *ns*. In discussing the arguments, many mentioned the extent to which they found the evidence convincing (e.g., "All he's talking about is the people didn't want to do anything and being lazy. That's not really a very good reason"). Others mentioned their impression of the extent to which facts were presented versus opinion (e.g., "I would give her a four because she had a lot of facts in it").

Nine participants did not mention the basic argument contrast when discussing similarities and differences. Of these, 6 indicated that both texts stated that the Romans were too lazy, and the others predominantly discussed details from specific sentences without precisely mentioning laziness or size. In fact, only one text mentioned laziness. Although it might appear that these 9 participants had not understood the two accounts very well, it was also possible for participants to demonstrate comprehension of the basic arguments in response to other questions. Six of the 9 participants who did not mention the basic arguments when discussing similarities and differences did mention the basic arguments either when rating the evidence strength, posing questions to the authors, or in response to the reasoning question. For example, one student, who focused only on laziness when discussing similarity, indicated understanding of the size argument when explaining strength ratings: "Because it went into a lot of things like how the Barbarians could attack them … and how they could conquer it easier because it was so big." Looked at this way, all but three of the students demonstrated basic comprehension of the arguments put forth in the texts.

TABLE 3
Categories of Questions Participants Would Ask Authors About Their Arguments

Category	*%*
Main claim	29
Details of argument	
More information	15
Causal or explanatory information	22
Criteria for including specific information	4
Additional information provided	1
Research strategy or source of information	20
Claims of other author	6
Purpose of doing the work	2
Author's personal life or history	1

Note. *M* questions generated = 2.68.

Additional information and questions for the historians. Another aspect of comprehension is recognizing what a text leaves unanswered. The additional information that we presented to participants was designed to stimulate such thinking and the evaluation of the strength of the arguments in each text. The analysis of the responses to "Questions for the historian," shown in Table 3, indicate that the additional information seemed to have little impact. Across the categories into which we classified participants' responses, the distribution shows that only 1% of the responses made any reference to events in the additional information. The predominant question categories showed that participants concentrated on asking the historians about the main causal claim in their account (29%, e.g., "Why do you think they were lazy?") and the causal–explanatory details of their arguments (22%). An additional 20% of the questions asked for more information about the historians' strategies for finding support for their claims.

Explanations for the Fall of Rome

Table 4 shows the distribution of students' responses for the three different dimensions that contributed to the reasoning score. This distribution resulted in a mean reasoning score of 2.77 and a mean number of causes of 1.48. Across the group as a whole, 76% of the causes were the main arguments (lazy or size), 24% were generated from prior knowledge, and no causes were taken from the additional information. In terms of individuals, 36% of the students mentioned at least one cause that was not one of the main arguments. The mean complexity score was 1.18, indicating that students tended to do more than list cause(s). As Table 4 indicates, only 10 (22%) simply stated the cause or causes. The remainder either elaborated (36%) or explained (41%) how the cause(s) related to the Fall of Rome or how the causes re-

TABLE 4
Distribution of Students' Reasoning About Why Rome Fell

	Stated	*Elaborated*	*Explained*
Single cause	7	9	8
Multiple cause			
Two	3	6	10
Three		1	
Integration			5

Note. Integration is scored only for those students whose responses contained more than one cause and is thus a subset of those who provided multiple causes.

lated to each other or the Fall of Rome. As well, five students brought multiple causes together, explaining how the two combined to create the Fall of Rome. (Please refer to Example 3 in the scoring section.)

