



---

What Happens When Students Read Multiple Source Documents in History?

Author(s): Steven A. Stahl, Cynthia R. Hynd, Bruce K. Britton, Mary M. McNish and Dennis Bosquet

Source: *Reading Research Quarterly*, Vol. 31, No. 4 (Oct. - Nov. - Dec., 1996), pp. 430-456

Published by: [International Reading Association](#)

Stable URL: <http://www.jstor.org/stable/748185>

Accessed: 18/08/2013 14:39

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



*International Reading Association* is collaborating with JSTOR to digitize, preserve and extend access to *Reading Research Quarterly*.

<http://www.jstor.org>

Steven A. Stahl

Cynthia R. Hynd

Bruce K. Britton

Mary M. McNish

*University of Georgia, Athens, USA*

Dennis Bosquet

*Clarke County School District, Athens, Georgia, USA*

# What happens when students read multiple source documents in history?

---

After many years of comparative neglect, the study of history has received renewed attention by cognitive psychologists (Wineburg, 1991a, 1991b). Cognitive analyses of history learning have appeared in symposiums presented at major national meetings, as well as in books devoted to the subject (e.g., Leinhardt, Beck, & Stainton, 1994; Perfetti, Britt, & Georgi, 1995) and a special issue of *Educational Psychologist* (Wineburg, 1994).

This renewed attention may presage an interest in new methods of presenting historical content. The traditional means of teaching history was to rely heavily, if not exclusively, on the textbook as a means of conveying information. In 1982, a survey found that roughly 90% of all social studies teachers use a textbook in their class (Patrick & Hawke, 1982). Approximately half of all teachers in that survey reported relying on just one text, with that text being reported as the major determinant of the content of their curriculum.

Currently, the single text approach to history learning and the model of learning upon which it is based are being challenged by those who espouse constructivist views of knowledge acquisition (e.g., Seixas, 1993) as well as those who espouse more traditional views of learning history (e.g., Ravitch, 1992). This article reports

on an attempt to examine an alternative approach to learning about historical events, using multiple source materials, and the processes used by students as they negotiate the information in the various documents.

## *Construction of meaning in history*

The textbook-based teacher can be caricatured as using a *transmission* model of learning, in which the information to be learned is contained in one vessel, the textbook, and transmitted to another vessel, the student's memory, via the teacher's lecture. Traditionally, many teachers have treated content area knowledge as Hirsch (1987) did, as a *basket of facts*, that must be gathered from text and lecture. These facts are stored in memory, the way information is stored in a computer database. As one history teacher quoted by Wineburg (1991b) put it, "History is the basic facts of what happened. What *did* happen. You don't ask how it happened. You just ask, 'What are the events?'" (p. 513, italics in original).

Such a transmission model is not supported by current views of the nature of knowledge and learning. More recent theories suggest that as information is learned, this information is not merely copied from one source to another but is transformed by the process of learning (Spiro, 1980). In this constructivist view of

### ***What happens when students read multiple source documents in history?***

SOME EDUCATORS (e.g., Ravitch, 1992) have suggested that students use multiple source documents to study history. Such documents could be primary sources, such as legislative bills or eyewitness accounts; secondary sources, such as editorials; or tertiary sources, such as textbooks. This study examined the processes used when high school students were presented documents about a controversial incident in U.S. history, the Tonkin Gulf Incident and its aftermath. These students were asked to read these either to describe or develop an opinion about the incident or the Senate action on the Tonkin Gulf Resolution. We were interested in (a) whether students could develop a rich, mental model of a historical event, (b) what they would do with the document information, (c) how the task influenced their processing of information, (d) how students integrated information across texts, and (e) whether students engaged in corroborating, sourcing, and contextualizing in evaluating historical

materials. We found that the mental models created by these students were more internally consistent after reading at least two documents, but did not become more consistent after that. When compared to knowledgeable readers, they failed to make any growth after a first reading. Examining their notes, we found that students tended to take literal notes, regardless of the final task, suggesting that they were using the initial readings to garner the facts about the incident or the resolution. If students were asked for a description, they tended to stay close to the text. If asked for an opinion, however, they tended to ignore the information in the texts they read, even though they may have taken copious notes. Our observations suggest that high school students may not be able to profit from multiple texts, especially those presenting conflicting opinions, without some specific instruction in integrating information from different texts.

### ***¿Qué sucede cuando los estudiantes leen documentos de diversas fuentes en historia?***

ALGUNOS EDUCADORES (por ej. Ravitch, 1992) han sugerido que los estudiantes usan documentos de diversas fuentes para estudiar historia. Estos documentos pueden ser fuentes primarias, como por ejemplo declaraciones del Congreso o relatos de testigos presenciales, fuentes secundarias, como por ejemplo editoriales, o fuentes terciarias, como los libros de texto. Este estudio examinó los procesos usados por estudiantes de escuela secundaria al presentárseles documentos acerca de un incidente controvertido de la historia de los Estados Unidos, el incidente del golfo de Tonkin y sus consecuencias. Se solicitó a los estudiantes que leyeran estos documentos para describir o desarrollar una opinión acerca del incidente o de la acción del Senado respecto de la Resolución del golfo de Tonkin. Nos interesaba investigar: (a) si los estudiantes podían desarrollar un modelo mental rico sobre un evento histórico, (b) qué harían con la información de los documentos, (c) cómo influenciaba la tarea el procesamiento de la información, (d) cómo integraban los estudiantes la información de distintos textos y (e) si los estudiantes procedían a corroborar, buscar las fuentes y contextualizar durante

la evaluación de los materiales históricos. Encontramos que los modelos mentales creados por estos estudiantes eran más consistentes internamente luego de la lectura de al menos dos documentos, pero su consistencia no aumentaba después de eso. Cuando se los comparó con lectores conocedores del tema, no lograron hacer ningún progreso después de una primera lectura. Al examinar sus notas, encontramos que los estudiantes tendieron a tomar nota literalmente, sin tener en cuenta la tarea final, lo que sugiere que usaron las lecturas iniciales para reunir el conjunto de los hechos acerca del incidente o la resolución. Cuando se les pidió una descripción, tendieron a permanecer muy ligados al texto. Cuando se les pidió una opinión, sin embargo, tendieron a ignorar la información de los textos, aún cuando hubieran tomado abundantes notas. Nuestras observaciones sugieren que los estudiantes de escuela secundaria parecen no poder beneficiarse con el uso de múltiples textos, especialmente aquellos que presentan opiniones conflictivas, si no se les da instrucción específica acerca de la integración de información de distintos textos.

### ***Was geschieht, wenn Schüler/innen unterschiedliche historische Quellen lesen?***

EINIGE PÄDAGOGEN (z.B. Ravitch, 1992) schlugen vor, daß Schüler/innen unterschiedliche historische Quellen zum Studium der Geschichte lesen sollten. Solche Dokumente können Primärquellen sein, z.B. Kongreßakte oder Augenzeugenberichte, Sekundärquellen, wie z.B. Zeitungsberichte, oder Tertiärquellen, wie z.B. Lehr- oder Geschichtsbücher. Diese Studie untersuchte die Vorgänge, als Highschool-Schüler/innen Dokumente über kontroverielle Ereignisse der amerikanischen Geschichte, z.B. den Vorfall am Tonkin-Golf und dessen Nachwirkungen, vorgelegt wurden. Die Schüler/innen wurden gebeten, diese Dokumente zu lesen und diese zu beschreiben, oder sich eine Meinung über den Vorfall oder die Maßnahmen des Senats in Form der Tonkin-Golf-Resolution zu bilden. Unser Interesse richtete sich darauf, ob a) die Schüler/innen eine reiche, geistige Vorstellung von dem historischen Ereignis entwickeln konnten; b) was sie mit den Informationsquellen machten; c) wie die Aufgabenstellung ihren Wissenserwerb beeinflusste; d) wie die Schüler/innen die Informationen der einzelnen Texte transferierten und e) ob die Schüler/innen sich in der Auswertung des historischen Materials engagierten, indem sie dieses bestätigten, auffindig machten und Sinnzusammenhänge herstellten. Wir fanden

heraus, daß das Geschichtsbild, das sich die Schüler/innen geschaffen hatten, gefestigter war, wenn mindestens zwei Dokumente gelesen wurden, aber mit mehr nicht besser wurde. Verglichen mit gutinformierten Lesern zeigten sie keinen Wissenszuwachs nach dem ersten Lesen. Bei der Untersuchung ihrer Bemerkungen fanden wir heraus, daß Schüler/innen dazu neigten, wortgetreue Bemerkungen zu machen, ohne auf das Lernziel zu achten, was den Schluß nahelegt, daß sie die ersten Leseindrücke verwendeten, um die Fakten über den Vorfall oder die Resolution im Gedächtnis zu speichern. Wenn die Schüler/innen um eine Beschreibung gefragt wurden, tendierten sie dazu, nahe am Text zu bleiben. Wenn sie um ihre Meinung gefragt wurden, neigten sie jedoch dazu, die im Text gelesenen Informationen zu negieren, auch wenn sie sich umfangreiche Notizen gemacht hatten. Unsere Beobachtungen legen den Schluß nahe, daß Highschool-Schüler/innen noch nicht imstande sind, von verschiedenen Informationsquellen zu profitieren, besonders wenn diese einander widersprechende Meinungen ergeben und wenn keine ausdrückliche Anweisung erfolgt, Informationen aus verschiedenen Texten miteinander zu kombinieren.

## 様々な歴史の資料を読む際に生徒に何が起ころのであろうか

ある教育者たち（例えば Ravitch, 1992）は歴史を勉強するにあたって様々な資料を活用することを生徒に勧めている。資料は議会の法案とか証人からの報告といった第一次資料であつてもよいし、論説といった第二次資料、或は教科書といった第三次資料であつてもよい。この研究は、アメリカ史において議論を呼ぶようなトンキン湾事件とその影響についての資料を読んだ時、高校生がどのようなプロセスを使うかを調査したものである。生徒たちはその資料を読んで、トンキン湾決議に関する事件や上院のとった行動についての意見を述べ、発展させるよう求められた。我々は (a) 生徒たちが史実に関して豊かな観念的モデルを発展させることができるかどうか, (b) 資料からの情報を使って何をするのか, (c) その作業が生徒たちの情報処理にどう影響するのか, (d) 様々なテキストから情報をどのようにまとめあげるのか, (e) 歴史の文献を評価する際に、裏づけをしたり、原因を調べたり、歴史的背景から流れを捉えようとするのかどうか、といった点に注目した。生徒の観念的

モデルは、少なくとも2つの資料を読んだ後は一貫したものであったが、それ以上になると一貫したものではなくなることがわかった。知識を多く持った読み手に比べると、最初に資料を読んだ後は、生徒たちに何の進歩も見られなかった。生徒たちのノート調べてみると、最終作業の段階であるにもかかわらず逐語的にノートを取る傾向があることがわかった。このことは生徒たちがその事件や決議についての事実を覚えておくのに最初に読んだ資料に頼っていることを示唆している。もし生徒に史実の描写をするよう求めたら、それは最初に読んだテキストの内容に似たものになる傾向があった。しかし意見を求めた場合には、たくさんノートを取ったにもかかわらず、テキストからの情報を無視する傾向があった。異なったテキストから情報をまとめるための特別な指導をしなかった場合には、特に議論を呼ぶような内容の多くのテキストからは高校生たちが恩恵を受けていない可能性があることを我々の観察は示唆している。

