

# 1 1 How Students Use Texts to Learn and Reason About Historical Uncertainty

Charles A. Perfetti, M. Anne Britt, Jean-François Rouet,  
Mara C. Georgi, and Robert A. Mason  
*University of Pittsburgh*

To learn history is to learn a story: to come to know the major characters, events, and simple causal relations among events. Of course, an historian or history educator may reject the suggestion that story learning is the heart of history learning. Real learning in history entails going beyond simple stories to interpret, construct explanations, and generally to negotiate uncertainty surrounding the events. In effect, learning history requires at least primitive use of some of the text and interpretive skills employed in historical analysis.

The *teaching* of history, however, because of limitations in time and resources and because of traditions of testing, often emphasizes learning the story to the exclusion of introducing the student to the complexity of historical analysis. The interplay of social forces, for example, is likely to be sacrificed in the classroom for a simple story about dates and names. Carretero, Asensio, and Pozo (1991) referred to the European “Discovery” of America to make this contrast. Is it a story about Columbus and his relationship to the Spanish King and Queen? Or a story of how 15th-century social and economic forces promoted explorations with long-term consequences in Europe and America? As Carretero et al. pointed out, the concepts needed to elaborate the social forces story “pose a rather strong cognitive demand” (p. 29) and make for difficult learning. In contrast, learning history as a story takes advantage of the compatibility of story forms with the cognitive dispositions of the learner. We suggest that an educational goal that emphasizes a higher standard of learning must accommodate rather than reject the cognitive advantage of story learning.

There is nothing incompatible about a story approach and the complex explanation approach to understanding history. Stories provide the basic

representations of history for the student. Knowledge of the stories of history is merely the minimum standard we would be willing to apply to answer the question of whether a student has attained competence in historical topics.

A higher standard applies to what we call "historical literacy." To engage in discourse on history goes beyond storytelling to include, as part of the analysis of the story, some implicit knowledge of the "methods" of history. Such knowledge means realizing, for example, that the received story is an interpretation, a simplification of complex events that may be distorting. It implies awareness that evidence counts, that elements of a story come from records of various kinds, and some awareness of the distinction between primary and secondary sources.

In our work we addressed both the basic cognitive representation issue—the understanding of history as a story—and the issue of the use of texts as evidence. At one level, these two issues are on either side of a boundary between basic and higher standards of historical literacy. At another level, the issues reflect complementary components of how a student uses a text in constructing an understanding of historical events. The story is the student's beliefs about the event structure. Textual evidence helps the student determine his or her belief in the story.

### STORIES AND CAUSAL MODELS

In claiming that history is understood as a story, we have a specific idea in mind: Understanding history is having a mental model, a temporal-causal model, of historical events. Our assumption is that this temporal-causal model is one of the type that serves story understanding generally. Trabasso and van den Broek (1985) developed a causal analysis of narratives and argued that the comprehension of stories depends on establishing story coherence through causal relations (Trabasso, Secco, & van den Broek, 1984, Trabasso & van den Broek, 1985). Earlier work on story understanding also took note of causal structures in accounting for what readers remember from stories (Black & Bower, 1979; Omanson, 1982). Although work on story grammars, following Rumelhart (1975), emphasized story components rather than event connections in their analyses (Mandler & Johnson, 1977; Stein & Glenn, 1979; Thorndyke, 1977), even these approaches imply a role for causal connections as part of the meaning of event categories. The point of departure of more recent work has been the explicit and fine-grain analysis of causal structures.

Causal connections, which link events through causal event chains, are what make a story coherent and memorable. The "psychological reality" of these causal structures is seen when people judge the importance of story statements, recall stories, and write summaries of stories (Trabasso & van den

Broek, 1985; van den Broek & Trabasso, 1986). Although the case for causal structures as mental representations of stories appears to be quite strong, it is not completely clear how comprehension processes make use of this structure during (as opposed to after) reading. Does a search for causal structures guide comprehension, or do such structures merely reflect what comprehension brings automatically by other means? The answer to this question is not clear and probably depends on particular reader and text circumstances.<sup>1</sup>

In our study of history learning, we developed a system of causal analysis related to that of Trabasso and van den Broek (1985). Our use of causal structures, however, assumes not that these structures guide comprehension, merely that they represent what a person knows (or believes) about some historical topic. Additionally, we see causal structures as research tools—a means of assessing what students learn when they learn historical material. To assess what someone has learned about the Panama Canal story, we examine which parts of the causal model have been learned. To ask whether two students have the same knowledge is to ask whether their causal models are the same.

For example, we assume that to understand the Panama Canal story is to have a mental model that links a sequence of events by temporal-causal connections.<sup>2</sup> The Panama Canal Story relates the events leading to the

<sup>1</sup>Although we believe that *when* things occur in comprehension is a crucial question for theories of comprehension, this "online" issue is less relevant for the topic we address here. Perhaps all causal connections are established "online." Perhaps only explicitly expressed connections are established. In any case, it is sufficient to assume that causal relations have some kind of status as mental representations, regardless of whether they are established immediately as each potential causal event is encountered.

<sup>2</sup>There is a bit of equivocation in the idea of temporal-causal connections. First there is uncertainty about causality. Was there a causal relationship between Colombia's rejection of a canal treaty with the United States and the Panama revolution? Or just an incidental temporal relationship? Such uncertainty is abundant in history, and the goal of historical scholarship is essentially to reduce that uncertainty. For the nonexpert, however, there is either uncertainty or a belief about a causal relation. That is, someone may believe that the Colombian action was a sufficient condition for U.S. support for the revolution and that this support was necessary for the success of the revolution.

Such uncertainty is specific to a particular case. But there is a second equivocation in the basic idea of temporal-causal relations. Whether one event *causes* another in any strong sense, that is, whether one event is necessary and sufficient for the occurrence of a second event, may be unknowable in principle in history. Some events seem clearly necessary for others to have occurred—for there to be normal international negotiations with Panama, it was necessary that Panama be independent of Columbia. Most of the more interesting historical episodes, however, appear to fail this test. Was it necessary for the United States to encourage an uprising in order for the Panamanian Revolution to be successful? For it to occur? Was the Spanish-American War a necessary cause of renewed interest in the Canal Project? Finding sufficient causes is even more difficult, and finding necessary and sufficient causes seems virtually impossible in history.

U.S. acquisition of the rights to build a U.S.-owned canal on the Isthmus of Panama. Our analysis connects states (e.g., the U.S. desire for a canal) and events (e.g., the Panamanian revolution) by temporal-causal connections. Some of the links that a student must have in his or her mental model are the following: the U.S. desire to have faster commerce and military routes *motivated* the United States to begin negotiations for a canal; Panama's status as a part of Colombia *motivated* the United States to negotiate with Colombia; Colombia's rejection of the U.S.-offered treaty *motivated* Panama to revolt; and the successful Panamanian revolution *enabled* the United States to negotiate with Panama.

Figure 11.1 illustrates a causal analysis of the Panama Canal Story. It captures the essential events of the story, linking them together with temporal-causal predicates. This causal representation was also used as a reliable scoring template (95% interscorer agreement) to indicate what subjects remember from any text that tells the story. (See Britt, Rouet, Georgi, & Perfetti, 1994, and Perfetti, Britt, & Georgi, in press, for a detailed discussion of the causal analysis.)