Relations Between Processing and Explanations for the Fall of Rome

A primary question of interest for this study is whether processing of the historical accounts was related to the complexity of the explanations that students provided for the Fall of Rome. To address this question, we first correlated processing variables derived from the think-aloud data with the reasoning scores, using the frequency of occurrence for the processing variables.[4] Table 5 presents the four variables that were significantly correlated with the reasoning score, the correlations, and the intercorrelations among these variables. Consistent with previous research using the think-aloud method (e.g., Chi et al., 1994), we found that three of these four variables involved self-explanatory processing—either using prior knowledge, the current text, or the previous text. Thus, the greater the number of causal self-explanations students generated during processing, the higher the reasoning score. The fourth variable reflects processing that juxtaposes two segments of a text being read and notices some type of connection or similarity. All four of these variables should tend to lead to the creation of integrated representations, and thus facilitate more complex explanations of the event. None of the other processing variables were significantly correlated with reasoning.

[4]Variables that were correlated were the reasoning score, frequency of paraphrase, evaluation, comprehension problem, comprehension success, four sources of information (prior knowledge; current text; prior text, time line, and map; previous comment) crossed with types of processing (causal self-explanations, comparative self-explanations, associations [not relevant], and surface text connections). Other variables that might have been included had an insufficient number of nonzero cells for correlations to be meaningful or encompassed variables that were included (e.g., elaborations is a superordinate to the 16 variables resulting from crossing source with processing type).

TABLE 5
Simple Correlations Between Reasoning Score and Processing Categories Coded From Think Aloud Data

Variable	*Current Text in Causal Self-Explanation*	*Previous Text in Causal Self-Explanation*	*Current Text in Surface Connection*	*Reasoning Score*
Prior knowledge in causal self-explanation	.167	.332*	.063	.379*
Current text in causal self-explanation		.536**	.260	.325*
Previous text in causal self-explanation			.199	.382*
Current text in surface connection				.309*

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

We submitted the four variables to a stepwise multiple regression analysis to determine the best model for predicting the reasoning score. The three-variable model shown in Table 6 accounted for 28% of the variance in the reasoning score, $F(3, 40) = 5.05$, $p = .005$. This model includes previous text in causal self-explanation, prior knowledge in self-explanation, and current text surface connection. The fourth variable from Table 5, current text in causal self-explanation, did not account for a significant amount of unique variance above that accounted for by the first three variables (an increase of 1%) and was not included in the model. Given that the first two variables pick up differences among students in self-explanation, it is not surprising that there was little unique variance remaining to be accounted for by the current text causal self-explanations variable.

What was somewhat surprising was that noticing surface connections within the same text did account for a marginally significant ($p = .08$) amount of variance in the model. We interpret these surface connections as an additional mechanism by which content is being connected, although the think-aloud comments do not

TABLE 6
Multiple Regression Predicting Student Reasoning Scores Using the Variables Shown in Table 5

Regression Model (Predictor Variable)	R^2	*F*	*Partial Correlation*	*t*	*p*
Previous text causal self-explanation + Prior knowledge causal self-explanation + Current text surface connection	.28	5.05			.01
(Previous text self-explanation)			.25	1.65	.11
(Prior knowledge self-explanation)			.30	1.99	.05
(Current text surface connection)			.27	1.77	.08

Note. R^2 and *F* values refer to the variance accounted for by the complete regression model. Partial correlation and *t* values refer to the unique variance accounted for by each predictor within the complete model.

permit a precise understanding of what similarity or contrast students are thinking about when they notice and report these connections. However, these think-aloud events do juxtapose sentences that are not located in proximity to one another in the surface text and thus may be a mechanism for creating greater within-text coherence. Parenthetically, it is interesting to note that associative reasoning that used prior knowledge that was largely irrelevant to the specific content of the texts was not predictive of more complex explanations for the historical event. In summary, those students who used relevant prior knowledge and the text in explanatory causal relations to make sense of the historical event were the ones who subsequently provided the more complex explanations for the Fall of Rome event.

DISCUSSION

In this experiment, we were interested in adolescents' processing of information from multiple texts, reasoning about that information, and the relation between processing and reasoning. The texts were parallel accounts of a historical event, the Fall of Rome. We first summarize the main findings and then discuss implications for each of these issues.