## *Que se passe-t-il quand les élèves lisent des documents de sources multiples?*

CERTAINS DIDACTICIENS (par exemple, Ravitch, 1992) ont suggéré que, pour étudier l'histoire, les élèves utilisent des documents de sources multiples. Ces documents pourraient être des sources primaires, comme des lois votées par le Congrès ou des comptes rendus de témoins oculaires, des sources secondaires, comme des éditoriaux, ou des sources tertiaires, comme des manuels. Cette recherche a analysé les processus mis en oeuvre quand des élèves de lycée reçoivent des documents relatifs à un incident controversé de l'histoire des États Unis, l'épisode du Golfe du Tonkin et ses suites. On a demandé aux élèves de lire ces documents et de décrire ou de développer une opinion à propos de l'incident ou de l'action du Sénat sur la Résolution du Golfe du Tonkin. Nous nous sommes demandé: (a) si les élèves étaient en mesure de développer un modèle mental riche d'un incident historique, (b) ce qu'ils feraient de l'information provenant du document, (c) comment la tâche influencerait sur leur traitement de l'information, (d) comment les élèves intégraient l'information en passant d'un texte à l'autre, et (e) si les élèves s'engageraient dans un travail demandant confirmation, re-

cours aux sources, contextualisation, lors de l'évaluation de matériaux historiques. Nous avons trouvé que les modèles mentaux créés par ces élèves ont plus de cohérence interne après la lecture d'au moins deux documents, mais ne gagnaient pas en cohérence en allant au-delà. Comparés aux lecteurs que l'on peut connaître, ils ne faisaient pas montre de plus progrès qu'après une première lecture. En examinant leurs notes, nous avons trouvé que les élèves avaient tendance à prendre davantage de notes littérales, sans tenir compte de la tâche finale, ce qui suggère qu'ils ont utilisé les lectures initiales pour engranger les faits relatifs à l'incident ou à la résolution. Si on demandait une description aux élèves, ils avaient tendance à rester près du texte. Toutefois, si on leur demandait une opinion, ils avaient tendance à négliger l'information contenue dans les textes qu'ils avaient lus, même s'ils avaient pris des notes copieuses. Nos observations suggèrent que des élèves de lycée peuvent ne pas tirer profit de textes multiples, en particulier quand ils présentent des opinions conflictuelles, s'ils n'ont pas reçu un enseignement spécifique relatif à l'intégration de l'information issue de textes multiples.



knowledge acquisition, new information can be retained in short-term memory through rote memorization or rehearsal, but this information is easily forgotten. This is evidenced by the often-experienced phenomenon of a student learning facts for a test and forgetting as soon as the test is through. For information to be learned and retained, it must be actively combined with previously learned information. The new learning is constructed from the new information and the old information into new knowledge, either through *assimilating* the new knowledge into already existing knowledge structures or *accommodating* the new information by creating new knowledge structures that would account for both the previously known and the new information (Rumelhart, 1980). Because every learner brings somewhat different knowledge and experience to the classroom, the knowledge that each learner retains is going to be somewhat different.

In this constructivist view of knowledge, the conveyance of content is more than merely ensuring that the students devote enough time and attention to memorizing the text or the teacher's lecture. Instead, the teacher must create the conditions that best allow the student to construct a mental model of the knowledge domain, incorporating into this mental model not only the information in the current curriculum, but also past knowledge.

The constructivist view of learning not only challenges the transmission model but also calls into question the relevance of those psychological models of learning based on the reading of a single text for examining the processes involved in learning history. Models such as Kintsch and van Dijk's (1978) may accurately describe how readers construct a propositional text base from the reading of a single text. However, texts that we read are understood in relation to other texts that we have previously read and other knowledge that we have acquired (Hartman, 1995). A psychological model of learning from texts, whether a single text or multiple texts, should include not only the text itself, but the reader's previous knowledge and how the student uses that knowledge in constructing a new mental model (see Kintsch, 1986).

#### *Content and disciplinary knowledge*

One goal of history instruction, then, should be for the learner to construct a well-articulated mental model of history, understanding the interconnections between various events and actors. Taking the topic of the present study, the origins of the Vietnam War, a student should have an understanding of the relations between the U.S. election of 1964, U.S. views of communism during that era, Lyndon Johnson, the Viet Cong, and the Gulf of Tonkin Resolution. These understandings should

be deep enough to understand why a possibly misunderstood incident, involving minor damage to two ships, could trigger a major conflagration. The mental model containing these understandings could be called *content knowledge* or knowledge about a particular domain (Stahl, Hynd, Glynn, & Carr, 1995).

Stahl et al. (1995) argue that, while content knowledge is important, it is not sufficient for the study of history. In addition, a person needs *disciplinary knowledge* or the ability to think like a historian, to evaluate materials and information in relation to their context and their source, and to integrate this information into a historical discourse (e.g., Greene, 1994).

Wineburg (1991a, 1991b) gave eight historians and eight high school seniors a series of historical texts about the Battle of Lexington and had them complete a variety of activities, including thinking aloud as they read, rating the trustworthiness of the documents, and evaluating the historical veracity of three paintings of the Battle. He noted that historians could be distinguished from students by their use of three processes:

- Corroboration, or comparing and contrasting documents with one another;
- Sourcing, or looking first at the source of the document before reading the text itself to consider how the bias of the source might have affected the content of the document; and
- Contextualization, or situating a text in a temporal and spatial context to consider how the time or place in which the document was written might have affected its content or the perspective taken.

The differences were not simply due to differences in content knowledge, since historians who did not know very much about the American Revolution still used the same reasoning processes in their think-alouds. Nor were the differences due to inability to detect bias. The college students in Perfetti, Britt, Rouet, Mason, and Georgi's (1993) study and the high school students in Stahl and Hynd's (1994) study were both able to detect bias in sources.

Instead, the differences between the students and the historians seem to be tied to differences in the way the historians and students viewed text. Wineburg (1991a, 1991b) inferred that students tended to view texts as repositories for facts, as bearers of information, as they might well have, given years of exposure to a transmission model of learning. For example, they tended to rate textbooks as more trustworthy than source documents, a finding replicated by Perfetti et al. (1993) and Stahl and Hynd (1994). Historians tend to view texts as speech acts, produced for a particular purpose by a particular person. To understand historical texts involves understanding both the person and the purpose, and to

get at the *truth* hidden within the texts involves comparing various perspectives, with an understanding of who produced the various texts and why. The students in Perfetti et al.'s (1993) study were able to grasp the basic story of the Panama Canal Treaty from documents describing the events leading up to the signing of the Treaty in 1903 but were less able to provide evidence about their stance on whether the treaty should have been signed.

### *Multiple texts and history learning*

A number of educators have suggested that the single classroom text be supplemented with or supplanted by multiple original source materials (e.g., Perfetti et al., 1993; Spoehr & Spoehr, 1994; Wineburg, 1991a). Providing students with multiple perspectives on a particular event can aid them in constructing a richer and more detailed mental model of that event, thus enhancing content knowledge. Spiro, Coulson, Feltovich, and Anderson (1994) likened the use of multiple perspectives to crisscrossing a conceptual landscape and suggested that seeing an event through different perspectives is necessary to create a rich understanding of an event or concept. This use of original material forces students to construct links across information presented in different texts, and this information and the links connecting the different sources are remembered better if students make their own constructions rather than relying on the constructions of a textbook author or teacher (Spoehr & Spoehr, 1994). The links based on this crisscrossing create a rich mental model, or what we are calling content knowledge.

The use of multiple texts can also increase students' disciplinary knowledge. If we consider the tasks that Wineburg (1991a, 1991b) found to distinguish between historians and high school students—corroboration, sourcing, and contextualization—to be at least part of the thought processes used by historians, they can be activated only by providing opportunities to compare and contrast different source materials with different and independent viewpoints. The single, omniscient view of a textbook cannot easily be used to develop disciplinary knowledge, since there is nothing to which the student can compare the information, and thus the student is usually unable to examine the bias of the textbook or the effects of the time and place in which it was written and to compare it to other sources. (However, McKeown, Beck, and Worthy, 1993, have developed procedures to elicit this information from critical examination of a single text.)

If conflicting information is presented in these texts, however, the conflict may impede learning. Perry (1970) examined the development of thought among

male college students and found evidence for development from a stance of looking for a single right answer to an understanding that knowledge is relative, depending on one's perspective, to the melding of information from different perspectives. Belenky, Clinchy, Goldberger, and Tarule (1986) replicated Perry's study with a broad range of women. They found that stances of knowledge can move from a belief that knowledge is received, or is transmitted from someone else, to a subjective stance, in which knowledge is seen as subjective and relative, to a procedural stance, in which rational processes are seen as a way to break through the subjectivity, to a stance they call *constructed*, in which knowledge is constructed through both rational processes and the acknowledgment of other perspectives.

It is this last stance that we expect students to take when looking at multiple documents, but it is one that is typically achieved in the later college years or graduate school, after exposure to the more open-ended discussions typical of college classrooms. It may be unreasonable to expect high school students, who tend to be exposed to more lecture and recitation, to think like this, at least not without some greater instruction in how to do it and some expectation that they engage in this kind of thought.

Despite the theoretical sense that multiple sources can enhance learning, there is very little information on *how* readers synthesize information across texts. Crafton (1983) had 11th graders read two science texts, either both on the same topic or on two different topics. She found that those who read two texts on the same topic comprehended significantly more about the topic than those who read texts on the different topics. Further, the students who had prior knowledge (from reading the first, relevant text) were able to focus on larger segments of text during a verbalization task and made more inferences, suggesting that they were better at integrating material in the text.

Spivey and King (1989) examined how 6th-, 8th-, and 10th-grade writers synthesized information across different encyclopedias. They found that older and more able students tended to be more adept at using information that was repeated in all three texts read and was presumed to be more important, better at reorganizing information from the different sources into a coherent whole, and more aware of the needs of their audience.

Greene (1994) gave college juniors and seniors a task either to write a report or solve a problem in history. He found that students given the problem-based task were more likely to bring their previous knowledge into their essays, to see the task as one of evaluation of the information in the articles, and to draw upon different kinds of information than the students who were given

the report writing task. The students who were asked to write a report had difficulty doing so, because they tended not to set their ideas in a context and justify the issues they chose to write about.

The purpose of this study was to examine the processes and outcomes of reading multiple original source materials. The materials relate to the Gulf of Tonkin Incident and the resultant Gulf of Tonkin Resolution passed by the U.S. Congress that eventually began the Vietnam War. We were specifically interested in the following questions: (a) Could students develop a rich mental model of a historical event? (b) What did students do with the document information? (c) Did the task students were given influence their processing of information? (d) How did students integrate information across texts to form a coherent essay? and (e) Did students engage in corroborating, sourcing, and contextualizing in evaluating historical materials?

The first two questions asked whether students can learn from multiple text documents and how they process these documents to aid their learning. The third question dealt with the effects of task on learning. We used two different tasks, having students write either a description or an opinion of either the incident itself or the events leading up to the Senate resolution. Some studies have found that encoding tasks influence processing (e.g., Reynolds, Trathen, Sawyer, & Shepard, 1993). We hoped to find processes that were used more often in one task than in another, thus illuminating how students process multiple text information. The fourth question dealt with how students put information from multiple documents together. The final question was intended to see whether the multiple text task induced students to use the operations found by Wineberg (1991a) to distinguish historians from high school students in their reading of original source materials.