### HISTORICAL LITERACY

Historical literacy involves not only the learning of historical events, but also the use of interpretive reasoning. Reasoning about historical topics, like reasoning in other domains, requires a use of evidence, argument, and interpretive strategies. Students, like historians, can come to appreciate that history stories may reflect different lines of evidence that are illuminated through different sources. These sources can be distinguished in a number of ways—documents versus nondocuments, witness evidence versus non-witness evidence, and deliberately transmitted information versus unintentionally transmitted information (Shafer, 1974). Students can learn to evaluate uncertainties in the available evidence, not necessarily in expert fashion, but in a manner that goes beyond a naive and unanalyzed acceptance of the story as truth.

When we speak of causal relationships, therefore, we are speaking loosely, not by the rules of logical inference, but by plausible inferences. The "temporal" in "temporal-causal" suggests the sense in which this is true. When one event follows another event in time, the temporal relation is taken to be temporal-causal if the circumstances seem to support a causal inference. The first event may have been necessary for the second, it may have been sufficient, or, probably the most common case, it may have been neither. The last case can be considered a probabilistic causal relation. Event 2 was made more likely to occur because of Event 1 in the context of other factors. Temporal-causal relations include all these possibilities. When someone knows an historical topic, what he or she knows is the temporal-causal relations that are the core of the story.

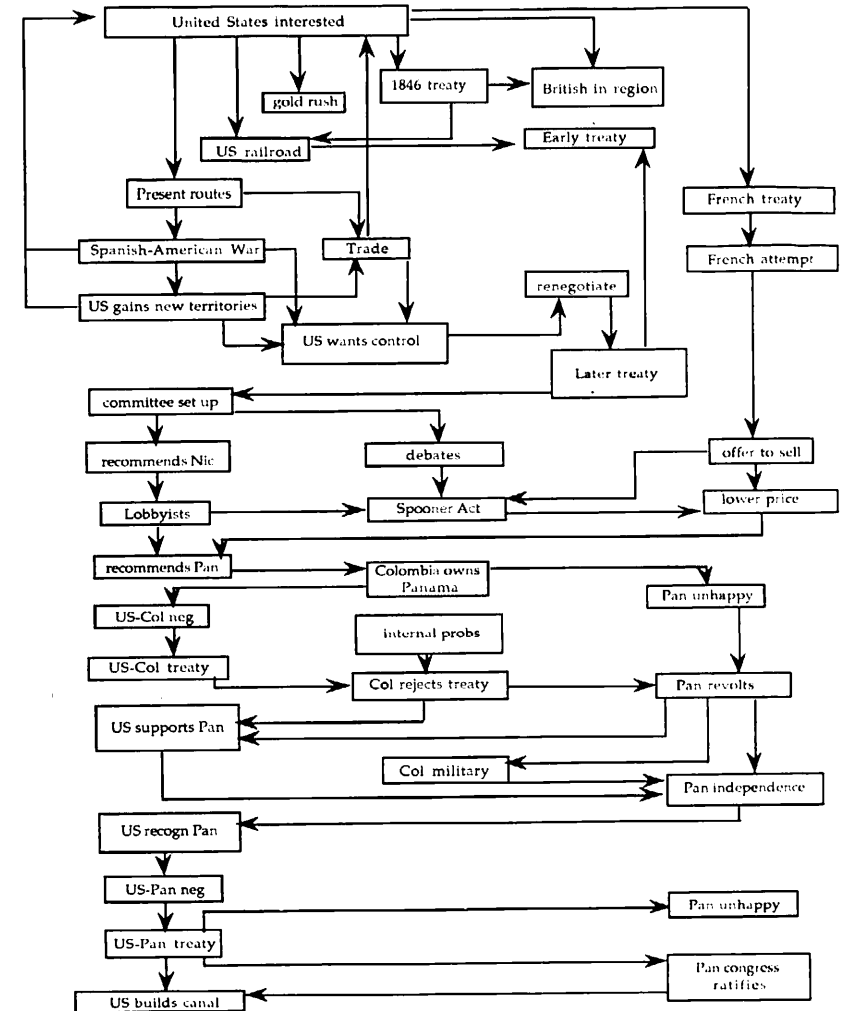


FIG. 11.1. Streamlined representation of the history of the acquisition of the Panama Canal.

An example of what it means to expect students to have a sense of historical evidence comes from Wineburg (1991), who studied differences between historians and novices (high school students). In examining paintings depicting the Battle of Lexington, students showed little tendency to use information from written historical documents as standards against which to compare the accuracy of the pictures. Asked to choose which

picture was the most accurate depiction of the battle scene, the students made limited use of historical documents that could provide critical information, even though they had just read the texts. Historians actively sought information from such documents, whereas students seemed unaware of the privileged status of the documents. As Wineburg concluded, more facts about the American revolution would not necessarily be helpful for such students

When they remain ignorant of the basic heuristic used to create historical interpretations, when they cannot distinguish among different types of evidence, and when they look to a textbook for the "answer" to historical questions—even when that textbook contradicts primary sources from both sides. (p. 84)

Wineburg (1991) identified three heuristics used by experts in reasoning from historical evidence that could be taught to students. Experts notice and evaluate the source of the document (sourcing), check the facts mentioned in the document against those in other documents (corroboration), and set events in a larger context (contextualization).

To repeat an important caveat, we do not suggest that the goals of historical literacy correspond to the full attainment of historical research skills. However, we take the view that there is value in acquainting students with the use of evidence and argument in history learning. This implies having an appreciation of historical evidence and an awareness of the origins of the stories of history.

In the following sections, we summarize two lines of research we have carried out over the past few years. We regard the two lines of work as complementary: One focuses on learning a history story, and the other examines the use of evidence in interpreting controversial parts of the story.

### THE LEARNING STUDY

The Learning Study attempted to mimic a realistic learning situation—one that approximates some of the kind of learning activities college students commonly encounter in history classes. The results of this study are reported elsewhere (Britt et al., 1994; Perfetti et al., in press) and are only summarized here.

This study had two basic parts. The first focused on the events leading up to the building of the Panama Canal (the Acquisition Story). The second part dealt with a recent historical issue related to the Acquisition Story: the 1978 negotiations between Panama and the United States over the future

ownership of the Panama Canal. Here, however, we limit consideration to the first part of the study.

In this study, six paid undergraduate students (4 males, 2 females; age range 18–25) participated in an extended learning situation in which, over several weeks time, they read four different texts on the U.S. acquisition of the Panama Canal Zone. The first text was a neutral and relatively short text (about 1,700 words) written by the experimenters to provide a brief introduction to the events and characters. The second text, written by The Center for Strategic Studies (1967), was also relatively neutral. However, it was clearly authored by Americans. Texts 3 and 4 were longer texts that were mostly polemical attempts to influence public opinion in the United States. Both were published in 1978, during the debates in the United States concerning the ratification of a new U.S.-Panama treaty that returned sovereignty of the canal zone to Panama. One, by Representative Philip Crane (a republican from Illinois), strongly argued against the treaty. The other text was written by a Latin American scholar at Cornell University, Walter LaFeber. LaFeber's (1978) text argued in favor of treaty ratification.

After every text assignment, subjects were asked a series of questions about the story. Their answers provided data on the course of their learning over a period of weeks. Because the focus here is on students' use and evaluation of history texts, we do not discuss the comprehension results in detail (see Britt et al., 1994, for more details about the comprehension results). Students learned the main story rapidly. Events, as defined by our causal analysis, were largely learned by the second text assignment. Learning the supporting facts of the story was also rapid, but somewhat slower than the learning of events.