Summary of the Major Findings

The processing, comprehension, and explanatory reasoning results provide clear evidence that 11- to 13-year-olds from diverse backgrounds and a range of classrooms could successfully negotiate multiple historical accounts as represented in the texts used in this study. They engaged in many of the same kinds of activities that characterized adolescents' single-text processing (Chi, 2000; Chi et al., 1994; Coté & Goldman, 1999; Coté et al., 1998) and, in addition, related information across the two texts. In particular, students generated causal self-explanation inferences that connected information in the current sentence with prior knowledge, prior information in the text they were reading, and information in the other historian's text, the time line, or the map. All of these types of causal self-explanations were positively correlated with the reasoning score. Finally, the number of surface text connections to information in the text the students were currently reading was positively correlated with reasoning, suggesting that even connections that do not involve the creation of new knowledge can potentially improve students' reasoning.

The pattern of correlations generally indicates that students who made more of an effort to establish connections within and across texts, and explained the connections, produced more complex explanations of the event they were learning about. These are the same characteristics of processing that have previously been found to positively impact various learning measures (e.g., Chi et al., 1994; Coté & Goldman, 1999; Coté et al., 1998). They serve the purpose of creating an inte-

grated mental representation of the situation being described within a text, and across the two texts, and contribute to the formation of a situation model (Coté et al., 1998; Coté & Goldman, 1999; Kintsch, 1998; Perfetti et al., 1995; van Dijk & Kintsch, 1983). Less integrative processing, such as paraphrasing, supports establishing a strong textbase and good memory performance, but this kind of processing is often not sufficient to support use of the information to reason about events (Kintsch, 1994; Kintsch et al., 1993; van Dijk & Kintsch, 1983). It is also noteworthy that associative prior knowledge accounted for about half of the elaborations that students made. Associative prior knowledge, as defined in this context, reflects personal experiences that do not contribute to understanding the text content per se. Although they may have provided the reader with a vague sense of topic familiarity, they did not further understanding of the historical event.

The positive relation between processing and reasoning also suggests that adolescents viewed the reasoning task as one in which they were to discuss what they learned from the two opposing texts, rather than one in which they were to gather information from any available sources, including the additional information. This interpretation of the task contrasts with one adolescents often make of the research process in classroom settings: Collect together as much information as possible. This interpretation is often supported by task instructions in the classroom (Goldman et al., 1999; VanSledright & Kelly, 1998).

We recognize, however, that these results must be interpreted in the context of the complexity—actually the lack thereof—of the historical accounts used in this study. These historical accounts were specifically designed to be comprehensible to students of this age and reading-skill range, and to encourage cross-text processing. Even with that intentional design, there was substantial variability in processing and reasoning among the adolescents, although all but three students understood the basic divergence between the accounts. Thus, though we sacrificed authenticity and limited the generalizability of our results, most of the students were reasoning from the same basic understanding of what each "historian" proposed as the explanation for the event.

Processing of and Reasoning About Historical Arguments

When processing two opposing arguments about the Fall of Rome, adolescents tended to make attempts to establish connections between the arguments. In addition, when adolescents were generating cross-text connections, they frequently explained the connection by generating a self-explanation (rather than merely making superficial associative connections). This evidence indicates that adolescents are indeed capable of processing information about a topic from multiple texts and thinking about relations across the texts. VanSledright and colleagues (VanSledright, 2002a, 2002b; VanSledright & Frankes, 2000; VanSledright & Kelly, 1998) argued that at least some adolescent students are capable of "practic-

ing" history by comprehending information across multiple texts and reasoning about conflicting evidence. The present data are consistent with this conclusion and add several additional dimensions to what we know about adolescents' processing of, and reasoning about, history.