## Method

### Participants

The participants were 44 students in two classes of 10th-grade Advanced Placement U.S. History taught by a single teacher. The school was one of two high schools in a small southern U.S. university town, drawing from a wide range of socioeconomic states. Approximately one quarter of the students participating were African American and the others were of European American origin. These students were enrolled in U.S. History so that they would be exempt from a required course in college. Therefore, only high-achieving students who were expecting to attend college were taking the class.

The topics used in this study, the Gulf of Tonkin Incident and the resultant Gulf of Tonkin Resolution, were on the Advanced Placement exam, but students had not yet studied the incident and resolution in class.

Of these 44 students, 18 students worked in groups. Students were randomly assigned to groups. These groups were used for a study of students' interactions around texts (Hynd, Stahl, Britton, & McNish, 1996). As will be discussed later, these students were included in the analysis of mental models but not in the analysis of notetaking, since their notetaking was not independent. Six students' notes and final products were not analyzable for a variety of reasons, such as not following instructions. Only 20 of the remaining students produced notes, and these were used in the analysis of notetaking. Sixteen students produced analyzable final products; four others produced notes but not analyzable final products. The analyses reflect these differing numbers.

This study was conducted in January, before the students' history teacher began preparing these classes for the document-based question on the Advanced Placement examination. Thus, these students had not yet had direct instruction in how to integrate information across documents. Instead, the teacher used primarily a lecture mode, believing that "History is a story." The teacher was widely regarded as an excellent history teacher, with high percentages of students passing the Advanced Placement examination.

### Materials

*Background questionnaire.* The background questionnaire asked students their political affiliation and their parents' political affiliation. It asked them whether they were liberal, conservative, or moderate on matters of national defense, the economy, and social issues. It also asked them about their stance on certain current affairs and issues debated in public forum, asked them to rate their knowledge of the Vietnam War, and asked them to describe their feelings about what was important to study in history. Finally, the questionnaire asked students to rate the U.S. Congress, U.S. newspapers, the President of the United States, army generals, historians, and history textbooks for their trustworthiness.

*Prior knowledge writing task.* In addition to the questions on the background questionnaire about students' knowledge of the Vietnam War, we included two additional measures of participants' prior knowledge. The first was an open-ended writing task. We asked the students to "Please write down everything you know about your assigned topic. If you are not sure, then write down what you think you know." This task was scored for number of accurate knowledge statements and ex-



**Table 1** Texts used in the study

Name of text	Brief synopsis
<b>Gulf of Tonkin Incidents</b>	
Text of telegram written by North Vietnamese to protest the mission of the U.S.S. Maddox	North Vietnamese call upon South Vietnamese to stop aggressive raids. Explain that U.S. ships were seen as aiding those raids.
<i>Another Gulf, Another Blip On the Screen</i>	An eyewitness account by James Stockdale, America's highest ranking prisoner of war during the Vietnam War. He was flying over the Gulf of Tonkin and did not see a torpedo attack.
<i>The Pentagon Papers</i>	The official history of the event written shortly following the incident. It said that the Gulf of Tonkin Incidents were clearly aggressive acts on the part of the North Vietnamese.
<i>Secrets of the Vietnam War</i>	An excerpt from a privately published book written by a retired Army colonel, claiming that the North Vietnamese were primarily responsible for the incident and that the U.S. was not overly aggressive.
"The Tonkin Gulf Crisis"	An editorial analysis claiming that the Gulf of Tonkin Incidents were largely trumped up by the U.S. as a way to widen the Vietnam conflict.
<i>Vietnam: A History</i> by Stanley Karnow	An in-depth historical analysis that explained the events leading up to the U.S.'s interpretation of the Gulf of Tonkin Incidents that resulted in the Gulf of Tonkin Resolution and ultimate widening of the Vietnam conflict.
<b>Gulf of Tonkin Resolution</b>	
The Tonkin Gulf Resolution	A copy of the actual resolution as voted on by Congress.
"The Vote that Congress Can't Forget"	A newspaper article that described members of Congress' retrospective thoughts about the Gulf of Tonkin Resolution as they voted to allow the President to attack the Persian Gulf. Most of the congressmen said they regretted voting for the Resolution and that they didn't realize the effect it would have.
<i>The Vietnam Hearings</i>	Text taken from the Congressional Record describing Senator Fulbright's celebrated hearings where U.S. involvement in Vietnam was discredited. Dean Rusk, Senator Fulbright, and others were attempting to decide if spending more money on Vietnam was justified.
<i>As I Saw It</i>	An excerpt from Dean Rusk's autobiography that attempted to exonerate both the President and himself from accusations that they had acted hastily in their decision to escalate the war after the Tonkin Gulf Incidents.
<i>Vietnam: A History</i>	Same text as for Gulf of Tonkin Incidents.

pressed as a percentage of accurate to total number of statements. We used this measure only to interpret individual differences in notetaking.

*Gulf of Tonkin relationships task.* The final measure of prior knowledge was a relationships task, used by Britton and Gulgoz (1991). In this task, students were asked to rate the strength of the relationship between all possible pairs of 10 key words or phrases—the Gulf of Tonkin Resolution, the Gulf of Tonkin, North Vietnam, South Vietnam, U.S. Congress, President Johnson, Vietnam War, U.S. Forces, Defense, and Aggression. This task was given before any of the reading, as a pretest, and after each reading was completed, as a measure of growth as a result of that reading.

Students rated the pairs on a 1 to 6 scale with 1 being *not very related* and 6 being *strongly related*. The purpose for this task was to determine the coherence (or harmony) of students' mental models before they read texts and as a result of reading. We expected students to have a more coherent way of rating the pairs after hav-

ing read texts, and we were interested in whether students would evidence steady growth in coherence or whether one or more texts were responsible for more coherent rating than other texts.

The measuring of harmony is described in Britton and Gulgoz (1991) and is expressed in the form of a decimal. For example, a harmony value of 1.00 would mean that an individual had rated the relationships between pairs in such a way that there were no conflicts between ideas. A harmony value of .50, however, would mean that there was a moderate degree of contradiction in the way the pairs were rated. If a student rated the Gulf of Tonkin Resolution and President Johnson as strongly related, and Aggression and President Johnson as strongly related, but Aggression and the Gulf of Tonkin Resolution as not very related, the person's mental model of those three items would be considered in-harmonious. As students learn, they sort out internal contradictions between different ideas and begin to generate stable relationships between ideas. A good mental



model would have high internal consistency; low internal consistency would indicate some confusion.

*Texts.* Students read multiple texts presented on Hypercard stacks on Macintosh computers. Computers were used to provide an orderly environment for the exploration of the texts. The computers were also used to provide online help, such as identification of key people and short biographical information about the author.

Before reading any of the texts, students viewed a map showing Vietnam and the Tonkin Gulf and read a one and one half card background information statement that described in objective terms the Gulf of Tonkin Incidents and resultant resolution. This background text provided an overview of the Vietnam War and the Gulf of Tonkin Incidents' role in that war. It was written to be neutral in terms of the two questions that were posed, providing just the facts that were verified in all selections. The text is reproduced in Appendix A.

After they had read the background information, students were directed to a screen with two buttons, one directing them to documents concerning the incident and one directing them to documents concerning the resolution. They were to refer to their assignment sheets to see which question they were supposed to address. Clicking on a button led to a menu that presented the titles of their assigned readings. Students could browse the readings before deciding which ones to actually read. Because we wanted this task to be as natural as possible, we did not control the order of readings.

Six readings were about the Gulf of Tonkin Incident and five were about the Gulf of Tonkin Resolution. We chose the topics because they have been hotly debated by historians and politicians. Different interpretations of the event and resolution exist, allowing us to choose texts that represented several perspectives. It was the integration of various perspectives that the researchers wished to study. The texts chosen represented a blend of primary to tertiary sources that were as evenly distributed as possible in terms of their stances. The texts are listed in Table 1.

Because part of the focus of the study was to see which documents students would choose, we included texts to represent a span of possible documents that might be used to study this incident. About one half of the texts we judged to be prowar and half antiwar. We included histories (*Vietnam: A History* by Stanley Karnow and *The Pentagon Papers*), newspaper opinion papers, autobiographies of participants (Commander James Stockdale and Dean Rusk), original documents (the text of the Gulf of Tonkin Resolution and the telegram sent from the North Vietnamese protesting the earlier raids in the Gulf) and secondary sources. We wanted to make sure that all viewpoints were represent-

ed and that students had a choice of different genres and styles of documents. We also used the information from a pilot study (Stahl & Hynd, 1994), choosing texts students rated as highly believable and those rated less believable.

As students read the texts, they had several options for help. For one, students could find out information about the author of the text. This information was basic, including the source of the document (newspaper, book, etc.) and the author's position (writer, former army colonel, Secretary of State, etc.). Further, if they put the cursor on selected vocabulary (mostly people and organizations), background information appeared on the screen. Students could also search for a keyword by choosing the FIND button and typing in the word for which they were searching. They could take notes on the computer if they wished (although only three tried and all decided against it), and, finally, they could move freely backwards and forwards within and across texts.

*Notetaking option.* While students read each text, they could take notes, if they wished, on paper provided in their packet. Although they were not required to take notes, researchers and written directions explained to them that they could use these notes for the final writing task but could not refer to the actual readings.

*Evaluation sheet.* We asked students to answer these questions about each text: (a) What do you feel the author's purpose was in writing this? (b) How useful would this be to help you learn about the origins of the Vietnam War? (rated from *Not Very*, 1, to *Very*, 6); (c) How unbiased do you think this account is? (also rated from 1, *Not Very*, to 6, *Very*); (d) How difficult was this text to read? (1 to 6 rating); and (e) How interesting was this text? (1 to 6 rating). Students answered these questions before engaging in the free recall task.

*Free recall task.* This task directed students to "Write down all the information you can remember from reading this text. Do not refer to your notes or the text before or during writing. Be as complete as possible." Students engaged in this activity after reading each of the texts.

*Final writing task.* We gave students a final writing task that mirrored their assigned purposes for reading. If students had been assigned to read in order to form an opinion about either the Gulf of Tonkin Incident or the Gulf of Tonkin Resolution, they were asked to write about their opinions. If students had been assigned to read in order to describe the Gulf of Tonkin Incident or the Gulf of Tonkin Resolution, they were asked to write a description. We gave students 30 minutes in class to complete this activity. All students finished before the 30 minutes were up.

### *Procedure*

Students who participated in the study met for three days in the computer room that was part of their school library. The librarian had equipped the room with 15 Macintosh SE30 desktop computers. As students came into the room on the first day of the study, the researchers handed each one a folder that included questionnaires, written directions for completing the study, and an introduction that assigned them to a topic and a purpose for reading. Researchers distributed these folders in a stratified fashion to students upon entry, resulting in random assignment. Four conditions represented two purposes and two topics. We asked students to read either to (a) form an opinion about the topic or (b) be able to describe the topic. We also told students that they would engage in a writing task related to their purpose for reading at the end of the study. Finally, we asked students to read texts about (a) the Gulf of Tonkin Incidents or (b) the Gulf of Tonkin Resolution. For these topics, students could choose six and five texts, respectively.