### Understanding Uncertainty

Although the students learned the events quickly and gradually built up a representation of the details of the events, there was an additional question: We were also interested in whether students without much history training used some of the interpretative skills involved in reading history.

*Detecting Author Bias.* One such skill is detecting the bias of the author of the text. After each text assignment we asked the subjects, "Did the author present a neutral coverage of the events? If not, what do you think the author's attitude was?" The students' first text was, in fact, as neutral as possible; there was no interpretation and no slanted or colorful wording. The second text was written for an American audience for policymaking and could be considered as taking an American perspective, but again there was very little interpretation and no biased phrasing. The final two texts were both selected because they presented a bias. The bias in text 3 was

achieved by selective omission of events, including author interpretation and opinion and using slanted and colorful language. The bias in text 4 was achieved by selective omission of events and included author interpretation and opinion.

We found that all but one of the subjects could detect bias in texts; the one exception thought that all of the texts gave a neutral presentation. The five remaining subjects detected bias in at least two of the three texts with bias, while correctly judging the first text to be neutral.

Although subjects could detect bias when asked, it is not clear that they used this skill actively when reading. One subject said, "I only thought about a real bias when I was getting questioned." Although this particular comment was in response to the unbiased first text, it does illustrate the possibility that asking questions made subjects more aware of bias during later assignments.

Most responses concerning the author's attitude were simple statements that the author had an attitude about something. Occasionally, however, subjects elaborated their judgments. They suggested three factors that constituted author bias: selectively omitting events, interpreting events to persuade the reader, and using slanted or colorful language. In two cases subjects mentioned that the author skipped information. One subject mentioned, "He skips U.S. businessmen; he doesn't mention Cromwell and their involvement like LaFeber did. He was conservative, didn't say enough about the Senators." In three cases, subjects mentioned that the author was trying to support a thesis and this was why they judged it as biased. As one subject put it: "[Crane was] pretty much U.S. supportive. He wanted to make it (pause) set the record straight. Show that the U.S. was completely fair in their dealings. He was pretty much trying to show that the U.S. was the good guy."

In only one case did a subject ever specifically mention biased or slanted wording as a reason for judging a text as biased. Failure to note biased language is especially noteworthy because the author of one of the texts was quite adept at colorful prose: "Three years later, the government of Colombia (Panama was then only a geographical expression, a poverty-stricken, pestilential province of Colombia racked by chronic internal disorder)" (Crane, 1978, p. 4). It is interesting that no subject mentioned the biased language in this text.

Subjects were not uniformly quick to attribute bias, however. One subject, defending a judgment that a text was neutral, noted that "He [LaFeber] threw his opinions in but he let it be known that those were his opinions. He showed the facts then he showed how he thought. He never really took a side." One subject considered a document to be neutral if the author used distancing language: "It was pretty much neutral. He didn't even refer to Americans as Americans but always as North Americans. I

found that rather odd. It was always the North Americans vs. the British or the North Americans dealing with the Panamanians." Finally, one subject seemed to think that vague writing meant that a document would be neutral: "Yes, kind of [neutral]. . . . I guess he was neutral by not giving all the facts. On the other hand, he can't sway you or give you an opinion." Interestingly, this same subject later said that text 4 was neutral because "I think he presented all the facts. I think he was pretty neutral."

*Handling Inconsistencies Among Texts.* A second skill needed to understand and reconstruct past events is that of reconciling inconsistencies among texts. Of special interest was a particular detail reported differently in the four texts: the time taken by the U.S.S. Oregon to get from its Pacific coast station to join battle in the Caribbean during the Spanish-American War. This long voyage around Cape Horn was mentioned to illustrate the military advantage of a shorter transoceanic route. Three of the texts gave durations of approximately 2 months (67 days in text 1, 2 months in text 2, and 68 days in text 4), but text 3 produced a much longer time (90 days).

Subjects learned this detail quickly and mentioned it quite often. The U.S.S. Oregon inconsistencies were handled in various ways, occasionally ignored, but usually assimilated either by using the most recently read text or the first-read text. One subject mentioned the Oregon only after the first text, appearing to ignore the later inconsistencies. A second subject gave the duration noted by the first text (67 days) on all but one occasion (after the second text he stated the duration, 2 months, given by second text). The other four subjects gave the duration from the most recent text. Two of these four subjects, however, also noted that there was an inconsistency in the texts about this point. One subject, after reading text 3, which gave the longest duration, said: "It took the U.S.S. Oregon 90 days or something like that—every time I read it it gets longer—to go all the way around South America."

*Detecting the Incompleteness of Text.* A third relevant text skill is detecting the incompleteness of the story. After each text assignment we asked the subjects, "What else would you like to know?" On only one occasion did a subject not want more information about something, and this was after having read all four texts. There were three classes of information that the subjects wanted. Some subjects wanted more information about the basic events and characters. They were still trying to be sure about details and facts of the basic story. Examples included: "They mentioned someone in Bogota and negotiations in Bogota but I don't know what Bogota is"; "I'm still hazy on how France got in there. I know we were there with Great Britain and we signed a treaty to get them out and then all of the sudden there was France." A second type of information subjects

wanted was more historical context. Some examples were: "Maybe more about the Panamanian people themselves. What type of place was it to live in?"; "What happened between 1850 and 1878?"; "What was [the] political situation in Congress at the time?" Finally, subjects wanted more information on the controversial parts of the story, such as: "I'd like to probe Bunau-Varilla. Prove to me why he did the right thing for Panama even though it looks so skewed in [the] U.S.'s favor"; "More background on the treaties, what was said in the treaties — just names given"; "If Panamanians had wanted their independence, did [the] U.S. really push them? It's one thing to say Panamanians wanted to do it or they did it because if [the] U.S. was going to help [the revolution], they might as well [revolt]."

*Resolving Conflicting Views.* A fourth text skill is negotiating the uncertainty prompted by contradictory views given by different authors. We asked subjects four questions involving controversial issues that were covered differently by the texts: (a) "Were the United States' dealings with Colombia fair to Colombia?" (b) "Were the United States' dealings with Panama fair to Panama?" (c) "Was the United States' role in Panama's revolution justified? Why or Why not?" and (d) "Do you believe the United States should have negotiated with Bunau-Varilla or waited for the two Panamanian negotiators?" The first text, by design, was scarce in details on these controversial events. The subjects' initial response to these probes therefore provided relatively uninformed opinions. The same questions were asked again after every text. Thus, at most, subjects could change their opinion on 12 occasions (4 questions and 3 opportunities to change the response to each). We classified responses as "yes" (the United States was fair/just), "no" (the United States was not fair/just), "no answer" (no opinion or commitment on fairness/justness), or "both" (the United States was fair/just in some respect and not fair/just in other ways).

For the most part subjects did take a position in response to the opinion questions. On only 13% of the opportunities did subjects fail to draw a conclusion about the controversial issues, and most of these occasions occurred after either the first or second text, both of which were free of an opinion on the matter. Overall, subjects believed that the United States was fair in its dealings with Colombia and Panama in both the treaty negotiations and the revolution. Fifty percent of the opinions were that the United States was fair, whereas 26% of the opinions were that the United States was unfair, and 11% of the opinions were that, in part, the United States was fair and, in part, they were not fair.