First, this study provides direct evidence of reasoning across texts from think-alouds that occurred during comprehension of the historical arguments. We found that 18% of the elaborations made during processing represented connections to information that was in a text read previously during the experimental task. In previous work on processing from multiple texts, inferences about reasoning across texts have been generated based on data collected after comprehension, such as in student essays or discussions (VanSledright, 2002a, 2002b; VanSledright & Kelly, 1998). Second, the data collected in this study come from five different classes with a large and diverse number of students. Third, most of the adolescents in this study did not participate in special instructional units in which comprehension of information from multiple texts was a focus, as was the case in the work of VanSledright and colleagues. Finally, in this study, we demonstrate a relation between the processing of information from multiple texts about a specific historical event and the complexity of the reasoning students did in subsequently providing their own explanations for that event.

These results on adolescents' processing and reasoning also provide information about the participants' general approach to the task of reasoning about history and their epistemological beliefs. During the course of the experiment, adolescents were offered information from several different sources, including a map and time line, two argumentative texts written by historians, and additional factual information about the Roman Empire. The information they processed, in conjunction with what they did not process, suggests that the adolescents interpreted the task demands in a narrow sense: Read the two texts and learn as much as you can about the event. That is, for the most part, students read the texts and engaged in an effort after meaning. However, they did not generate connections between the texts and the additional information, nor did they include the additional information as potential causes when reasoning about the Fall of Rome. They were also very willing to offer their own explanations of the event on the basis of the information in the texts; they did not discuss additional information they would have liked to have had.

Findings from prior research indicate that adolescents tend to approach the study of history as a process of remembering factual information and judge the usefulness of multiple sources in terms of their ability to gather more information related to the topic of study (Goldman et al., 1999; VanSledright, 2002a, 2002b; VanSledright & Kelly, 1998). This research represents a particular aspect of this approach in that the adolescents focused on the information they viewed as directly relevant to the task at hand but paid little attention to information relevant to the more general topic. This is not surprising for two reasons. First, we did not attempt

to address students' epistemological beliefs about "doing" history. Second, due to the school district's mandates about coverage and sequencing, classroom instruction had included only very limited kinds of inquiry projects. The "inquiry" projects focused on finding factual information relevant to specific aspects of civilization but did not require students to explain relations among the different systems (e.g., economic and government systems) or conjecture about relations between systems (e.g., geography and economy). In contrast, a more expert approach to studying a historical topic would involve attempting to incorporate all available information to construct a general understanding of the topic, as in the situations model of Perfetti et al. (1995).

History Instruction and Adolescents' Thinking

These findings also suggest two important points related to history instruction and students' approach to the study of history. One conclusion is that adolescents can, indeed, establish connections among information across multiple texts. A potentially important factor regarding this conclusion is the extent to which the texts are structurally aligned with each other. In this research we intentionally made the connections across the two texts as transparent as possible for purposes of inviting cross-text connections and comparisons. We suggest that the admittedly artificial but transparent structural alignment acts as a scaffold and that cross-text connections and comparisons may be precursors of corroboration processes. In a typical classroom situation, students may read multiple texts relating to the same historical time period or event, but without either sufficient background knowledge or an "expert" view of how to study history. In such cases, students would be unable to generate the types of cross-text connections that lead to deeper understanding of the topic. In this experiment, students needed neither extensive background knowledge nor the specific goal of establishing connections across texts. The texts invited these connections and comparisons through their highly parallel structure and the explicit reference to the "other" account in the very opening of each account. This strategy in text construction is similar to Palincsar and Magnusson's (2001) creation of Lesley's Science notebook, a developmentally appropriate set of artifacts that scientists use in investigations. The difficulties VanSledright (2002b) encountered in figuring out how to encourage historically meaningful cross-document processing suggest that the "engineered text" approach might facilitate or jumpstart this kind of thinking.

Furthermore, the positive relation between cross-text self-explanations and reasoning complexity suggests that these connections are leading to a more integrated understanding of the text, not merely being generated because they are invited by the think-aloud process. As a pedagogical issue, the results suggest the potential benefit of having students "practice" learning about history with a small set of documents constructed in a deliberate manner to encourage "corroborative process-

ing." History teachers could also use such materials as a way of illustrating or modeling these processes of cross-text reasoning.