Students filled out the background questionnaires, read the introduction that explained their task, wrote down everything they already knew about the topic they were assigned, completed the Gulf of Tonkin relationships task used as a pretest, and read the instructions for accessing the texts from the computer screen while one of us explained those directions out loud and answered questions. All students were familiar with the computers and with using the mouse so that they did not need basic directions for managing the computer. The researcher told students that they could read the texts in any order they wished, and that they could take notes if they wished. After they had completed reading each text they were to write a free recall without looking back to the text they had just read, complete the Gulf of Tonkin relationships task, and fill out a questionnaire about the text.

We did not allow students to look back because we wanted to assess ongoing learning, rather than the strategic use of text. After completing those tasks, students could then proceed to their next chosen text. Students started reading on the first day of the experiment, read through the 50-minute period on the second day, and stopped reading on the third day, approximately 30 minutes before the end of the period.

After students stopped reading, we told them to read the directions for their writing task and to follow those directions. The directions asked students to state their opinion about either the Gulf of Tonkin Incident or the Gulf of Tonkin Resolution, or they asked students to describe the Gulf of Tonkin Incident or the Gulf of Tonkin Resolution. We allowed them to consult their notes if they wished but did not allow them to return to

the actual texts on the computer screen while they were writing.

*Analysis of notes and final products.* Because we were interested in identifying processes that students used as they read each text and then formed an essay incorporating some or all of the texts they had read, we developed a format for recording the notes, text, and idea units from the essay so that their correspondences could easily be seen. A sample can be found in Appendix B. We divided pages into three columns, one for the text, one for the notes, and one for the essay. In the middle column, we wrote down the notes (in idea units), in the order in which they were taken. In the left-hand column, we recorded the section of the corresponding text. Although our judgment was sometimes needed to determine the textual basis for the notes, this task was relatively easy to perform because students generally took notes in the same linear order in which they read the text. Further, the majority of their notes were paraphrases or copying of the text. We also recorded idea units from the free recalls in this column, using the same procedures. The free recalls were clearly marked as such so that they would not be analyzed as notes.

In the right-hand column, we recorded idea units from the final essay next to corresponding notes or text. Because the essay was an incorporation of several different texts, sometimes these idea units were recorded in several different places. If no corresponding note or text was found, we placed the idea unit at the end of the third column. Again, we used judgment in deciding whether or not an idea unit represented an idea taken from notes or text. We tried to be inclusive; that is, if there was a possibility that students may have had a certain text in mind when they made the statement, we placed it accordingly.

After each student's notes and essay had been recorded in this manner, three researchers read all protocols. We divided each text into idea units, which were defined as single pieces of information. Usually, there was one idea unit per sentence. However, some sentences contained more than one idea unit, and, of course, students did not always write complete sentences. We had a 95% agreement in breaking protocols into idea units.

We then created a system for categorizing idea units for the notes, free recalls, and essays. The system was not developed with an a priori set of categories. Instead, the categories emerged from the data (Glaser & Strauss, 1967). We were also concerned about the reliability of the categories and the replicability of these categories. To develop this system, three of the five authors read through the protocols and discussed what we felt they revealed about what the students were doing as

they were taking notes and creating their final products. We attempted to codify these processes into a system that could be reliably used to categorize the processes we found. We went through a number of different systems before we found an approach that we could apply with greater than 90% interrater reliability and that seemed to produce useful interpretations of the data. This categorization system is described below.

We classified each idea unit as (a) copying, (b) paraphrasing, (c) reducing, (d) making a gist, (e) evaluating, or (f) distortion/misreading. We classified an idea unit as *copying* if it was word for word or nearly word for word with close synonym replacement or minimal reordering. An example of copying is when the text said, "Gulf of Tonkin Resolution passed by the Congress on 7 August 1964," and the notes said, "Gulf of Tonkin Resolution—August 7, 1964 passed by Congress."

*Paraphrasing* was a more radical replacement of words that included within sentence reduction or elaboration. An example of a paraphrase is when the text said, "Vietnamese coastal targets—this time the Rhon River Estuary and the Vinh Sonh radar installation, which were bombarded on the night of 3 August," and the notes said, "On the night of August 3, Vietnamese coastal targets were bombarded."

We described *reducing* as a summarization process across two or more sentences, so that the writing contained markedly fewer words and details than the original. An example of reducing is when the text said, "At 1940 hours, 4 August 1964 (Tonkin Gulf time) while 'proceeding S.E. at best speed,' Task Group 72.1 (Maddox and Turner Joy) radioed 'RCVD INFO' indicating attack by PGM P-4 imminent," and later "Just before this, one of the PT boats launched a torpedo, which was later reported as seen passing about 300 feet off the port beam, from aft to forward, of the C. Turner Joy." The notes merely said, "On 4 Aug 1964, the Maddox & Turner Joy were attacked by PT boats, who launched a torpedo."

We described *making a gist* as radical reduction in which nouns were replaced with superordinates or more general terms. We noted that gists were often blanket statements that were more topical in nature than reductions, such as, "the text was about the resolution," or made blanket interpretations of details, such as "LBJ uses attack to get control of Congress," when several paragraphs had described the President's dealings with Congress in getting the resolution passed.

We described *evaluating* as stating an opinion about the ideas in the text that were not merely the copied opinion of authors or the opinion of people the authors described. For example, we classified the statement "Johnson was an idiot," as an evaluation.

We described *distortion/misreading* as being either inaccurate textual interpretations or statements that, although not evaluative, were simply not found in the text. An example of a misreading is when the notes said, "South Vietnam mistakes U.S. for South Vietnamese ship," but the text said that the North Vietnamese mistook the Maddox for a South Vietnamese vessel.

As noted earlier, this coding system was developed after much discussion among three of the five authors. After the system was developed, the researchers reached 92% agreement after coding notes on 5 of the 20 protocols. From that point, the researchers coded the remaining notes and free recalls separately.

*Analysis of final essay.* We read the final essays and coded idea units as coming from a single text or two or more texts. If a significant number of statements had come from two or more texts, we assumed that students were either integrating ideas across texts, or paying attention to information that was repeated across texts. We also coded each idea unit as being copied, a paraphrase, a reduction, a gist, an evaluation, or a misreading, as we did with the notes and free recalls. The purpose of this categorization process was to analyze what processes students were using to form a coherent essay.

In addition, we noted the order of statements in relation to the order of the texts they read and performed Kendall's Coefficient of Concordance ( $W$ ) to obtain a measure of the overall agreement in order. This coefficient helped us decide whether students were radically restructuring ideas or merely reporting them in the form in which they were first perceived. A low  $W$  indicated that students were reordering from the texts in the final product. A high  $W$  indicated that they were generally preserving information in the order that they had read it.

Finally, we calculated a ratio of information found in the text to that which could not be found in the text. This ratio would reveal whether students were sticking to the task of describing or stating their opinions, as they were assigned. It would also reveal whether students who were asked to state opinions would back up these opinions with factual information.

*Sourcing, corroboration, and contextualization.* Wineburg (1991a, 1991b) observed that historians used (a) sourcing (looking first at the source of the document before reading the text itself to consider how the bias of the source might have affected the content of the document), (b) corroboration (comparing and contrasting documents with one another), and (c) contextualization (situating a text in a temporal and spatial context to consider how the time or place in which the document was written might have affected its content or the perspective taken) when thinking about information in the texts they read, while students used these to a lesser degree, if at all.



**Table 2** Harmony levels

	Read 1 text	Read 2 texts	Read 3 texts
Pretest	.67	.67	.71
Text 1	.70		
Text 2		.76*	
Text 3			.81*

\* Probability of difference between last reading and pretest,  $p < .05$ . Different numbers of subjects read different numbers of texts.

**Table 3** Harmony after reading specific texts

Text read	Resulting harmony
Pretest	.67
<i>The Pentagon Papers</i>	.72
<i>Secrets of the Vietnam War</i>	.71
"The Tonkin Gulf Crisis"	.75
<i>Another Gulf</i>	.78
Text of telegram	.71
<i>Vietnam: A History</i>	.81*
<i>The Vietnam Hearings</i>	.77
<i>As I Saw It</i>	.79
"The Vote that Congress Can't Forget"	.75
The Gulf of Tonkin Resolution	.74

\* $p < .05$

In this study, we looked for instances of sourcing, corroboration, and contextualization in the notes, in the free recalls, and in the final essays. Sourcing was indicated by an explicit reference to the author or source, corroboration was indicated by a reference to another text in the series, and contextualization was indicated by an explicit reference to the time that the article was written. Because we counted only explicit references in the text, this may have underestimated the amount of sourcing, corroboration, and contextualization that the students were actually doing.

## Results and discussion

### *Can students develop a rich mental model of a historical event?*

The data from the relationships task were used to track how students developed a mental model of the events surrounding the Gulf of Tonkin Incident and the Gulf of Tonkin Resolution. We took two approaches to examining the development of a mental model. First, we examined the growth of harmony, which we used as a proxy for the internal consistency of the mental model

developed by the students. Second, we compared the structures generated by the students to those of experts in order to trace the growth of students' mental models toward those held by experts.

**Harmony.** The harmony ratings (using Kintsch's 1986 system [cited in Britton & Gulgoz, 1991] of calculating those ratings) are shown in Tables 2 and 3. This analysis included all 44 students. Some students worked alone and some in groups. We analyzed these two sets of students separately. The results were indistinguishable, so the two sets were combined for the analysis presented here.

As noted on Table 2, there was a significant growth in harmony from the pretest to the second reading, and from the pretest to the third reading. The growth in harmony after the first reading was not statistically reliable. Comparing the growth in harmony after each reading, only the difference in harmony between the first and second reading was statistically significant. There was a further small increase in harmony after a third reading. This difference was low (.79 vs. .81). This finding suggests that a student needs to read at least two different texts to develop a coherent mental model and that the majority of the growth occurs with two readings, but not much occurs after that.

To examine the effects of individual texts on the growth of harmony, we examined the gains in harmony from their prereading ratings after students read each of the texts (see Table 3). Of the 10 texts we used, only the section from the history text, *Vietnam: A History*, produced a significant gain in harmony by itself. This might be expected, since it was the longest and most detailed text we used.

**Expert ratings.** Another way to examine mental models is to compare the structures generated by the students with the structures generated by experts. We assume that knowledge consists of knowledge of relations among concepts, and that, as a person's knowledge grows, her/his knowledge of the relations among concepts will resemble that of experts. We used three expert raters to generate structures, using the same terms and procedures that we used with the students.

The first rater was the students' high school teacher, an experienced history teacher. The second rater was an amateur military history buff who read extensively about the war in Vietnam. The third rater was one of the authors of this study, who had majored in history, taken graduate level courses, and read the documents thoroughly and responded to the task based on her reading of the texts.

We chose these raters rather than studied experts on the Vietnam War because they represented the expertise that we wanted our students to have. There was not



**Table 4** Correlations with expert raters

	Correlation with average of three experts	<i>r</i> with Expert 1	<i>r</i> with Expert 2	<i>r</i> with Expert 3
Pretest	.26	.21	.15	.29
After text 1	.42*	.33*	.26*	.41*
After text 2	.36*	.28	.23*	.37*
After text 3	.43*	.33*	.30*	.41*
After text 4	.38	.33*	.23	.35

\* Correlation significantly different from pretest  $p < .05$ .

enough information in the texts to allow the students to obtain as full a representation of the events in Vietnam as a scholar would. Because these tasks focused on a small incident embedded in a larger context, it would be unrealistic to compare the knowledge obtained from these readings to that of scholars who were immersed in the larger context. The level of expertise that our raters had was about that which could be reasonably expected on this task.

