How to support learning from multiple hypertext sources

ANJA B. NAUMANN AND INA WECHSUNG

Deutsche Telekom Laboratories, Berlin University of Technology, Berlin, Germany

AND

JOSEF F. KREMS

Chemnitz University of Technology, Chemnitz, Germany

In the present study, we investigated three factors that were assumed to have a significant influence on the success of learning from multiple hypertexts, and on the construction of a documents model in particular. These factors were task (argumentative vs. narrative), available text material (with vs. without primary sources), and presentation format (active vs. static). The study was conducted with the help of the combination of three tools (DEWEX, Chemnitz LogAnalyzer, and SummTool) developed for Web-based experimenting. The results show that the task is the most important factor for successful learning from multiple hypertexts. Depending on the task, the participants were either able or unable to apply adequate strategies, such as considering the source information. It was also observed that argumentative tasks were supported by an active hypertext presentation format, whereas performance on narrative tasks increased with a passive presentation format. No effect was shown for the type of texts available.

A great feature of the World-Wide Web is that it provides easy access to a wide variety of information from different texts on one topic (Wiley, 2001). Moreover, hypertext, which is the common text format on the Web and is characterized by nodes and links, seems to be the appropriate medium for dealing with multiple texts: Information from one text can be compared with information from other texts more easily than with conventional text (Perfetti, Britt, & Georgi, 1995). Multiple hypertexts are hypertexts on the same topic from different authors or sources. Studies concerning learning with hypertext have predominantly been conducted using a conventional, single text segmented into hypertext. Only in recent years has there been an increase in the number of studies investigating the potential benefit of learning from multiple hypertexts (e.g., Goldman, 2004; Stadtler & Bromme, 2005). The factors theoretically assumed to support learning from multiple hypertexts still need to be validated (Stadtler & Bromme, 2005). The aim of the present study is to bridge this gap between text comprehension research and recommendations for hypertext design, and to identify factors that support efficient knowledge acquisition with multiple hypertexts.

To date, most studies on this topic deal with learning and knowledge acquisition from history texts (Perfetti, Rouet, & Britt, 1999; Stahl, Hynd, Britton, McNish, & Bosquet, 1996; VanSledright, 2002; Wiley & Voss, 1999; Wineburg, 2000). One advantage of this subject is that historical events are often documented from different per-

spectives; thus, a critical approach to historical documents is required from the reader. He or she has to apply adequate reading strategies to be able to build up a coherent mental representation of the text content (Wineburg, 1994). Empirical evidence regarding learning with multiple texts is rather vague. On the one hand, some studies report more benefit for multiple documents in comparison with single documents (Rouet, Britt, Mason, & Perfetti, 1996; Stadtler & Bromme, 2005; Voss & Wiley, 2000; Wiley & Voss, 1999). This is explained by better comprehension of the content due to the necessity of identifying and correlating different relations and opinions (Voss & Wiley, 2000). On the other hand, it has been shown that the use of multiple texts or multiple documents does not result in higher learning performance per se (Stahl et al., 1996; Wineburg, 2000). Therefore, it is necessary to clarify the factors that influence comprehension.

When reading, readers actively build up a mental representation of the text (Graesser, Millis, & Zwaan, 1997; Kintsch, 1998). Perfetti et al. (1999) presented a theoretical framework describing possible mental representations of multiple texts. Based on Kintsch and van Dijk's model of text discourse comprehension (Kintsch, 1998), the framework posits four different theoretical models of a document's representation: the *separate representation model*, the *mush model*, the *tag-all model*, and the *documents model*. The models differ from each other in the degree to which the information is integrated in the *situation model*

and in the degree to which the source information is considered. The mental representation they proposed as “most typical of a good reader’s model of multiple learning” (Britt, Rouet, & Perfetti, 1999, p. 220) is the documents model. The documents model consists of an intertext model and a situation model. The intertext model contains links between different source information and links from source information to information represented in the situation model. The situation model (see also Johnson-Laird, 1983; Singer, Graesser, & Trabasso, 1994; Zwaan & Radvansky, 1998) is built up while reading. When reading different texts, readers may be confronted with different, even conflicting, situations. In the ideal case, readers update their situation model and are able to also integrate inconsistent information. For that purpose, the information described has to be linked to the corresponding nodes in the intertext model. Figure 1 shows an example of the documents model representing a segment of text from two authors (Clark and La Costa). Shown are the representation of the situation (white boxes and solid arrows) and the intertext links to (dotted lines) and between (fat solid line) document nodes (shaded boxes).