Another important issue related to the study of history arises from the adolescents' focus on the two parallel texts and lack of consideration of the additional information. This finding suggests that teaching history should involve not only instruction about the content, but instruction about what it means to study a historical issue. In particular, adolescent students would need to realize that interpretations of events are often complex, with multiple opinions and varying evidence. They also need to understand that a criterion for evaluating the quality of an explanation is the degree to which it accounts for all of the available evidence, not just a selected subset. Furthermore, no single text or pair of texts is likely to provide a complete account of an event, so the task of the learner will be to gather information from all sources that are available. These realizations may be closely connected to Lee's report of a progression in how children understand the relation between the past and accounts of the past (Lee, 2001). He reported that early conceptions of the relation (prevalent in 8-year-olds) were that the past and a story about the past were isomorphic. Older children (10- to 12-year-olds) understood that the relation was more complex and that stories reflected the biases and viewpoints of their authors, making it difficult to decide among them. It was largely the oldest sample (14-year-olds) that expressed the idea that it was part of the nature of historical accounts that they would differ; hence there was a need to establish criteria for deciding among them.

In conclusion, this research, in combination with previous investigations of adolescents' processing of historical information, indicates both that they are capable of engaging in historical reasoning across texts and that there is variation in the degree to which they engage in this kind of processing. These studies establish that certain tasks and certain texts seem to be able to promote intertextual processing. Important next steps involve more systematic exploration of the processing and reasoning that can be promoted by systematic design of both tasks and texts in ways that scaffold and promote not only corroboration but sourcing and contextualization, albeit in developmentally appropriate ways. Such research needs to examine processing, reasoning, and their relations, in conjunction with students' epistemological beliefs about the domain of history and how such beliefs evolve in conjunction with changes in how students "do" history.

ACKNOWLEDGMENTS

The research was supported, in part, by a Spencer Foundation grant to David M. Bloome and Susan R. Goldman. However, the opinions expressed in this article are ours and should not be attributed to the Foundation. We thank historian David Steiner for invaluable assistance with this study. We take responsibility for

any inaccuracies or distortions regarding the history information in this study. Peter Meyerson and Cynthia Mayfield, members of the project team, made important conceptual and methodological contributions to the research reported here.

REFERENCES

Ashby, R., Lee, P., & Dickinson, A. (1997). How children explain the why of history. *Social Education, 61,* 17–21.

Barton, K. (1996). Narrative simplifications in elementary students historical thinking. In J. Brophy (Ed.), *Advances in research on teaching* (Vol. 6, pp. 51–84). Greenwich, CT: JAI.

Beck, I. L., & McKeown, M. G. (1992). Young students' social studies learning: Going for depth. In J. J. Dreher & W. H. Slater (Eds.), *Elementary school literacy: Critical issues* (pp. 133–156). Norwood, MA: Christopher-Gordon.

Beck, I. L., & McKeown, M. G. (1994). Outcomes of history instruction: Paste-up accounts. In M. Carretero & J. F. Voss (Eds.), *Cognitive and instructional processes in history and the social sciences* (pp. 237–256). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Bloome, D., Goldman, S., Cote, N., Meyerson, P., Mayfield, C., & Wolfe, M. (1999). *Reconceptualizing reading as intertextual practice*. Paper presented at the meeting of the American Educational research Association, Montreal.

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

Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J.-F. (1999). Content Integration and source separation in learning from multiple texts. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 209–233). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161–238). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Chi, M. T. H., de Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18,* 439–477.

Coté, N. C., & Goldman, S. R. (1999). Building representations of informational text: Evidence from children's think-aloud protocols. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 169–193). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Coté, N. C., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes, 25,* 1–53.

Cuban, L. (1991). History of teaching in social studies. In J. Shaver (Ed.), *Handbook of research on social studies teaching and learning* (pp. 197–209). New York: Macmillan.