All three experts tended to cluster the terms around two axes, one separating the terms Aggression and Defense, and the other roughly separating terms into domestic (U.S. Congress, U.S. Forces, etc.) and foreign (N. Vietnam, S. Vietnam, Gulf of Tonkin, etc.). Experts tended to have a strong separation between Aggression and Defense and clustered the other terms in the middle, roughly equidistant between these two poles.

A gain in knowledge would be evidenced by an increase in the students' correlation of their mental structure with that of the experts. Examining these correlations, shown in Table 4, there was a significant growth in knowledge after the first reading, but no significant gain subsequently. The initial correlations between the students and the individual experts ranged from .15 to .29, and initial correlation between the students and the composite was .26. These are small and not statistically reliable, suggesting that the students' initial knowledge was low and essentially random. The gain to .42, a moderate correlation, suggests that students learned some of the initial relationships after a single reading. Since subsequent readings tended to view the same facts from different perspectives, it is not surprising that there was little gain from these readings. There is also some evidence, as will be discussed below, that students read the first reading more closely than subsequent readings.

In contrast with the experts, students tended to cluster N. Vietnam, Gulf of Tonkin, and the Gulf of Tonkin Resolution with Aggression, and S. Vietnam, U.S. Force, and U.S. Congress with Defense. This may reflect

a different world view than our experts, who all lived through the Vietnam era. These students tended to see the U.S. and its allies in a positive light, and its enemies in a negative light. In contrast, the experts tended to view both sides in a more balanced regard, as neither side being more defensive or aggressive than the other.

One explanation for the lack of growth after a first reading may be in the nature of the texts and the task. We deliberately chose texts that contradict each other. It could be that students would read a first text to get the basic facts. Beyond these basic facts, which were accepted by all authors, were the different interpretations. Students may have ignored these interpretations and thus did not construct an increasingly complex mental model that might get closer to that of the experts. This supposition needs to be tested in a further study.

*Background information.* We gave the students an extensive background questionnaire, asking them information such as their political orientation, their parents' political orientation, their views about current events, their views of the reliability of various people and institutions such as Congress, the President, historians, and so on. We found no relation between any of our background variables and students' responses to these measures or any of the other measures. As will be discussed later, we did find that the students with moderate to high knowledge of the Vietnam War took different types of notes than other students.

#### *What do students do with the document information?*

As noted previously, we attempted to follow the flow of ideas from each document through the notes to inclusion in the final product. This model suggests that students initially selected ideas from the text as they were reading, deciding which ideas were important and which were not. They may have made a note of a selected idea, either copying it, paraphrasing, reducing two or three sentences, or reducing a paragraph or more into a single gist statement. They may also have noted an opin-

ion or reaction to information in the text. In producing the final product, they used similar operations, with ideas from a single text or ideas combined or repeated from multiple texts. These analyses used only those 20 students who worked individually and availed themselves of the notetaking option.

*Choosing texts.* The excerpts from two texts—*Vietnam: A History* and *Secrets of the Vietnam War*—were chosen by more students to be read first than any other texts. Each was chosen by about one third of the subjects; the remaining texts were chosen by the remaining third of the subjects. We speculated that the history was chosen because it seemed to provide an overview, and because students would perceive it as neutral in tone. We do not know why the *Secrets* text was so popular. It seemed to be an important source of information in the students' final products as well.

*Selecting information.* When reading a text, students must first select which information is important. Given that these were natural texts, they varied considerably in how well they were constructed. Two texts were especially poorly constructed. *The Pentagon Papers*, for example, is a detailed history of the Vietnam War, written for internal purposes by the Army, and contains many gaps (indicated by a notation reading "Several Paragraphs Missing"). The text was written by and for bureaucrats and is highly inconsiderate of the reader (Anderson & Armbruster, 1984). *The Vietnam Hearings* are a transcript of the hearings, written in a play format. Other texts were written for different purposes than those given to our participants. Commander Stockdale's account used his experiences in the Tonkin Gulf to comment on the unreliability of radar data in a more recent incident. The difficulties of using naturally occurring documents is that the student has to cull through a great deal of irrelevant information to find what is important.

There were differences among documents in how consistent students were in selecting information. We recorded how many students annotated each statement in each of the texts and looked for patterns. In some documents, students tended to select the same idea units in their notes. These tended to be shorter and more focused documents. The statements themselves tended to be clear statements, strongly stating an opinion about the incident or the resolution. For example, 7 of the 9 students who read "The Tonkin Gulf Crisis" annotated the statement "The accumulated evidence makes it reasonably certain that the alleged North Vietnamese PT boat attack of Aug. 4 was a figment of the U.S. government's imagination." Two other statements were annotated by 5 students and another by 4. In other documents, students diverged widely in terms of what information they selected. In *The Pentagon Papers*, one statement was annotated by 4 of

the 6 students who read it. ("Upon first report of the PT boats' apparently hostile intent, F-8E aircraft were launched from the aircraft carrier Ticonderoga, many miles to the south, with instructions to provide aid cover but not to fire unless they or the Maddox were fired upon.") Fifty-three statements in total were annotated, but no other statement was noted by more than half of those reading. Few students read *The Vietnam Hearings*, the other text we judged to be poorly structured.

Thus, it appears that the nature of the text affected how students selected information. Students tended to be more consistent in what information they selected from short, well-constructed texts. In these texts, they tended to choose strong, clear statements of a position. In *The Pentagon Papers* excerpt, a longer, less well-structured text, students chose many different statements, with only one statement chosen by more than half of those who read it.

However, students rarely chose irrelevant information. In the Stockdale article, no student mentioned the current incident, annotating only information dealing with the Gulf of Tonkin Incident. In *The Vote that Congress Can't Forget*, which looked back on the Gulf of Tonkin Resolution by contrasting it with the authorization of the Persian Gulf War, only two students annotated information dealing with the Persian Gulf War. Thus, students seemed good at filtering out information they did not need.

*Does the task the students are given, describing an event or forming an opinion about that event, influence their processing of information?*

*Notetaking.* We hypothesized that the task, either describing or forming an opinion, would affect processing, as evidenced by the notes they took. Students given a task of describing would concentrate more on details in their notes and might include more copying and paraphrasing. Students asked to form an opinion might reduce larger chunks of text into main idea statements and might include more statements classified as reduction or gist in their notes.

Of the students who took notes, 11 were asked for an opinion and 9 were asked to describe either the incident or the resolution. We examined the differences using discriminant analysis, a fairly sensitive multivariate analysis technique. Neither this analysis, nor other appropriate analyses, found significant differences between either those given different tasks or those given different topics.

The lack of differences is surprising, because we expected that students asked to form an opinion would concentrate on more global information and construct more gist statements and evaluative statements, and students asked to describe the incident would concentrate

**Table 5** Notetaking of individual subjects

Task	Topic	Copying	Paraphrasing	Reducing	Making a gist	Evaluating	Distortion/misreading	Total	Number of texts read
6	Opinion	Incident	1	12	2	2	0	17	4
12	Description	Incident	17	20	7	0	4	48	3
16	Opinion	Incident	15	4	2	0	1	22	4
18	Opinion	Incident	11	17	3	0	0	31	4
19	Description	Resolution	17	10	2	0	2	31	3
21	Opinion	Incident	10	4	0	3	4	23	4
23	Description	Resolution	1	2	0	2	0	5	4
26	Opinion	Incident	0	0	0	5	0	6	3
27	Description	Resolution	7	12	0	3	1	23	2
28	Description	Incident	24	21	2	2	1	52	4
29	Description	Incident	11	8	0	21	2	43	4
30	Opinion	Resolution	3	21	3	1	0	28	4
31	Opinion	Resolution	2	7	3	8	0	20	4
33	Description	Resolution	0	1	0	0	0	1	4
34	Description	Incident	14	7	3	3	0	27	5
37	Description	Resolution	2	10	2	3	0	17	2
40	Opinion	Resolution	1	10	1	3	1	20	3
41	Opinion	Incident	6	0	0	0	0	6	6
42	Opinion	Resolution	12	2	5	3	0	22	5
43	Opinion	Incident	7	14	12	18	9	61	6

on details and copy more information directly or in paraphrase. Even those asked for an opinion included few evaluative statements. Of the 11 students asked for an opinion, 1 student (#43) made 9 evaluative statements; 1 other (#21) made 4; 2 students made 1 apiece.

The student who made considerably more evaluation comments than the others, #43, indicated that he had relatively high knowledge of the Vietnam War on the pretest and many of his comments reflect that knowledge. For example, his evaluative comments tended to reflect a strong bias such as "That U.S. was not wrong in firing on the Vietnamese/and that Vietnamese started War./ Johnson's an idiot" and "So it sounds like it's a bunch of idiots playing with their guns." This bias appeared to be based on foreknowledge, rather than developed through reading. This student also read through all six texts in the time allotted for the study, the only student to read this many.

There were strong individual variations in how students approached the task. Some students took copious, detailed notes, no matter which task they were assigned. As noted on Table 5, three students asked to write a description (#12, #19 and #28) copied or paraphrased a great deal of information, as did #18, who was asked to form an opinion. Others tended to write gist statements, condensing a great deal of information into brief, even telegraphic, notes, such as #43 (opinion) and #29 (description).

The order of the texts also seemed to affect how many notes were taken. A repeated measures analysis of variance, looking only at the first three text readings,

found a statistically significant difference among readings,  $F(2, 32) = 9.07, p < .001$ . Students averaged taking 11 notes for the first text, 5 for the second text read, and 7 for the third. The greater amount of notes taken for the first text may indicate that more effort was expended in reading the first text. Recall that only after the first text was read did students make statistically significant growth toward the experts' knowledge structures. There appeared to be little effect of task or topic on readings of the different texts.

#### *How do students integrate information across texts to form a coherent essay?*

**Free recall.** Students were more likely to reduce and make gist statements in the free recalls than in the notes, regardless of whether they were asked to write a description or form an opinion. This behavior seems reasonable, in that students were relying on memory and were not able to easily paraphrase information in the texts. It also argues for the idea that students processed information in similar ways, regardless of the final task.

**Final product.** On the final product, students tended to stick to the task. As can be seen in Table 6, the students who were asked to describe engaged more in paraphrasing, reducing, and making overarching gist statements from a particular text than did students who were asked to form an opinion. Students who were asked to form an opinion rarely paraphrased or reduced. Rather, their final essays were replete with evaluative/gist statements such as, "I believe that the U.S. was too quick to pass the Tonkin Gulf Resolution."

**Table 6** Processes used in final product

	Paraphrasing (single text)	Reducing (single text)	Making a gist (single text)	Paraphrasing (multiple texts)	Reducing (multiple texts)	Making a gist (multiple texts)
Opinion	.07 (.07)	.00 (.00)	.03 (.07)	.01 (.04)	.05 (.08)	.60 (.24)
Description	.20 (.16)	.13 (.13)	.22 (.27)	.10 (.13)	.07 (.08)	.23 (.22)

Note. Numbers are mean percentages. Standard deviations in parentheses.