To date, the conditions under which the documents model is established have barely been tested. Findings from text comprehension research show that sometimes more effortful processing can lead to a better understanding of the subject matter (e.g., Einstein, McDaniel, Owen, & Côté, 1990), because the active preoccupation with the text material makes it easier to draw inferences and results in mental representations that are better linked (Voss &

Wiley, 2000). However, for text recall, more effortful processing is rather disadvantageous (McNamara, Kintsch, Songer, & Kintsch, 1996; Wiley & Voss, 1999). Recall performance tends to be better for narrative tasks than for tasks that require more processing (Hemmerich & Wiley, 2002; Stahl et al., 1996; Wiley & Voss, 1999). Therefore, argumentative tasks, because they require active interaction with the text, should promote the construction of a documents model. Construction on the intertextual level is especially supported by argumentative tasks (Voss & Wiley, 2000) because they contain the implicit hint that different opinions or perspectives exist regarding the topic (Wiley, 2001). Narrative tasks, on the other hand, will involve a more superficial representation. Consistent with studies using individual texts, the advantage of argumentative tasks was shown for learning with multiple texts (Voss & Wiley, 2000). Therefore, in the present study we hypothesized that argumentative tasks support the construction of a documents model and, with it, a more efficient navigation and a better comprehension of multiple hypertexts and sources. Narrative tasks, on the other hand, should support the recall of a text.

As for argumentative tasks, an active interaction between reader and text is assumed for primary sources (e.g., Rouet et al., 1996; Voss & Wiley, 2000). Primary sources often contain no causal inferences and do not show a clear line of argument. Thus, the reader needs to draw inferences by him- or herself. Therefore, primary sources should support the construction of a mental representation with more causal connectives during the read-

Figure 1. The documents model representing a segment of text from two authors. From *Narrative Comprehension, Causality, and Coherence: Essays in Honor of Tom Trabasso* (p. 221), by S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), 1999, Mahwah, NJ: Erlbaum. Copyright 1999 by Lawrence Erlbaum Associates. Adapted with permission.

ing of multiple texts. To develop his or her own opinion on the described situation, the reader has to evaluate the reliability and credibility of the documents on the basis of their source information (e.g., Britt et al., 1999). Primary sources are preferred by readers and rated as most reliable and most useful (Rouet, Favart, Britt, & Perfetti, 1997). Therefore, we hypothesized that the construction of a documents model is supported by the use of primary sources. This effect should be enhanced when an argumentative task is given because results have shown that primary sources are taken more into account with an argumentative task (Rouet et al., 1996).

Another factor in learning from electronic texts is the presentation format. The results of studies conducted with single hypertexts show that linear, static presentations are superior to active presentation formats for reading and knowledge acquisition (Naumann, 2003; Naumann, Wanek, & Krems, 2001). The causes seem to be orientation problems ("lost in hyperspace") and cognitive overhead when interacting with hypertext (Conklin, 1987; McDonald & Stevenson, 1998; Naumann, 2003). In contrast, studies with multiple texts show that hypertext environments can support the construction of a mental representation, but this is possible only when the task demands that an integrative approach be taken and when the browser design supports that process (Wiley, 2001). In addition, special functions in hypertext environments (e.g., Stadtler & Bromme, 2005) can enhance the ability to search for and consider source information. It seems that hypertext environments reduce the effort required for comparing relevant information from text to text. Therefore, we assumed that an active hypertext format would help the reader to construct a mental representation of multiple texts when an argumentative task was given. The narrative condition, on the other hand, combined with a static hypertext format, should lead to the best results regarding the recall of a text.

In summary, the construction of a documents model should be supported when an argumentative task is given (e.g., Wiley & Voss, 1999) and if the text material is presented in a way that promotes an active interaction between text and reader—that is, primary sources (e.g., Voss & Wiley, 2000). Furthermore, the reader is more likely to build up a documents model if source information about the documents is available (Rouet et al., 1996) and if the presentation format is active hypertext (Wiley & Voss, 1999).