Foltz, P. W., Britt, M. A., & Perfetti, C. A. (1996). Reasoning from multiple texts: An automatic analysis of readers' situation models. In G. W. Cottrell (Ed.), *Proceedings of the 18th annual Cognitive Science Society* (pp. 110–115). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Goldman, S. R., Meyerson, P., Wolfe, M. B. W., Mayfield, C., Coté, N. C., & Bloome, D. (1999, April). *"If it says so in the text book, it must be true": Multiple sources in the middle school social studies classroom.* Paper presented at the Conference of the American Educational Research Association, Montreal, Canada.

Goldman, S. R., & Rakestraw, J. (2000). Structural aspects of constructing meaning from text. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research, Vol. III* (pp. 311–335). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Graesser, A. C., Gernsbacher, M. A., & Goldman, S. R. (1997). Cognition. In T. Van Dijk (Ed.), *Discourse: A multidisciplinary introduction* (pp. 292–319). London. Sage.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101,* 371–395.

Greene, S. (1993). The role of task in the development of academic thinking through reading and writing in a college history course. *Research in the teaching of English, 27,* 46–75.

Greene, S. (1994). The problems of learning to think like a historian: Writing history in the culture of the classroom. *Educational Psychologist, 29*(2), 89–96.

Holt, T. (1990). *Thinking historically: Narrative, imagination, and understanding*. New York: College Entrance Examination Board.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49,* 294–303.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Kintsch, W., Britton, B. K., Fletcher, C. R., Kintsch, E., Mannes, S. M., & Nathan, M. J. (1993). A comprehension-based approach to learning and understanding. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 30, pp. 165–214). San Diego, CA: Academic Press.

Lee, P. (2001). *Understanding history.* Centre for the Study of Historical Consciousness, University of British Columbia, Vancouver, B.C. Retrieved April 30, 2004, from http://www.cshc.ubc.ca

Levstik, L., & Barton, K. (1997). *Doing history: Investigations with children in elementary and middle schools*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

McKeown, M. G., Beck, I. L., & Sandora, C. A. (1996). Questioning the author: An approach to developing meaningful classroom discourse. In M. Graves, P. van den Broek, & B. M. Taylor (Eds.), *The first R: Every child's right to read* (pp. 97–119). New York: Teachers College Press.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes, 22,* 247–288.

Palincsar, A. S., & Magnusson, S. J. (2001). The interplay of first-hand and text-based investigations to model and support the development of scientific knowledge and reasoning. In S. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 151–194). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Perfetti, C. A. (1989). There are generalized abilities and one of them is reading. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 307–333). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Perfetti, C. A., Britt, M. A., & Georgi, M. C. (1995). *Text-based learning and reasoning: Studies in history.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

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

Sandora, C., Beck, I., & McKeown, M. (1999). A comparison of two discussion strategies on students' comprehension and interpretation of complex literature. *Reading Psychology, 20,* 177–212.

Seixas, P. (1999). Beyond "content" and "pedagogy": In search of a way to talk about history education. *Journal of Curriculum Studies, 31,* 317–333.

Smith, J., & Niemi, R. (2001). Learning history in school: The impact of course work and instructional practices on achievement. *Theory and Research in Social Education, 29,* 18–42.

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

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

VanSledright, B. (1995). "I don't remember-the ideas are all jumbled in my head": Eighth graders' reconstructions of colonial American history. *Journal of Curriculum and Supervision, 10,* 317–345.
VanSledright, B. (2002a). Confronting history's interpretive paradox while teaching fifth graders to investigate the past. *American Educational Research Journal, 39,* 1089–1115.
VanSledright, B. (2002b). *In search of America's past: Learning to read history in elementary school.* New York: Teachers College Press.
VanSledright, B. A., & Frankes, L. (2000). Concept- and strategic-knowledge development in historical study: A comparative exploration in two fourth-grade classrooms. *Cognition and Instruction, 18,* 239–283.
VanSledright, B. A., & Kelly, C. (1998). Reading American history: The influence of multiple sources on six fifth graders. *The Elementary School Journal, 98,* 239–265.
Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology, 83,* 73–87.
Wineburg, S. S. (1994). The cognitive representation of historical texts. In G. Leinhardt, I. L. Beck, & C. Stainton (Eds.), *Teaching and learning in history* (pp. 85–135). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Wineburg, S. S. (1998). Reading Abraham Lincoln: An expert/expert study in the interpretation of historical texts. *Cognitive Science, 22,* 319–346.