We found reasoning about the uncertainties of the story to be tentative and affected by the text read most recently. There was an average of 7 opinion changes out of 12 opportunities. Five of the six subjects changed their opinion at least half the time. The opinion changes moved toward

agreement with the most recently read text. The other subject was much more stable, changing his opinion on only two occasions. Additionally, this subject consistently was also very pro-American. He believed that the United States was fair on 88% of the questions, for example, "Fair is fair. If they didn't like it they didn't have to sign it"; "I think we knew we should've negotiated with Bunau-Varilla because the other guys might have gone for the other terms. I think the U.S. was just trying to get the best deal we could. Fair is fair. You signed it, that's it"; "Yes I think so. How it went for our advantage was good. If they had the advantage, they would've done it." Other than this one subject who was more driven by his prior beliefs, the subjects' opinions were rather changeable and tentative, affected by the most recent text they read.

Although subjects were swayed in their opinions, they showed little spontaneous interest in the possible role that other documents might have in helping them come to conclusions. Subjects did not mention that treaties, cables, official orders, or other types of documents might aid in their forming their own opinion or checking the validity of the author's interpretation, either in response to the fairness question, or in response to the specific question "What else would you like to know?" Only one subject read the treaties given as optional reading, and that was a subject who, after the first reading, mentioned that he would like to know more about the treaties. No other subject asked for primary documents or used the primary documents given them.

### Summary

Subjects learned the basic history story rapidly, while more slowly acquiring the supporting details over successive readings. The causal model effectively captured the essentials of the story that they had learned. Subjects demonstrated mixed abilities in text-based reasoning. Although all subjects, except one, were able to detect author bias, their opinions on controversial events showed a high degree of malleability and tentativeness, using information from texts to modify judgments previously stated. When given the opportunity to ask for additional information, subjects did not request primary documents.

## THE DOCUMENTS STUDY

Subjects' failure to use provided evidence (treaties) and their failure to ask for other documents (e.g., treaties, cables, letters) directed our attention toward examining their understanding of and use of documents as evidence in historical reasoning. It is not surprising that our students were not aware

that they were missing evidence (i.e., primary documents) while trying to understand the story of the Panama Canal. They were satisfied learning the story as told by each author. The results are in line with those from bright high school students who rated textbooks as the most trustworthy of documents, in contrast to historians who, detecting both the text's inconsistencies with better evidence and the bias of its "subtext," rated it the least trustworthy of documents (Wineburg, 1991).

This problem leads to an examination of how students understand the relative privilege of certain types of documents when forming opinions about controversies in history. The use and interpretation of documents is of critical importance to historical researchers. Are college students aware of the privileged status documents have as evidence? How do they use texts as evidence in dealing with the uncertainties of the stories?

Twenty-four undergraduate students (17 males and 7 females) from the University of Pittsburgh participated in a 4-hour experiment. There was also one graduate student in history who served as an expert.

The undergraduate subjects were between the ages of 18 and 31, comprising 13 freshmen, 5 sophomores, 3 juniors, and 3 seniors. Reading ability, as measured by the Nelson-Denny Reading Test, ranged from the 16th percentile to the 99th percentile, with an average of the 54th percentile (based on the scale for sophomores; Brown, Bennett, & Hanna, 1981).

The subjects varied greatly in their knowledge of history, in general, and of Panama, in particular. Only 10 of the subjects had taken a college history course (one subject had taken 10, one subject had taken 3, and eight subjects had taken 1). Subjects' historical knowledge was evaluated by a three-part history knowledge test. General knowledge was assessed by subjects' answers to 15 two-part open-ended questions on general history topics (e.g., What was the Weimar Republic? When was it?). Their scores ranged from 10% to 80%, with a mean of 40% correct. The expert was able to answer every question correctly. Specific Panama history knowledge was assessed by a 21-question test covering the Panama Canal and its history (e.g., Why is the canal used?). The subjects scored between 6% and 71%, with a mean of 44% correct. The expert answered 81% correctly. Finally, geographic knowledge was assessed using a 9-question map test. Most subjects accurately located the targets relevant to the history of the Panama Canal: United States, Panama, Colombia, and the Pacific Ocean. Overall the scores ranged from 11% to 100%, with a mean of 76%. Again the expert had a perfect score.

The period of the acquisition of the Panama Canal was segmented into four controversial problems for which a large number of documents were available. For example, one controversy involved the extent to which the United States participated in the Panamanian revolution. For each of these problems, we searched the document literature and sorted documents into

three very broad categories: secondary, intermediate, and primary. These categories were differentiated according to the following four features: when the document was written (before, during, slightly after, or significantly after the events), whether the author was a participant (or not), the document status (official, public, private), and the bias of the author of the document (defending own actions, commenting and interpreting events, or neutrally stating and describing events). There are, of course, additional features one can use. Our choice of these features allowed us to create a small set of documents that could be managed within the format of our study.

The features of each document for the four problems are summarized in Table 11.1. The *secondary documents* were documents written after the events, by a nonparticipant (historian or politician); they were published and were considered biased (authors supported opposing interpretations of the events). The *intermediate documents* were documents written either at the time or slightly after, by a participant in the events, with a biased perspective of either defending his or her own actions or commenting on others' actions. We considered these documents "intermediate" because each had features in common with primary documents but at least one feature that qualified it as a interpretation of the events. The *primary documents* were excerpts from letters, treaties, and official correspondences. They were written at the time of or before (e.g., 1846 Treaty) the controversy, the authors were always participants, the purpose of the document was to state rather than comment, and the document was official (except for the Roosevelt letters). These documents were unbiased (again with the exception of the Roosevelt letters) but could be interpreted to support one side or the other. All the documents were very short excerpts of between 76 and 383 words, with an average of 177 words.

In addition to these documents, we also prepared a *textbook document* and a list of facts for each problem. The facts we included represented the basic, noncontroversial version of the story and were agreed on by all the

TABLE 11.1  
Features of the Documents in the Document Study

	<i>Time written</i>	<i>Written by a participant</i>	<i>Document status</i>	<i>Author bias</i>
Secondary	After	No	Public	Biased
Textbook	After	No	Public	Neutral
Intermediate	During or slightly after	Yes	Variable	Biased
Primary	Before or during	Yes	official (*)	Neutral (*)

\*Except in problem 1 (Roosevelt letters).

sources. Like the secondary documents, the textbook was written after the events, by a nonparticipant. Like the primary documents, the purpose was to describe and state the events in a neutral way.

The experiment was carried out in three sessions. The first session included the history knowledge tests. Also, in order to give subjects an understanding of the basic story of the U.S. acquisition of a canal in Panama, they were given a 2,200-word text to read (a modified version of text 2 from the Learning Study). We edited the text to ensure that the material related to the four controversies was very general or vague. Afterward, subjects were given 11 comprehension questions (e.g., "Name two reasons why the French attempt to build a canal failed"; "What actions did the U.S. take during the revolution?") to ensure basic comprehension of the story. Subjects were also given each of the four controversy problem statements (e.g., "To what extent did the United States participate in the planning of the Panama revolution?"). We informed them that we merely wanted their opinion because the text was not very detailed.

Based on subjects' combined scores on the history test (general history, Panama knowledge, and geography), 12 subjects were assigned to the primary group. This group was given access to primary documents for each of the four controversy problems. The 12 other subjects were assigned to the secondary group<sup>3</sup>. The subjects in this group did not have access to primary documents. Instead, they received additional secondary documents. The expert was assigned to the primary group.