These statements can only be seen as conclusions reached from reading more than one text, although evidence backing up these statements was scanty at best. Note the low number of paraphrased statements or reductions relating to either one text or a combination of texts. These types of statements would count as evidence backing up their opinions. Interestingly, student #29, who wrote many gist statements in his notetaking despite the fact that he was asked to write a description, stuck to description when he composed his essay. And many students who mostly copied or paraphrased their notes despite the fact that they were asked to write an opinion, wrote opinion-like statements when they composed their essays.

Students who were asked to write an opinion tended to move away from the text, toward broader generalities and statements without providing much apparent factual grounding from the texts read. Even though they had indicated a depth of reading through their notes, their final products seemed to disregard that depth. For example, student # 6 wrote:

(1) *My opinion is that the USS Maddox did get attacked by the North Vietnamese the first time (2) but was not attacked in the second "incident".* (3) The reason for the first attack was that the North Vietnamese thought the Maddox was a South Vietnamese ship (4) and since the South had attacked the night before they defended themselves. (5) Later on the South Vietnamese attacked the North Vietnamese again. (6) The Maddox was again patrolling (7) and the US government thought prematurely that the North Vietnamese would once again attack. (8) *The US government reacted.* (9) *I'm not sure if Johnson lied or what happened.* (10) *In my opinion, something wrong happened. It sounds like it might have been the US fault.* (11) *It might be this because several of the texts said the same thing.* (12) That nothing was out there when the Maddox and the Turner Joy were patrolling. (13) *I am not sure exactly why the USA would do this.* (14) *They might not have.* (italics and numbering added)

This student had taken notes throughout the text, but half of his statements (italics) could not be reconciled directly with any one of the texts that he read. Rather, he

appeared to look at the texts in a global fashion. The task of giving his opinion was viewed as being disassociated from obtaining evidence from the text to support that opinion. The first statement is a clear thesis statement; the following statements do support that thesis. However, by the eighth statement ("The US government reacted."), he gets vague and speaks in generalities, ending in confusion. This may be because of lack of experience with writing coherent texts using an argument structure or because he is still confused by the contradictory texts and has not yet examined the evidence to form an opinion.

As might be expected, the description texts tended to stay closer to the readings. Students provided few evaluative statements. An example would be that of student #28:

(1) The Tonkin gulf incident occurred due to a series of events such as the first battle in which the Maddox was legitimately involved in (due to the attack made by the North Vietnamese). (2) The second battle which some feel never really happened because no one actually saw any PT boats, also had a large effect on the Tonkin gulf situation. (3) It led to Congress passing of the Tonkin gulf Resolution, the retaliatory acts wanted by the Sec. of Def. and other officials that were allowed by President Johnson. (4) These things combined led to the N. Vietnamese feeling that war would occur in the South and moved troops down the Ho Chi Minh trail, resulting in what would possibly be interpreted by US officials as aggression. (Numbers added)

Statements 2 and 3 were supported by three references apiece in the text. The fourth statement was supported by a section in "The Tonkin Gulf Crisis," the last text read. However, this student took copious notes, and very little of the information in his 52 annotations were actually used in this short essay.

**Integration.** Students did appear to use more than one source of information in forming their final essays, and they engaged in rearranging ideas from single texts as they wrote. To examine how students integrated information across texts, we did two analyses. First, we categorized each statement in the text as to whether it



had one source in the readings or whether the idea could be found in multiple readings. (This may have overestimated the number of statements classified as coming from multiple readings, since we categorized an idea as coming from multiple readings whether or not that idea appeared in the students' notes in two places.) We found that students asked to write an opinion tended to use more ideas that came from multiple texts (64% of statements) than students asked to write a description (40% of statements). Students asked to write descriptions used more ideas that could only be found in a single text (55% of statements) than students asked for opinions (10% of statements). (Totals do not add up to 100% due to rounding. Framing and evaluative statements are also not classified here.)

Next, we looked at the ordering of ideas in the texts read and in the final product, using only those ideas that could be identified with a single text. We compared the order of the statements in the final product with the order of those statements in the texts the students read, in the order in which they read them. We used Kendall's Coefficient of Concordance ( $W$ ), a measure of interrater reliability, to compare the different orderings. If students had merely written ideas from the texts in the order in which they were presented, the mean Coefficient of Concordance would have been 1.00. However, the mean of the total final essays was .76, and the range was between .38 and 1.00. Most of the participants made little or moderate changes in how the texts were used in the essay compared to how they were read. Only 2 of the 16 students made drastic reorganizations. There was essentially no difference in the coefficients of the students who were to write descriptions and the students who were to write opinions. The essays from both groups were coherent, in that they had discernable beginnings, middles, and endings.

These results argue for the idea that some students reading multiple texts are able to form a more elaborate network of ideas, in that they seem to be integrating information across multiple sources. For example, Subject #25, whose protocol is reproduced in Appendix B, began his essay with a thesis statement, "President Johnson was definitely justified when he asked Congress to pass the Resolution." His next sentence, "He saw that North Vietnamese were being hostile toward the South Vietnamese, American allies," seemed to draw from two sections of the Resolution as well as two sections of Dean Rusk's autobiographical recollections. The remainder drew from all three sources that he read. The information in the essay was generally not in the same order as the information in the texts, suggesting that he integrated across texts. However, this student was atypical.

Most students reproduced ideas nearly in the same order that they read them.

*Do students engage in corroborating, sourcing, and contextualizing in evaluating historical materials?*

As noted on Table 7, which contains the number of comments in the notes classified as either sourcing, corroboration, or contextualization, few students had comments that could be classified as reflecting the processes used by the historians studied by Wineburg (1991a). What is interesting is that the students who included a great many gist statements also tended to include some sourcing statements. This may reflect the common influence of prior knowledge on both notetaking and critical analysis.

Student #43, who made 18 gist statements, made no references to the source of the documents but made one statement about documents corroborating each other. Student #29 made 21 gist statements and 2 corroborative statements (but only 2 evaluative statements).

Student #42, who included 10 statements dealing with the source of the documents, wrote short, telegraphic notes, copying a key phrase ("America keeps her word") to stand for a larger idea. These were classified as copying in our system, since he used the same words as the text, but is closer to gist than other students' copying. This student wrote a lot of notes, covering five texts. His notes tended to be telegraphic, using just a few words to cover the main ideas. He began staying closer to the text, mainly paraphrasing in the first two texts he read. In his recall, he produced global statements about the text, such as "This text was basically dialogue that outlined how the Senate felt about the current situation in the Gulf of Tonkin" (notes on *The Vietnam Hearings* excerpt). We also classified this as sourcing, since it makes reference to the text, but this is not sourcing in the same sense that Wineburg suggests. Instead, he refers to the text and the participants, not from foreknowledge of their roles, but instead as placeholders representing sides.

These three students (#29, #42, and #43) indicated that they had at least moderately high knowledge of the Vietnam War prior to the readings and could give a reasonably accurate identification of the Tonkin Gulf incident. It may be that some degree of topic knowledge is required to demonstrate sourcing or corroboration, but we did not have enough subjects with moderate or high knowledge to test this statistically. Wineburg (1991a), however, found that even high school students with high amounts of knowledge about the topic he examined (the American Revolution) did not engage in these behaviors as much as historians did, even historians with less factual knowledge. The differences in our findings

**Table 7** Sourcing, corroboration, and contextualization, by subject

ID	Task	Topic	Sourcing	Corroboration	Contextualization
12	Description	Incident	2	0	0
19	Description	Resolution	0	0	0
23	Description	Resolution	1	0	1
27	Description	Resolution	2	0	0
28	Description	Incident	0	0	0
29	Description	Incident	0	2	0
33	Description	Resolution	2	0	0
34	Description	Incident	3	0	0
37	Description	Resolution	0	0	0
6	Opinion	Incident	0	0	0
16	Opinion	Incident	2	0	0
18	Opinion	Incident	0	0	0
21	Opinion	Incident	4	1	0
26	Opinion	Incident	1	0	0
30	Opinion	Resolution	0	0	0
31	Opinion	Resolution	0	0	1
40	Opinion	Resolution	3	1	0
41	Opinion	Incident	0	0	1
42	Opinion	Resolution	10	0	2
43	Opinion	Incident	0	1	0

may result from the different topics. The vast majority of our subjects knew next to nothing about the Vietnam War; all of Wineburg's subjects (students and historians) could be assumed to have at least moderate knowledge of the American Revolution.

With the exception of #42, the number of sourcing, corroboration, and contextualization comments were a small percentage of the total number of comments. That students included very few comments that would be considered as sourcing, corroboration, or contextualization, as shown in Table 7, suggests that they lack the knowledge of the discourse patterns of historical analysis. There were more comments classified as sourcing, which simply involves noting the source, than corroboration and contextualization, which involve more complex operations—comparing the information in the text with either information in other texts (corroboration) or knowledge that students have about the time (contextualization). The comments classified as sourcing did not use the source to understand the text, as a historian would, but merely noted it.

As Wineburg (1991a) pointed out, professional historians approach the task of reading documents as members of a discourse community, but high school students do not. This was true even for historians who scored lower than high school students on a test of the factual content. The knowledge of discourse patterns represents the disciplinary knowledge of history, or the ability to think as a historian might, and may need to be directly taught.

## Conclusion

We studied the processing of students who read multiple historical documents about a controversial event in history—the Gulf of Tonkin Incident and the subsequent Gulf of Tonkin Resolution—that led to heavy U.S. involvement in the Vietnam conflict. We wished to understand what happens to students' mental structures when they read more than one text about an incident, particularly when those texts propose alternate interpretations of the event. We also wanted to see if students would employ different strategies for processing the texts if they were given different purposes for reading.

This study was intended to be exploratory. There were a large number of possible variations in the study—students differed in terms of task and topic and also differed in what texts they read and in what order. From our data, we want to propose a possible model of students' processing of multiple texts, based on our interpretation of our data, and then discuss why studying documents alone might not lead to the disciplinary knowledge, as has been proposed by Wineburg (1991a) and others.

Our basic model suggests that we can break the process down into *Selection* of ideas in each text read, *Processing* of ideas within that text, *Constructing* a mental model of the information, and *Integrating* ideas across texts to produce a final product. These will be discussed in turn.

*Selection.* Our students tended to be strongly influ-

enced by text features in their selection of ideas. Students consistently chose the same ideas to note from short, well-structured texts. These ideas tended to be clear and strong statements of opinion or topic sentences encompassing a great deal of detail. With long and ill-structured texts, such as *The Pentagon Papers*, few students chose the same ideas. With such texts, it was difficult to pick main points, since there was so much detail and little attempt at organizing it. Original source documents, however, tend to be more like *The Pentagon Papers* and *The Vietnam Hearings*, than like the shorter, more focused pieces. Students need to learn how to cull information from longer documents if they are going to be used in units such as this.

Students, however, were able to concentrate on relevant information. In two pieces written for a purpose different than the purpose given the students, they consistently ignored irrelevant information, focusing instead on information suited to the purpose.

*Processing.* We found that task had little effect on how students read the information in the different texts. Students asked for an opinion did not differ from those asked for a description in the types of notes they took. There were few evaluative statements given at all, even by those asked for an opinion. Students tended to take many more notes on the first text than on subsequent texts, suggesting that they were expending more cognitive effort in constructing a mental model of the information in the text.

There were, however, strong individual differences in notetaking. Some students tended to take copious, detailed notes, relying on copying and paraphrasing. Others tended to rely on gist statements, noting only main points, often only telegraphically. These differences in notetaking strategy do not seem to be related to the task or which text was read, but seem to be an individual difference.