APPENDIX A
Documents

Historian Michael Hatcher

There is an explanation of why the Roman Empire fell that I disagree with. I disagree that the empire covered too much territory. I have a different theory and have collected evidence to support it. My theory is that the Roman Empire fell because the Roman people became lazy. They didn't start out that way, but as the empire got larger the people became more lazy. Early writings about Roman society show that the citizens felt it was important to work hard for the good of their empire. However, as the empire grew in size and riches, the people spent more and more time in pursuit of their own selfish pleasure. They had to worry less about basic survival and had more time to engage in leisure activities. For example, paintings from that time show that the Roman people spent a lot of time laying around at the Coliseum watching violent sports rather than working for the good of the empire. Being lazy meant the Roman people took little interest in fighting for their empire. So, when the barbarians attacked in 476 A.D., they encountered a lazy Roman army. The Roman's laziness allowed the barbarians to defeat the Roman army and ended the power of the Roman Empire.

Historian Rebecca Walsh

There is an explanation of why the Roman Empire fell that I disagree with. I disagree that Rome fell because the people were too lazy. I have a different theory and have collected evidence to support it. My theory is that the reason Rome fell was

because the empire got too big. The Empire grew to its largest size under the Emperor Hadrian. It covered almost 6 million square miles. Troops had to be stationed in each place that the Romans ruled to make sure there were no rebellions. Covering so much land took all the troops the empire had. The emperors had no new troops to conquer new countries. Diaries written by soldiers tell us that there were many troops in Egypt, England, and Syria. These places were very far away from Rome, the center of the empire. The distances made it difficult to put together a large army quickly. This made it harder to defend the empire against attack from outsiders. So when the barbarians attacked in 476 A.D., the Roman army was very spread out. The large size of the Roman Empire allowed the barbarians to defeat the Roman army and ended the power of the Roman Empire.

APPENDIX B
Other Events on the Time Line of the Roman Empire

200s–400s A.D.	The slave population and the wealth of Roman Empire decreases
285 A.D.	The Empire is split into two separate empires—Eastern and Western
337 A.D.	Constantine inherits the Roman Empire from his father Chlorus
360 A.D.	Barbarians invade Roman colony in England
392 A.D.	Christianity becomes the official religion of the Roman Empire
455 A.D.	Barbarians attack Rome

APPENDIX C
Interview Questions

I. Similarities and differences:

I want to ask you about how similar YOU think the two historians are. Are their arguments similar? Do these two texts you read say the same thing about why the Romans couldn't defend themselves?

How are they the same? How are they different?

Now I would like you to tell me how similar you think each of the arguments are to each other. As you just did, can you give me a number from 1 to 5 where 1 means the arguments are very different and 5 means the arguments are very similar?

Walsh and Hatcher

1 2 3 4 5

II. Argument strength:

Next I would like you to tell me how strong you think each of the arguments are. Can you tell me a number for each one on a scale of 1 to 5 where 1 means the argument is terrible and 5 means the argument is very good?

Hatcher
1 2 3 4 5

Walsh
1 2 3 4 5

III. Additional information:

Now here is some more information about the events in the Roman Empire. All of these events are discussed in your textbook.

IV. Questions for authors:

If the authors of each of the two texts were here right now and you could ask them some questions, what would you ask them about the arguments they made?

V. Construct your own explanation:

If someone were to ask you why the Roman Empire could not defend themselves against the Barbarian invasion, what would you say to that person?

You can look back at any of the material we have here.