In the second session, subjects were given the first two controversial problems in a hypertext environment run on an Apple Macintosh IIsi. They were first given the problem statement—"To what extent did the United States participate in the planning of the Panama revolution?" Then they were given a screen that had a list of main facts presented in chronological order. When they had read these, subjects pressed a button to get a menu of the sources of the seven available documents (author name, credentials, type of document and date). Figure 11.2 shows the menu screen for problem 2. The first five documents were the same for both the primary and secondary group: two secondary documents, two intermediate documents, and the textbook passage. The final two documents were two additional secondary documents (secondary group) and two primary documents (primary group). Any selection in the menu resulted in the presentation of an extended source of the document. Subjects then could decide to read the actual documents or to go back to the menu. Previous selections were

<sup>3</sup>The groups had similar prior history knowledge as determined by the combined score on the history knowledge test. The secondary group averaged 29 correct of 60 (with a variance of 13), whereas the primary group averaged 28 correct of 60 (with a variance of 12).

marked in the menu to ease the study process. Subjects were required to select each source at least once for inspection.

For each of the four problems, subjects were given 15 min to study the list of main facts and the set of seven documents. When the subjects had studied the documents and believed they had an informed opinion on the controversy, they could exit the presentation system. Then they were to write a 1-page essay expressing their opinion. They were not able to use the documents while writing, but were given a list of the names and dates mentioned in the documents and the sources of the documents to avoid memory problems. After completing their essay, the subjects were asked to rank the seven documents according to their usefulness ("the extent it helped you build up your informed opinion") and their trustworthiness ("the extent you trust what the author says") on a 7-point scale. They were also asked to briefly justify their ranking of each document. In the third session, subjects followed the same procedure for the final two controversies.

Our focus here is restricted to the results of document evaluation. (For a

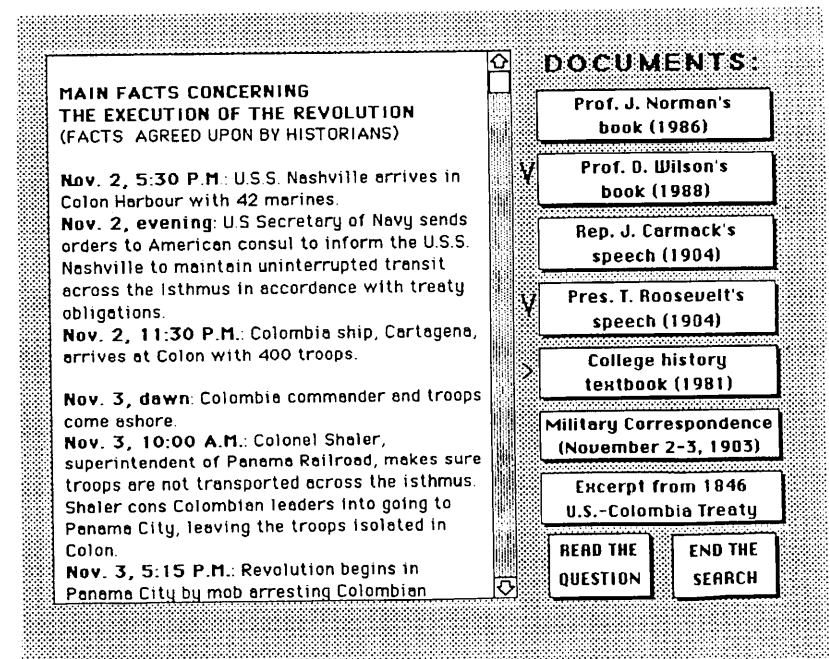


FIG. 11.2. List of facts (left) and document menu (right) in the prototype hypertext system (the marked items have been previously selected).



more complete discussion of the results see Rouet, Britt, Mason, & Perfetti, in preparation.) First we report subjects' ranking of the various documents and their justification of those rankings. Then we examine how the results of the document study address students' ability to deal with uncertainty. In particular, we address three questions answered in the previous study: Do students detect author bias? Do they detect incompleteness of texts? Are they able to resolve conflicting views?

### Evaluating Documents

We analyzed subjects' judgments of the trustworthiness and usefulness of documents. Subjects ranked the documents for each problem on a 7-point scale, with a 7 corresponding to the most trustworthy or most useful document and a 1 corresponding to the least trustworthy or least useful document. Although the results were consistent across problems, the pattern of ranking of the primary documents in problem 1 (letters) differed from the other three problems (cables and treaties). This difference may reflect the primary document features for problem 1 (type of document and the bias of the author; see Table 11.1). Because of its uncharacteristic pattern, problem 1 was excluded from any analysis about primary documents.

**Ranking of Trustworthiness.** Figure 11.3 shows the average ranking of document trustworthiness collapsed across problems 2 to 4. One interesting comparison is how the textbook was judged by the two groups. The secondary subjects ranked the textbook as most trustworthy for all three problems (average ranking = 6.5). The primary group also trusted the textbook (average ranking = 5.1), but they always ranked the primary documents as more trustworthy (average ranking = 6.1). Thus, when subjects are given primary documents (e.g., treaties and cables), their trust in the textbook is reduced. It is also worth noting that the ranking pattern of the expert was similar to the pattern of the primary group: In problems 2 to 4, the expert gave the textbook an average rank of 4.7, and primary documents an average rank of 6.5.

Although there are some patterns across types of documents in their ranked trustworthiness, subjects did not make these judgments on the basis of document type alone. Their judgments showed sensitivity to the role of the document in the problem space. In each of the four problems, one of the intermediate documents was written by a participant defending his own actions (as one subject put it: "[It was] very biased, written in his best interest to cover his own ass."). Although the other intermediate document was also written by a participant, the author was commenting on the actions of another character rather than defending his own actions. As shown in Fig. 11.4, subjects in both groups were sensitive to this feature when

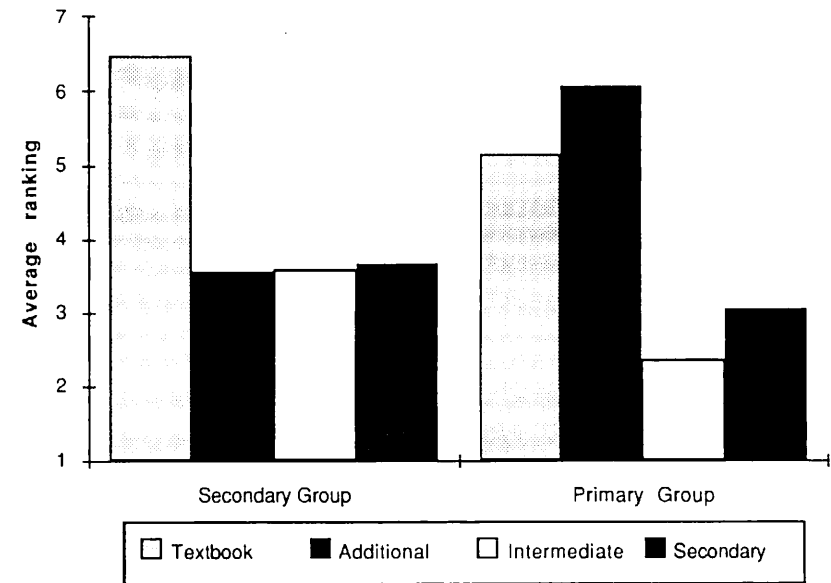


FIG. 11.3. Average ranking of document trustworthiness (problems 2-4).

judging the trustworthiness of the document. The intermediate document written by a participant defending his own actions was clearly trusted less than the other intermediate document. Again, the expert's average rankings were similar (1.8 for defending authors; 4.8 for nondefending authors).