*Constructing.* Our analysis of the students' ratings of relatedness among key terms suggests that students' mental structures tend to grow in two ways while reading multiple texts. Students' structures tend to be more internally consistent after reading a single text, and then still more consistent after reading a second text. The history text tended to produce the greatest gains in harmony or internal consistency. Students' structures also tended to become more similar to those of experts after reading a single text, with no further growth after reading two or more texts.

Since we are using similarity to experts as our metric of the growth of knowledge, this suggests that students did not grow in their knowledge after more than one reading, but they did become more consistent in their understandings. This lack of growth (or failure to

move closer to the knowledge structure of experts) may be simply because they did not process the subsequent texts as well. There is evidence of a clear decline in notetaking after the first text read.

Another complementary explanation might lie in the nature of the texts read. Because the texts were chosen to contradict each other, students may have looked for overlap between texts, rather than for new knowledge. The overlap would reinforce the basic knowledge acquired by reading the first text but might not add very much to the student's understanding. In fact, to be consistent, some contradictory information would have to be ignored. Because the internal consistency of our students increased, we might posit that students were looking for overlap and ignoring this contradictory new information.

*Integrating.* The task students were given strongly influenced their final product. Students asked for a description tended to stay close to the texts, with most of their statements coming clearly from information provided, usually in a single text. Students asked for an opinion tended to produce more global statements, not clearly tied to any single text, that could be found either in multiple texts or not in any text.

#### Limitations

Because the study was intended to be naturalistic, we did not control much that one might ordinarily control. Because we let students make decisions about what they did, we had different numbers of students in different analyses. In future studies, first, we need to systematically vary the texts that students read. We wanted to see what would happen if students were given freedom to choose whatever texts they wanted to. We expected to see if there was a pattern to their choice. We need to vary the texts in a principled manner to see how different types of texts—histories, opinions, source information—affect students' learning.

Second, we need to vary background knowledge of our topic. Students' knowledge was uniformly low. A great deal of effort seemed to be expended on constructing a basic understanding of what went on in the Gulf of Tonkin and in the U.S. Senate during the discussion of the Resolution. This may have hampered students in evaluating the information in the texts as we hoped they would.

Third, the problem of the Tonkin Gulf is a problem of perception—which of two clearly contradictory sides is correct? The processes described here might not be found in a less polarized topic, such as the Panama Canal Treaty as studied by Perfetti and his colleagues (1993). We need to compare different types of historical problems to examine their separate effects on students' learning.

It should be noted that the task used differed in at least one important way from that of professional histori-

ans. Students were asked to read each paper and recall the information from it. Although students could go back and forth between documents, we had no data about whether they actually did. Historians, on the other hand, read and reread texts. Also, in a more natural setting, students would have had access to the texts in writing their final product.

### *Thinking like a historian using multiple source documents*

Some students did engage in some of the processes described by Wineburg (1991a) as being typical of professional historians—contextualizing, corroborating, and sourcing. This was also an individual difference, with some students doing this frequently. Most did not evidence these processes at all.

For most of these students, though, simply presenting them with multiple texts did not encourage them to think like historians. In fact, the greatest growth of knowledge came after the reading of the first text, and the text that had the greatest influence on growth of harmony was a well-organized history textbook, albeit a text devoted entirely to Vietnam. Students read the first text to get basic facts and information and read subsequent contradictory texts trying to sort out that information.

One reason that many students did not seem to develop disciplinary knowledge from reading multiple texts was their lack of initial knowledge about the topic. Students' initial reliance on the history text and their tendency to take paraphrase-type notes may have been reflections of their need to gain a literal understanding of the content before attempting to produce an opinion. Alexander and Judy (1988) argued that students become able to use more sophisticated strategies for learning new information when they already have some content knowledge. The students we studied may have initially been taking notes in paraphrase fashion because they lacked background knowledge and were reading to gain this knowledge, regardless of the final task they had been assigned. They may not have been sophisticated enough to develop an opinion, if that was their task, until they had read at least two documents. Students began by paraphrasing the texts closely and were more likely to reduce information as they read subsequent texts. This tendency to move towards reduction may have been a result of their growing background knowledge.

A second reason that students did not seem to benefit from just reading multiple texts is that they may need to be taught what it means to think like a historian, and that, without this teaching, students will be less able to engage in historical analysis. In other words, students who know more about historical analysis may be more able to engage in it. It is possible, for instance, that the

four students who exhibited more gisting, evaluating, sourcing, and corroborating may have been more sophisticated readers of historical text, regardless of whether they were familiar with the Gulf of Tonkin Incident. Textbooks in history are written so that the author's background and stance and his or her methodology are hidden (Luke, DeCastell, & Luke, 1983). Therefore, interpretations of events are presented as fact, not analysis, and two or more interpretations to an event are rarely shared. While original documents and argumentative essays positing different interpretations should help students come to the realization that history is interpretation rather than fact, this idea may be less obvious to students who have relied mainly upon history texts for information and who have been taught to think of history as merely a series of chronicled events. The teacher in this study generally presented history as a story, stressing the relations among events, rather than the interpretations.

In this study, students tended to respond to the description task in a transmission mode. That is, they looked at the task as one of getting information from the text and writing it down. In contrast, they looked at the opinion task as a prior knowledge task and did not rely on the information in the texts to support their opinion. In neither task did the students use the texts to respond critically and evaluatively, as intended.

The lack of critical response may suggest that students need to be taught to write persuasive essays, with a warrant and evidence supporting that warrant. Chambliss (1994) found that there are differences in how students evaluate persuasive essays to formulate their own opinions. Our students made many unsupported statements when asked to form an opinion, even though their notes indicated that they had attended closely to the information in the texts and did have that information at hand. It is possible that these students did not know that they were supposed to provide support for an opinion, even though they clearly learned information that would be appropriate. This is another aspect of the disciplinary knowledge of history, and of other disciplines as well.

A final possible reason for the apparent lack of benefit from reading multiple texts may be a lack of experience with the task of working with multiple texts. As noted earlier, their teacher did not provide such experience but planned to do so later in the year. Experience (and teacher guidance) may improve students' ability to integrate information from different original source documents.

### REFERENCES

- ALEXANDER, P.A. , & JUDY, J. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of*



*Educational Research*, 58, 375–404.

ANDERSON, T.H., & ARMBRUSTER, B.B. (1984). Content area textbooks. In R.C. Anderson, J. Osborn, & R.J. Tierney (Eds.), *Learning to read in American schools* (pp. 193–226). Hillsdale, NJ: Erlbaum.

BELENKY, M.F., CLINCHY, B.M., GOLDBERGER, N.R., & TARULE, J.M. (1986). *Women's ways of knowing*. New York: Basic Books.

BRITTON, B.K., & GULGOZ, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*, 83, 329–345.

CHAMBLISS, M.J. (1994). Why do readers fail to change their beliefs after reading persuasive text? In R. Garner & P.A. Alexander (Eds.), *Beliefs about text and instruction with text* (pp. 75–89). Hillsdale, NJ: Erlbaum.

CRAFTON, L.K. (1983). Learning from reading: What happens when students generate their own background information? *Journal of Reading*, 26, 586–592.

GLASER, B.G., & STRAUSS, A.L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

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.

HARTMAN, D.K. (1995). Eight readers reading: The intertextual links of proficient readers reading multiple passages. *Reading Research Quarterly*, 30, 520–561.

HIRSCH, E.D. (1987). *Cultural literacy*. Boston: Houghton-Mifflin.

HYND, C.R., STAHL, S.A., BRITTON, B., & MCNISH, M. (1996, April). *Group processes involved in studying multiple documents in history*. Paper presented at the annual meeting of the American Educational Research Association, New York.

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

KINTSCH, W., & VAN DIJK, T.A. (1978). Toward a model of discourse comprehension and production. *Psychological Review*, 85, 363–394.

LEINHARDT, G., BECK, I.L., & STANTON, C. (1994). *Teaching and learning in history*. Hillsdale, NJ: Erlbaum.

LUKE, C., DECASTELL, S., & LUKE, A. (1983). Beyond criticism: The authority of the school text. *Curriculum Inquiry*, 13, 111–128.

MCKEOWN, M.G., BECK, I.L., & WORTHY, M.J. (1993). Grappling with text ideas: Questioning the author. *The Reading Teacher*, 46, 560–566.

PATRICK, J.J., & HAWKE, S. (1982). Social studies curriculum materials. In *The current state of social studies: A report of Project Span* (pp. 105–185). Boulder, CO: Social Science Education Consortium.

PERFETTI, C.A., BRITT, M.A., & GEORGI, M.C. (1995). *Text-based learning and reasoning: Studies in history*. Hillsdale, NJ: Erlbaum.

PERFETTI, C.A., BRITT, M.A., ROUET, J.F., MASON, R.A., & GEORGI, M.C. (1993, April). *How students use texts to learn and reason about historical uncertainty*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

PERRY, W.G. (1970). *Forms of intellectual and ethical development in the college years*. New York: Holt, Rinehart, & Winston.

RAVITCH, D. (1992). *The democracy reader*. New York: HarperCollins.

REYNOLDS, R.E., TRATHEN, W., SAWYER, M., & SHEPARD, C.R. (1993). Causal and epiphenomenal use of the selection attention strategy in reading comprehension. *Contemporary Educational Psychology*, 18, 258–278.

RUMELHART, D.E. (1980). Schemata: The building blocks of cognition. In R. Spiro, B. Bruce, & W. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, NJ: Erlbaum.

SEIXAS, P. (1993). The community of inquiry as a basis for knowledge and learning: The case of history. *American Educational Research Journal*, 30, 305–324.

SPIRO, R. (1980). Constructive processes in prose comprehension and recall. In R. Spiro, B. Bruce, & W. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 245–259). Hillsdale, NJ: Erlbaum.

SPIRO, R.J., COULSON, R.L., FELTOVICH, P.J., & ANDERSON, D.K. (1994). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In R.B. Ruddell, M.R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (pp. 602–615). Newark, DE: International Reading Association.

SPIVEY, N.N., & KING, J. (1989). Readers as writers composing from sources. *Reading Research Quarterly*, 24, 7–26.

SPOEHR, K.T., & SPOEHR, L.W. (1994). Learning to think historically. *Educational Psychologist*, 29(2), 71–78.

STAHL, S.A., & HYND, C.R. (1994, April). *Selecting historical documents: A study of student reasoning*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

STAHL, S.A., HYND, C.R., GLYNN, S., & CARR, M. (1995). Beyond reading to learn: Developing content and disciplinary knowledge through texts. In P. Afflerbach, L. Baker, & D. Reinking (Eds.), *Developing engaged readers in home and school communities* (pp. 139–163). Hillsdale, NJ: Erlbaum.

WINEBURG, S.S. (1991a). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, 83, 73–77.

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

WINEBURG, S.S. (1994). Introduction: Out of our past and into our future—The psychological study of learning and teaching history. *Educational Psychologist*, 29(2), 57–60.

Received June 13, 1995

Final revision received January 2, 1996

Accepted February 8, 1996

## AUTHOR NOTE

The work reported herein was funded in part by the National Reading Research Center of the University of Georgia and University of Maryland. It was supported under the Educational Research and Development Centers Program (PR/AWARD NO. 117A20007) as administered by the Office of Educational Research and Improvement, U.S. Department of Education. The findings and opinions expressed here do not necessarily reflect the position or policies of the National Reading Research Center, the Office of Educational Research and Improvement, or the U.S. Department of Education.