METHOD

Participants

A total of 123 students of Chemnitz University of Technology participated in the experiment for course credit or cash. Because 1 participant was not a native German speaker and 3 participants misunderstood the instructions, 4 cases were excluded from the analysis. Thus, data from 119 participants (34 male, 85 female) with an average age of 23 years ($M = 23.4$, $SD = 4.5$) were analyzed.

Materials

Texts. Ten text excerpts about the beginning of the revolution in Panama in 1903 served as text materials. These texts were adapted from Rouet et al. (1996) and translated into German. The text material consisted of one list of facts, two authoritative primary sources, two reports from people involved in the event, one textbook passage,

and four essays from historians. In addition, a short introduction to the topic was given.

The participants read eight texts altogether. In the condition with primary sources, participants received one list of facts, two authoritative primary sources, two reports of people involved in the event, a textbook passage, and two essays from historians. In the condition without primary sources, the primary sources were replaced by another two secondary sources (essays from historians). The number of texts was the same for both conditions. The eight texts and the introduction were presented one at a time.

Pretests. Before the reading task, participants filled out a brief demographic questionnaire and a questionnaire about their attitudes toward computers. Additionally, they completed a knowledge test containing four questions about the history of Panama, to check their previous knowledge about the topic.

Posttests. After the reading session, the first task was to write an essay. It was followed by a sentence recognition task and an inference verification task. These tests were constructed according to Wiley and Voss (1999). The aim was to measure whether the participants constructed a situation model or a more superficial representation of the text during reading. The essays were typed into a text field with 20 lines of 100 characters per line. Depending on the condition, participants were asked either to form an opinion about a controversy on the Panama topic or to write a description of the events. The essays were analyzed regarding their structure, their connections, and the origin of their sentences. For all of the above analyses the coding was adapted from Wiley and Voss. The coding was conducted double-blind by two interraters. Pursuant to Wiley and Voss, the organization of the essays was either a linear structure in which events or ideas were simply listed, or an analytical structure in which a thesis was stated and supported. For this analysis, the interrater reliability was Cohen's $\kappa = .89$. The more causal connections an essay contains, the deeper the understanding of the subject matter is assumed to be (Wiley & Voss, 1999). First, we counted all connections listed in Pasch, Brauße, Breindl, and Waßner (2003); then we divided them into the two categories *causal connections* and *other connections*. On the basis of the total sum of all connections, the percentage of causal connections was calculated (Cohen's $\kappa = .82$). To determine the origin of the information, we assigned the content of each sentence to one of four categories: *borrowed*, *transformed*, *added*, and *added but irrelevant*. The last category was not used by Wiley and Voss but seemed necessary in light of the large amount of new but irrelevant information in the participants' essays. Borrowed information was taken directly or paraphrased from the text material. The category *transformed information* included sentences that contained presented information combined with some new information. Sentences were also coded as transformed when they consisted of presented information previously not connected. If a sentence contained novel and relevant information, it was coded as added; if the information was novel but had nothing to do with the topic (e.g., "I think now I have written enough"), it was coded as irrelevant (Cohen's $\kappa = .84$).

Moreover, we investigated whether the participants paid attention to the source information. The first step was to analyze whether the essays included references. Second, the references were, according to Rouet et al. (1996), classified into two types: *general references* (e.g., "Supporters of the opponent opinion argued") and *specific references* (e.g., "According to J. Wilson . . ."). To analyze how many documents were explicitly cited per essay, only one specific reference per text was counted (Cohen's $\kappa = .87$ for general references, Cohen's $\kappa = .88$ for specific references).

The *sentence recognition task* as well as the *inference verification task* consisted of 10 items. In the *sentence recognition task*, participants read sentences and were asked to verify whether they appeared in the exact wording as in the texts they had read. The gist of the sentence was always the same—only the phrasing was changed. The *sentence recognition task* was used to test the recall of the superficial structure of the texts. The *inference verification task*

required the participants to classify 10 statements as valid or invalid inferences on the basis of the information in the texts. For each test, half of the items were true, the other half false. This test was used as an indicator of text understanding.

To assess the acquisition of factual knowledge, a multiple-choice test was given after the inference verification task. The multiple-choice test included 21 questions, each with one correct answer and three incorrect alternatives. It was followed by a questionnaire about orientation and navigation problems using a 5-point rating scale (*does not apply at all to applies completely*). The final task was to answer 5 open-ended questions about the texts' content. To see the experiment from the participants' perspective, please visit <http://e-learning.qu.tu-labs.tu-berlin.de/Mq/start.html> (fill in "A001" for the ID).