**Ranking of Usefulness.** Figure 11.5 shows the average ranking of document usefulness collapsed across problems 2 to 4. By design, all the documents were relevant for all problem statements. Perhaps for this reason, there was less distinction among types of documents in the ranking of usefulness than in trustworthiness. The secondary group considered both the intermediate documents and textbook useful. The primary group ranked the primary documents as most useful. Thus, although the secondary group clearly trusted the textbook the most, they did not find it more useful than all other documents. Subjects given the primary documents consistently found these treaties and cables both more useful and more trustworthy than the other documents. The expert found the primary documents even more useful than the college students did (6.3 vs. 5.0).

**Justification of Ranking Data.** Subjects were given the extended sources and asked to justify their trustworthiness and usefulness rankings. These revealed that subjects' answers in both groups were sensitive to

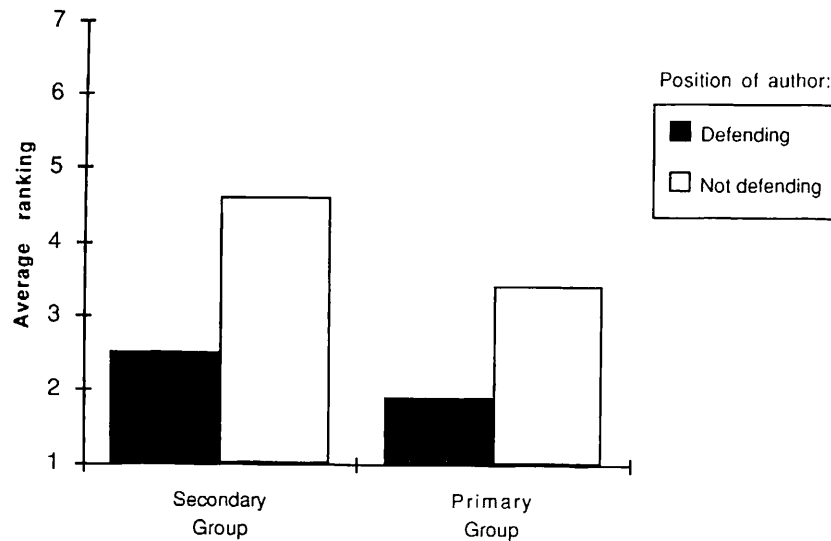


FIG. 11.4. Average ranking of trustworthiness for the intermediate documents (problems 1-4).

factors available in the extended source statement (e.g., the author or the type of document) and factors available only from their recall of the document (e.g., specific content statements). Features used by the subjects in their justifications fell into five categories. Figure 11.6 shows the proportion of judgments (collapsed over subjects and problems) in each category for both trustworthiness and usefulness justifications.

The first category concerned characteristics of the *author*; it was more frequently used in justifications of trustworthiness (33% of the responses) than in justifications of usefulness (16% of the responses). Several characteristics of the author were mentioned in these justifications, such as: author's credentials (e.g., "professor," "President"), author's motivations (e.g., "has his reputation at stake," "the President would say anything to keep from being impeached"), author's participation (e.g., "eyewitness account," "he had firsthand information"), and occasionally the subject's opinion of the author (e.g., "blind fool," "author is biased"). Characteristics of the author were mentioned equally often by both the secondary and primary subjects.

The second category was the type of *document*. Subjects mentioned both the type of text (e.g., "it's a textbook," "actual treaty") and an evaluation of the type of text (e.g., "college text would not print false facts," "as a treaty it has no bias"). Comments about the text were more frequent in justifica-

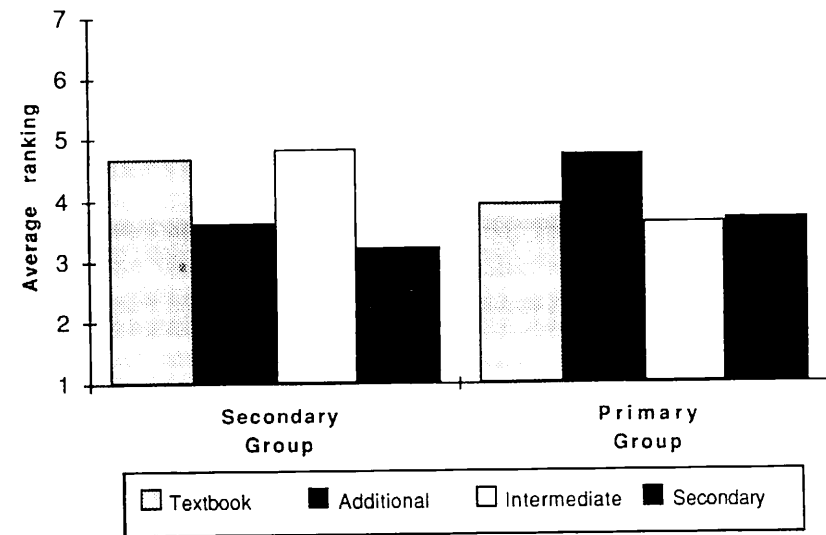


FIG. 11.5. Average ranking of document usefulness (problems 2-4).

tions of trustworthiness (19% of the responses) than in usefulness (9% of the responses). Also the primary subjects mentioned the type of document more frequently than the secondary subjects. This increase in focus on the document type, however, occurred mainly for primary documents, rather than an overall increase across documents.

The third category of justification was a mention of the *content* of the document. Content-based justifications included: a statement of the position taken (e.g., "argued they influenced"), an evaluation of the perspective or tone (e.g., "biased," "too one-sided"), an evaluation of the argument (e.g., "nothing to support his claim," "well-supported argument"), and a statement or evaluation of the content ("mentions the New Panama Company," "denies guaranteeing intervention, but not hinting at it"). Unlike the author and document characteristics, comments about the content were more frequent in justifications of usefulness (49% of the responses) than trustworthiness (36% of the responses). Both primary and secondary subjects mentioned content characteristics equally often.

The fourth justification category was a reference to the subject's own *opinion* as part of his or her justification. A typical use of this category was the subject's explicit statement of his or her own opinion and the document's agreement with this opinion. For example, subjects made statements such as: "I don't believe it was a good deal so I don't believe him." This category of justifications was used more frequently for usefulness (10% of

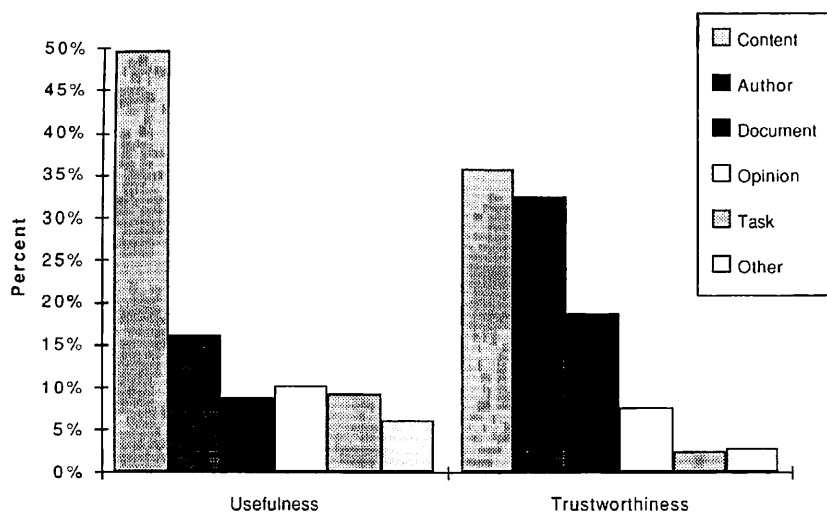


FIG. 11.6. Percentage of each type of justification for document ranking.

the responses) than for trustworthiness justifications (8% of the responses) and was mentioned equally often by secondary and primary subjects.