## APPENDIX A

### Background information

---

The war in Vietnam has been called the United States' longest war, because, even though the U.S. was not in combat the entire time, it was involved in the affairs of Vietnam for approximately 25 years, from 1950 to 1975. The U.S. became involved during the Truman administration, when it supported the French (who controlled the Vietnamese government) against a group of Communist rebels fighting for Vietnamese independence. By the time Lyndon Baines Johnson had taken over the presidency (in 1963), Vietnam had been divided into North and South, with the North being governed by the Communist president Ho Chi Minh, and the South being governed by a U.S.-supported president. In South Vietnam, a civil war had broken out in an attempt to topple the existing government and reunite North and South Vietnam under communism. This movement was led by a group the U.S. labeled the Viet Cong. The U.S. sent monetary aid, equipment, and advisors to the South Vietnamese government to support their fight against the Viet Cong and monitored, with concern, North

Vietnamese support of the rebels. It was against this backdrop that the Gulf of Tonkin Incident took place.

On August 2, 1964, shots were fired toward the U.S.S. Maddox by three PT boats while on patrol off the North Vietnamese coastline in the Gulf of Tonkin. Two days later, while the Maddox and a companion ship, the C. Turner Joy, were again on patrol, there were reports of another attack. President Johnson ordered a retaliatory strike and asked Congress to pass the Southeast Asia Resolution (also known as the Gulf of Tonkin Resolution) to give him the authority to "take all necessary steps, including the use of armed force, to assist any member or protocol state of the Southeast Asia Collective Defense Treaty requesting assistance in defense of its freedom." This resolution was passed. Johnson used this approval to commit the U.S. to heavy involvement in the Vietnam War. "Hawks" (those who were supporters of the war) and "Doves" (those who were against the war) disagreed about what actually happened and about President Johnson's motivations in handling the incident.

---

**APPENDIX B****Analysis sheet for notes and final product**

Text	#25—Take side in a debate (pro resolution) Notes	Final product
<b>1. Gulf of Tonkin Resolution</b>		
Three days after the second incident in the Tonkin Gulf, the Administration submitted a Joint Resolution to Congress which approved in advance the President's taking "all necessary steps" to assist South Vietnam or any other member or protocol state of the Southeast Asia Treaty Organization.	Administration submitted a Joint Resolution to Congress which approved the President's request to assist the Southeast Asia Treaty Organization	(1) President Johnson was definitely justified when he asked Congress to pass the Resolution.
It was approved unanimously by the House and by a vote of 88 to 2 in the Senate.		(7) The vote passed unanimously in the House of Representatives and eighty to two in the Senate in approval of military action against the North Vietnamese (Viet Cong).
Whereas naval units of the Communist regime in Vietnam, in violation of the principles of the Charter of the United Nations and of international law,	Communist Regime violated the Charter	
have deliberately and repeatedly attacked United States naval vessels lawfully present in international waters, and have thereby created a serious threat to international peace;	attacked UN vessels creating a threat to international peace	(2) He saw that North Vietnamese were being hostile toward the South Vietnamese, American allies (3) and in the process were attacking U.S. ships in the Gulf of Tonkin.
Whereas these attacks are part of a deliberate and systematic campaign of aggression that the Communist regime in North Vietnam has been waging against its neighbors and the nations joined with them in the collective defense of their freedom;	attacks are part of a deliberate and systematic campaign of aggression that the Communist regime has been waging against its neighbors	(2) He saw that North Vietnamese were being hostile toward the South Vietnamese, American allies

*(continued)*

**APPENDIX B (cont'd.)****Analysis sheet for notes and final product**

Text	#25—Take side in a debate (pro resolution) Notes	Final product
Whereas the United States is assisting the peoples of Southeast Asia to protect their freedom and has no territorial, military or political ambitions in that area, but desires only that these peoples should be left in peace to work out their own destinies in their own way:	US protecting the people's freedom	
Consonant with the Constitution of the United States and the Charter of the United Nations and in accordance with its obligations under the Southeast Asia Collective Defense Treaty, the United States is, therefore, prepared, as the President determines, to take all necessary steps, including the use of armed force, to assist any member or protocol state of the Southeast Asia Collective Defense Treaty requesting assistance in defense of its freedom.	the US is prepared to take all necessary steps to assist the Southeast Asia Collective Defense Treaty requesting assistance in defense of its freedom	
This resolution shall expire when the President shall determine that the peace and security of the area is reasonably assured by international conditions created by action of the United Nations or otherwise, except that it may be terminated earlier by concurrent resolution of the Congress.	this power of the President will expire when peace has returned to South Vietnam or as Congress sees fit	
<b>2. "The Vote that Congress Can't Forget"</b>		
For more than two decades, the Congressional vote that lawmakers most often cite as the one they would like to take back is their 1964 vote for the Gulf of Tonkin Resolution, the resolution that was used as authority for the war in Vietnam. Only two Senators and no Representatives voted no.	The vote that lawmakers would most like to take back is the Gulf of Tonkin Resolution in which two Senators and no Representatives noted no.	(7) The vote passed unanimously in the House of Representatives and eighty to two in the Senate in approval of military action against the North Vietnamese (Viet Cong).

*(continued)*



**APPENDIX B (cont'd.)****Analysis sheet for notes and final product**

Text	#25—Take side in a debate (pro resolution) Notes	Final product
Twenty-seven of those lawmakers are still in Congress.	27 of these lawmakers are still in Congress	
And as they prepared for Saturday's vote on what was even more clearly the equivalent of a declaration of war, the resolution to authorize the use of military force in the Persian Gulf, the memories of that earlier vote weighed heavily on the minds of the seven senators and 20 representatives who served then and still serve now.	They thought about this when making the decision to invade the Persian Gulf.	
He recalled the earlier resolution as one President Johnson had "distorted," and one whose repeal he accomplished as a freshman Senator in 1970.	the resolution was repealed one year after it was passed	
He added, "The Gulf of Tonkin Resolution was used as a declaration of war and plunged this country in eight or nine years of really disastrous war in Vietnam."	was used as a declaration of war	
Representative Charles E. Bennett, Democrat of Florida, said; "I am 80 years of age, I have been in this chamber 43 years. Out of the 17,000 votes I have cast, the only one I really regret is the one I cast for the Bay of Tonkin Resolution.	Charles E. Bennett regrets it.	
Representative Dan Rostenkowski, Democrat of Illinois, did not offer second thoughts about his old vote, but said the Persian Gulf decision was even more difficult for him. "Today's situation is clearer," he said in a statement in the Congressional Record, "The possibility of armed conflict, casualties and even death is much more apparent.	Persian Gulf situation is clearer	

*(continued)*

**APPENDIX B (cont'd.)****Analysis sheet for notes and final product**

Text	#25—Take side in a debate (pro resolution) Notes	Final product
But few avoided comparisons with Vietnam more generally.	Several lawmakers avoided comparisons of the situations	
When it came time to vote, 12 of the Tonkin veterans voted to authorize force and 14 voted against it.	12 authorized 14 no for the Gulf	(8) Even twenty years after the vote in Congress, twelve Congressmen still believed that the Resolution was justifiable.
The Tonkin Democrats sided against authorizing force.	The Tonkin Democrats sided against authorizing force in the Gulf	
So did Representatives Jack Brooks of Texas, John D. Dingell of Michigan, Dante B. Fascell of Florida, and Jamie L. Whitten of Mississippi, the only current member of Congress who was also on hand in December 1941, to vote the last formal declarations of war, against Japan, Germany and Italy.	James L. Witten of Mississippi, the only current member of Congress who was on hand in December 1941, to vote the last formal declarations of war, against Japan, Germany, & Italy voted Yes.	
<b>3. As I Saw It</b>		
He consistently favored strong American involvement, arguing that "aggression" must be stopped. (from "About the Author")	Dean Rusk favored American involvement	
Dean Rusk was Secretary of State under Presidents Kennedy and Johnson, 1961–68. (from "About the Author")	Dean Rusk was Secretary of State under Kennedy and Johnson	
On August 2 and 3, 1964, we received reports that the U.S.S. Maddox and U.S.S. C. Turner Joy, American destroyers operating in the Gulf of Tonkin off the coast of North Vietnam, had been attacked by North Vietnamese torpedo boats in two separate incidents.	On August 2 and 3 1964 the USS Maddox and USS C Turner Joy on the coast of North Vietnam had been attacked	(2) He saw that North Vietnamese were being hostile toward the South Vietnamese, American allies (3) and in the process were attacking U.S. ships in the Gulf of Tonkin.

*(continued)*

**APPENDIX B (cont'd.)****Analysis sheet for notes and final product**

Text	#25—Take side in a debate (pro resolution) Notes	Final product
The Republic of Vietnam today celebrates August 2—the day of the Tonkin Gulf attacks—as part of its national war effort against the Americans, so whatever happened that night in the Tonkin Gulf, evidently it takes credit for it now.	The Republic of Vietnam takes credit for something happening because they celebrate August 2 as part of its efforts against America	
North Vietnam was using coastal waters to infiltrate men and arms into South Vietnam;	North Vietnam was using the coast to infiltrate men and arms into South Vietnam	(2) He saw that North Vietnamese were being hostile toward the South Vietnamese, American allies
South Vietnam under the doctrine of self-defense was trying to block this infiltration and mount retaliatory raids of its own—a secret operation called 34-A, supported by the American Navy. But the destroyers attacked in the Gulf of Tonkin were on intelligence-gathering missions, not participating in South Vietnamese actions along the coast. It is entirely possible that the North Vietnamese thought that our destroyers were involved in these 34-A raids and in blockading operations along North Vietnam's coast to stop their infiltration of the South by sea. But even if Hanoi thought this, it isn't valid to call the exercise of self-defense a provocation.	North Vietnamese could have thought that Americans were part of a South Vietnamese operation called 34-A	(9) The navy needed to protect the South Vietnamese and American intelligence vessels.
		(4) President Johnson could not allow for the continued meaningless destruction of governmental property by North Vietnamese without doing something about it.

*(continued)*

**APPENDIX B (cont'd.)****Analysis sheet for notes and final product**

Text	#25—Take side in a debate (pro resolution) Notes	Final product
Indeed, the Tonkin Gulf Resolution, in which Congress declared its support for the United States' willingness to come to the assistance of those protected by the SEATO Treaty, including the use of armed force "as the President shall determine," was passed rapidly: 88-2 by the Senate and 416-0 by the House.	The resolution was also unanimously voted	(6) Congress also felt he was justified in his actions. (7) The vote passed unanimously in the House of Representatives and eighty to two in the Senate in approval of military action against the North Vietnamese (Viet Cong).
Some later complained, "We didn't anticipate sending a half million men to South Vietnam," but neither did Lyndon Johnson.	but shortly after people began to change their minds	(5) He would not foresee the tragic death and destruction in the future Vietnam War.
I felt the Tonkin Gulf Resolution was not congressional evasion of its war powers responsibility, but an exercise of that responsibility.	Resolution was an exercise of congressional powers not an evasion of them	(10) Thus, Lyndon B. Johnson and Congress acted rightly and in good faith regarding the approval of the Gulf of Tonkin Resolution.