Design

In a $2 \times 2 \times 2$ design, the factors presentation task (argumentative vs. narrative), presentation format (active vs. static), and type of available documents (with primary sources vs. with only secondary sources) were manipulated.

In the argumentative task condition, participants were instructed to form an opinion about a controversy on the Panama topic. In the narrative task condition, they wrote a description of the events.

In the active condition, participants were able to navigate through the texts via the text menu (see Figure 2) in a self-directed order. The static navigation format allowed navigation only via "back" and "forward" buttons; thus, the text order was fixed. The text menu was always presented on the left-hand side of the screen; the text frame was on the right-hand side. The actual position within the hypertext was marked with a colored background. In both presentation format conditions, the list of facts and the task instructions could be reached by clicking on the corresponding link. Including the task instruction page and the list of facts, 15 text nodes were presented in all, because some of the texts consisted of two text nodes.

Apparatus

The study was carried out in a computer lab at Chemnitz University of Technology, in which up to 20 participants at a time took part in each session.

The tools DEWEX, Chemnitz LogAnalyzer, and SummTool supported every phase of the study. The experiment was set up with DEWEX, an environment for developing Web-based experiments (Naumann, Brunstein, & Krems, 2007). All of the instructions, questionnaires, texts, and tests were constructed with DEWEX. The abilities to generate text structure overviews and to use the "post"

Figure 2. Text menu for the static presentation format. In the active presentation format, no back and forward buttons were available.

method of communication between browser and Web server were essential to the present study. DEWEX also recorded text reading order in log files.

For the data analysis we used the Chemnitz LogAnalyzer (Brunstein, Naumann, & Krems, 2005) and SummTool (to be published). The Chemnitz LogAnalyzer allows easy interpretation of the log files with a quick glance as well as detailed analysis. The identification and visual (graphical) presentation of group-specific navigation strategies made it easy to determine whether participants in the active condition used a nonlinear navigation strategy (see Figure 3 for an example). The log files were transformed into an SPSS-ready format with SummTool.

Procedure

The participants attended one 2-h session. The experiment began with a short overview of the test procedure, and the experimenter explained the relevant hardware and software. Subsequently, all of the

Figure 3. Example of a navigation strategy in the active condition visualized with the Chemnitz LogAnalyzer.

Table 1
Percentages of Causal Connections Included in the Essays, by Task and Presentation Format

Task	Presentation Format	Causal Connections	
		<i>M</i>	<i>SD</i>
Argumentative	Active	40.67	13.52
	Static	32.37	16.13
Narrative	Active	17.05	7.43
	Static	21.90	8.81

Table 2
Mean Percentages of Borrowed, Transformed, Added, and Irrelevant Sentences in the Essays, by Task

Task	Borrowed Sentences		Transformed Sentences		Added Sentences		Irrelevant Sentences	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Argumentative	44.97	19.51	3.33	7.55	31.82	18.75	19.53	15.79
Narrative	56.38	16.73	18.45	13.57	21.89	19.16	3.28	5.15
Mann-Whitney <i>U</i>	603.5***		1,024.5***		1,167.5***		659***	
Cohen's <i>d</i>	1.39		0.63		0.52		1.36	

*** $p < .001$.

instructions, questionnaires, texts, and tests were presented on the computer screen. The input devices were the mouse and keyboard. Participants were first asked to complete pretest questionnaires. Next, the experimenter explained the reading task. Participants read the short introduction first; when finished, they clicked "Next" to go to the text menu. Participants were instructed to read every text at least once. After 35 min, they were automatically forwarded to the posttest by DEWEX.

RESULTS

In the present study, we investigated the influence of the factors *task*, *presentation format*, and *type of information* (source) on learning from multiple sources. For all results presented, significant differences have $p < .05$. The pretest (control) variables showed no effect. A log file analysis using the Chemnitz LogAnalyzer showed that the majority of participants reading the text in the active hypertext presentation format first read the text in a linear sequence from beginning to end. During the second reading, 51 of 59 participants in the argumentative group chose a nonlinear reading sequence. Thus, if the active presentation format was available, it was used by most of the participants. Figure 3 shows the navigation strategy of 1 participant in the active condition. In the static condition, most participants read the texts in a linear order one or two times from beginning to end.