The final category of justifications was *task* characteristics. Subjects occasionally justified a particular document rank by referring to another document or the document set as a whole. Examples included: “[It] did not agree with other documents he had written”; “Supplemented Stanton’s arguments well”; “He informed the reader of events that the others lacked.” This category was infrequent, but slightly more usefulness justifications (9% of the responses) were based on task characteristics than trustworthiness justifications (2% of the responses). Primary and secondary subjects mentioned characteristics of the task equally often.

In summary, subjects were sensitive to specific source factors in justifying both trustworthiness and usefulness of documents. Subjects judged the usefulness of a document based on its *content*, whereas they judged the trustworthiness of a document based on both its *content* and its source (i.e., *author* and *document type*).

### Understanding Uncertainty

Students were sensitive to various features of documents in evaluating their trustworthiness and usefulness. They trusted primary documents more than others and mentioned many relevant aspects of the source and content in their justifications. Can we link these observations to those text-based reasoning skills we probed in the Learning Study? Although the tasks in the

Learning Study and the Documents Study were very different, we reexamine three of the skills probed in the Learning Study to see whether the Documents Study sheds any further light.

*Detecting Author Bias.* In the Learning Study we found that most of the students were able to detect the author’s bias when reading multiple texts. Although the Document Study did not specifically ask subjects about the author’s perspective, we can examine subjects’ justifications to find evidence of sensitivity to author bias. Of the 169 times that subjects mentioned the perspective of the author in the justifications, 100 of these were statements that the content was biased, and 69 were statements that the content was neutral. Of the 100 cases of author bias, 99 were from either secondary or intermediate documents that were biased by design. One of the 100 cases was a judgment that one of the Roosevelt letters (considered primary here) was biased. On the other hand, the documents judged neutral were more often (90% of the time) neutral by design: the textbook and the primary documents. These were documents that did not take a stance on the controversy. In only 10% of the cases did subjects state that the author was neutral when the document was arguing for one side or the other of the controversy, that is, intermediate and secondary documents. Thus, consistent with the results of the first study, students showed a sensitivity to authors’ perspectives in these texts.

*Detecting Incompleteness of a Text.* As in the Learning Study, we asked the subjects to indicate whether they detected incompleteness in the material given. Because the subjects’ essay task went beyond merely learning the story, their responses were expected to be different from those of the Learning Study. The questions of interest concerned subjects’ sensitivity to missing primary documents and missing but related parts of the story.

We asked subjects if there was anything else they would like to know about this story, if they lacked any documents, and if they would like more information on some aspect of the story. The subjects in the secondary condition were less satisfied with the documents they were given. On 60% of the occasions, the secondary subjects wanted more information, whereas on only 35% of the occasions did primary subjects state that they wanted more information. Responses of both groups of subjects included references to primary documents (e.g., treaties, orders, and correspondences). Subjects were also interested in having other points of view (e.g., the President of Panama’s view and Colombia’s view). Thus, subjects, even those not given any primary documents, showed an interest in having treaties and other primary documents to help them gain an informed opinion of the controversies that they addressed.

There is an interesting comparison that can be made between the two

studies with respect to this question. When the task was to learn the story, as in the first study, subjects asked for information that would help fill in the gaps in their understanding of the events and relations among events. They did not ask for primary documents and, with one exception, did not read the primary documents they were given as optional reading. When the task, however, was to read multiple documents to come to an opinion about the events, as in the second study, subjects asked for both evidence and opinions.

*Resolving Conflicting Views.* As in the first study, we asked subjects their opinion of controversial events for which they were provided conflicting views. Preliminary analysis of the opinion essays showed that subjects were more likely to take a stronger stance after studying the documents than before. In addition, the subjects who were given access to primary documents were more likely to change their overall opinion than secondary-documents subjects. Although both groups used facts to support their claims, the primary group was more likely to explicitly cite references. In fact, 65% of the primary group subjects made at least one reference in their essays, compared to 39% in the secondary group. Also primary group subjects referred mostly to primary documents (62%, compared to 38% in the secondary group). Interestingly, all document types (primary, secondary, intermediate) were cited, with the exception of the textbook. Thus, having access to primary documents seemed to increase the likelihood that the students would cite references for the facts that they produce in their essays.

### Summary

The results of the Documents Study demonstrate that college students are able to use some relevant text-based reasoning methods. They were sensitive to bias in documents; they judged primary sources and texts as more trustworthy than sources with explicit bias (secondary and intermediate). In addition, students' understanding of and reasoning about a controversial problem in history was affected by having access to primary documents. Perhaps most interesting is the impact that such documents appeared to have on students' willingness to use documents of all types in constructing arguments. It seems clear that when several documents are simultaneously available and the nature of the historical controversy is made explicit, students understand the importance of primary documents.

## GENERAL DISCUSSION

The two studies we described represent our two-dimensional approach to the question of historical literacy. On the first dimension, historical literacy

was studied as the ability to understand the stories of history. On the second dimension, we approached historical literacy as a reasoning skill, a skill that involves interpreting multiple sources of historical evidence.

### History As Story Understanding

History learning, at all levels, centers around the cognitive representation of events in terms of temporal-causal structures. The stories of history are learned, understood, and expressed in these terms. This is not a claim about the nature of history as a discipline, but the nature of cognitive structures. Indeed, so basic are these structures that they are accessible to young children, who show a natural grasp of stories that have the basic temporal-causal structure (Mandler & Johnson, 1977; Stein & Glenn, 1979). To put this point specifically in the context of history, we found that fifth-grade children, although they show some variability, are adept at learning the basic story of the Panama Canal (see Britt et al., 1994). (This study also found that some subjects learned an interesting tangential episode about problems with malaria encountered by the French in trying to build the canal somewhat better than the events linking the United States with Colombia and Panama. We are reluctant to overinterpret this observation, because we think it is largely a function of the text itself.)

The adult Learning Study reinforces this view of history as story learning. In addition, it provides an enriched view of the course of real learning with multiple texts. We observed the rapid learning of core events and the gradual acquisition of relevant (and even less relevant) details. But we found also that continued reading with additional assignments led to increased learning of core events. What students learned after one assignment, although quite representative of the whole story, was not in fact the whole story. Reading later assignments continued to help the student fill in the essential story. After each of the first two texts, subjects learned the early part of the story best. They focused on the U.S. motivations for building a canal and early obstacles (e.g., the Spanish-American War, the Gold Rush, the French attempt, and the British involvement). After each of the final texts, subjects began learning the end of the story much better. Here, they focused on the later obstacles and their resolution, for example, the U.S.-Colombian negotiations, the Panamanian revolution, and the U.S.-Panamanian negotiations. Although there are other possible explanations for this learning pattern, one plausible explanation is that understanding later causally related events is contingent on understanding the initial events.

An interesting question is raised by the results of the Learning Study. Is the value of multiple text assignments to increase the students' learning because of exposure to different texts or simply because of more reading? In other words, would it have mattered if our students had read the same text

each time? We believe the multiple text format is uniquely responsible for effective learning on theoretical grounds. First, a new text controls students' attention in a way that a repeated text cannot. This is a simple matter of the arousal value of novel stimuli, a very general cognitive fact. More interesting is the possibility that differences in the manner of presenting information promote a more accessible cognitive representation. A representation based on a single text has a limited number of "access" points, that is, specific linguistic cues that connect to the representation. A representation based on different texts has a richer set of linguistic access points, reflecting the different ways in which the texts presented similar and identical information. Finally, there is a third level of possible advantage, related to the second. Texts with different perspectives may promote a deeper learning of the core events on which the texts agree. Having to confront contradiction may increase the stability of a representation, reinforcing the core of the story and delineating the points of contention. Our study is silent with respect to evidence on the value of multiple texts. We are pursuing the issue, however, in other work with high school students.

### Reasoning with Multiple Documents

Our Learning Study tells us that college students with minimal history background are quite ready to engage history as interpretation. Our subjects were quick to engage the controversies and to detect bias in authors. Perhaps most interesting is the extent to which they showed an appropriate tentativeness in their opinions, in general. Although one student did live up to the expectation that ideologues are not about to be influenced by alternative perspectives, the majority used what they read each time to modify their understanding of the controversial issues. Between "wishy-washy" and "unalterable ideologue," our students were divided 5 and 1, respectively.

Overall, students had little trouble recognizing that history stories are interpreted events, and that controversies not only surround the interpretation but are perpetuated by historically based texts. However, they showed a dimmer awareness that additional documents might be useful to help them decide what to conclude. Part of this might be the factual and quasi-scholarly nature of the texts they read. When students have the impression that they are getting lots of real data through the author's citations, they might be less inclined to anticipate better sources. But the larger factor is probably simply the demands of the typical learning task. Although our learning study went well beyond what a typical text learning exercise might entail by forcing both multiple text reading and questioning designed to elicit awareness of controversy, it did not confront students with problems for which document use would have an obvious value. In

contrast, the document study, by posing problems and presenting document choices in connection with the problems, allowed students to show their appreciation of the role of documents.

And what was the level of this appreciation? Like Wineburg (1991), we did find that students considered textbooks trustworthy. But here the question was different, because there was no built-in contradiction between the textbook and other information sources. The textbook was trustworthy, not because it was considered beyond reproach, but because it had useful information and a lack of "bias." (Students' inability to detect textbook bias is an interesting additional issue. There does seem to be a general assumption favoring the neutrality of textbooks.) More interesting, however, is the context sensitivity of students' judgments of documents. When students were given primary documents, they considered them trustworthy. When students did not have primary documents, they considered documents that have some of the features of primary documents ("intermediate" documents) as trustworthy. It is clear that college students do have a concept of document privilege that can be demonstrated in some situations.

Finally, we note the interesting value that primary documents have on how students develop arguments. We found that students provided with primary documents not only used them more in writing essays, but they made greater use of other documents. Our hypothesis was in fact that exposure to primary sources increases students' sensitivity to the possibility of citing sources in connection with arguments. The evidence supports this hypothesis.

A good deal of our work is directed to the development of accounts of causal events (Britt et al., 1994) and argument analysis (Rouet et al., in preparation). In both cases, there is a need for a clear methodological tool based on a theoretical analysis. In the first case, we linked the development of an analysis tool—the temporal-causal template—to a model of the cognitive representation of historical events. In the second case, we linked an analysis of the argument structure of an essay to assumptions (we are reluctant to term these assumptions a "theory") about the level of discourse that comprises the heart of short arguments. We believe that further progress in describing both history-as-story and history-as-interpretation depends on having such tools.

### Learning to Learn History

We turn finally to the implications of our two-dimensional approach to instruction in history. If the history-as-story is justified on the grounds of cognitive compatibility, history-as-interpretation is justified by an appeal to a higher standard of historical literacy. The contrast between what is expected at the middle grades and what is expected at college is quite dramatic. Although not even college history courses typically challenge students with

documents in the introductory courses, there is a clear expectation that students should gain competence in interpretation, explanation, and argument with whatever texts make up instruction. The lack of such emphasis in grade school history and even in high school history stands in contrast.

The question is whether such a division of cognitive labor is warranted by cognitive constraints, instructional resources, or both. The answer may prove to be complex. We suspect there is nothing in principle to prevent a document-based interpretation approach to history even in middle school—that is, so long as specific targets of breadth of coverage are relaxed. On the other hand, there are some constraints to deal with, including the demands that documents make on reading skill, domain knowledge, and a nondomain reasoning skill, based on something like “rules of evidence,” that has quite general application in education.

Although we think there is a fairly general skill that underpins the ability to appreciate the role of evidence across domains, we suspect it is a skill fostered initially in some specific domain. Of course, domain-specificity versus generality is a complex issue, and we leave it at that—except to suggest that there is little specifically historical about realizing that evidence counts and that some kinds count more than others. How to explain such an ability short of an appeal to general scholarship-based education is a challenge. Whatever it is, there is no reason not to expect our history instruction to include it at least in secondary education. The obvious strategy is to take advantage of the superior position of story learning and to supplement it with complexity as early as possible.

#### ACKNOWLEDGMENTS

The research reported in this chapter was primarily supported by a grant from the Mellon Foundation. Additional funding was also provided by the Office of Educational Research and Improvement (OERI), United States Department of Education to the Center for Student Learning, Learning Research and Development Center. The opinions expressed do not necessarily reflect the position or policy of the Mellon Foundation or OERI and no official endorsement should be inferred.

#### REFERENCES

- Black, J. B., & Bower, G. H. (1979). Episodes as chunks in narrative memory. *Journal of Verbal Learning and Verbal Behavior*, 18, 309-318.
- Britt, M. A., Rouet, J.-F., Georgi, M. C., & Perfetti, C. A. (1994). Learning from history texts: From causal analysis to argument analysis. In G. Leinhardt, I. L. Beck, & C. Stainton (Eds.), *Teaching and learning in history* (pp. 47-84). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brown J. I., Bennett, J. M., & Hanna, G. (1981). *The Nelson Denny Reading Test*. Chicago: The Riverside Publishing Company.
- Carretero, M., Asensio, M., & Pozo, J. I. (1991). Cognitive development, historical time representation and causal explanations in adolescence. In M. Carretero, M. Pope, R. Simons, & J. I. Pozo (Eds.), *Learning and instruction: European research in an international context* (Vol. 3, pp 27-48). Oxford: Pergamon.
- The Center for Strategic Studies. (1967). *Panama: Canal issues and treaty talks*. Washington, DC: Georgetown University Press.
- Crane, P. M. (1978). *Surrender in Panama: The case against the treaty*. Ottawa, IL: Green Hill Publishers.
- LaFeber, W. (1978). *The Panama Canal: The crisis in historical perspective*. New York: Oxford University Press.
- Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 9, 111-151.
- Omanson, R. C. (1982). The relation between centrality and story category variation. *Journal of Verbal Learning and Verbal Behavior*, 21, 326-337.
- Perfetti, C. A., Britt, M. A., & Georgi, M. C. (in press). *Learning and reasoning from “real” texts: Case studies of history text learning*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rouet, J. F., Britt, M. A., Mason, R. A., & Perfetti, C. A. (in preparation). *Using multiple sources of evidence to study historical controversies*.
- Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 211-236). New York: Academic Press.
- Shafer, R. J. (1974). *A guide to historical method*. Homewood, IL: The Dorsey Press.
- Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing* (pp. 53-120). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 9, 77-110.
- Trabasso, T., Secco, T., & van den Broek, P. (1984). Causal cohesion and story coherence. In H. Mandl, N. L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 83-111). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, 24, 612-630.
- van den Broek, P., & Trabasso, T. (1986). Causal network versus goal hierarchies in summarizing text. *Discourse Processes*, 9, 1-15.
- 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.