Influence of the Task

As expected, more causal connectives were used by the argumentative group than by the narrative group (Mann-Whitney $U = 584.5$, $p_{\text{one-tailed}} = .000$, $d = 1.36$; see Table 1).

Further differences between the narrative and the argumentative group were shown for the origin of the information as well as for the organization of the essays (see Table 2): The narrative group wrote essays with a higher

percentage of borrowed and transformed sentences. The essays written under the argumentative task, in contrast, included a higher percentage of added and irrelevant information.

These findings do not provide clear evidence of an advantage for one of the two tasks. However, regarding the text structure, a significant influence of the task was observed [Pearson's $\chi^2(1, N = 119) = 83.02$, $p = .000$, $w = .84$]. Except for two essays, all essays in the narrative group showed a linear structure. In the argumentative condition, in contrast, all except 8 of the participants stated a thesis and provided support for it. This is an indication of a deeper understanding of the content.

No significant effect was found for the inference verification task, nor did the sentence verification task show any effect for the sentences that appeared in the texts. However, for the sentences that were not in the texts, contrary to our expectations, the argumentative group ($M_{\text{arg}} = 3.56$, $SD_{\text{arg}} = 1.12$) performed better than the narrative group [$M_{\text{nar}} = 3.01$, $SD_{\text{nar}} = 1.11$; $t(117) = 2.67$, $p_{\text{two-tailed}} = .009$, $d = .49$].

To enable us to measure the factual knowledge participants had gained, participants answered 5 open-ended questions and 21 multiple-choice questions. As expected, the argumentative group ($M_{\text{arg}} = 3.15$, $SD_{\text{arg}} = 1.29$) showed slightly poorer performance for answering open-ended questions correctly than did the narrative group [$M_{\text{nar}} = 3.61$, $SD_{\text{nar}} = 1.21$; $t(120) = 2.05$, $p_{\text{one-tailed}} = .068$, $d = .27$]. The multiple-choice questions were answered equally well by both groups. However, concerning the source information, the results show clear differences. Figure 4 illustrates the percentages of essays with the respective references.

The argumentative group paid significantly more attention to the source information: In the argumentative group only one eighth of the essays were without any

Figure 4. Percentages of essays containing any reference, at least one general reference, and at least one specific reference, by task.

reference, as compared with approximately one third in the narrative condition [Pearson's $\chi^2(1, N = 119) = 6.76$, $p_{\text{one-tailed}} = .008$, $w = .23$]. Also, general references [Pearson's $\chi^2(1, N = 119) = 17.458$, $p_{\text{one-tailed}} = .000$, $w = .38$], as well as specific references [Pearson's $\chi^2(1, N = 119) = 4.37$, $p_{\text{one-tailed}} = .018$, $w = .19$], were used in more essays when an argumentative task was given. Furthermore, the argumentative group cited more documents explicitly [$M_{\text{arg}} = 1.05$, $SD_{\text{arg}} = 0.67$; $M_{\text{nar}} = 0.71$, $SD_{\text{nar}} = 0.52$; $t(117) = 3.13$, $p_{\text{one-tailed}} = .001$, $d = .56$]. That means that considering the source information of multiple texts is supported by argumentative tasks.

Influence of the Presentation Format

Concerning the connections in the essays, a nonparametric two-way ANOVA (Kubinger, 1986) showed no main effect. However, in support of our hypothesis, we found a significant interaction effect between task and presentation format: The combination of argumentative task and active presentation format led to the highest number of causal connections [$\chi^2(1, N = 119) = 8.5$, $p_{\text{one-tailed}} = .002$]. The percentages of causal connectives are presented in Table 1.

For the transformed sentences, no significant main effects were evident. Moreover, performances in the inference and sentence verification task were similar for all groups. Only for the sentences not in the text was an interaction effect found [$F(1,115) = 3.74$, $p_{\text{one-tailed}} = .028$, $\eta_p^2 = .031$]. In Table 3, the means and standard deviations for each group are presented. According to these results, the narrative task group benefited from the static presentation format.

No main effect of the presentation format was found for factual knowledge. But as hypothesized, an interaction effect was observed for the multiple-choice questions

[$F(1,119) = 5.03$, $p_{\text{one-tailed}} = .014$, $\eta_p^2 = .042$]. The group with the argumentative task showed the best performance in the active condition; in contrast, the narrative group achieved the best results with the static presentation format. This interaction effect is depicted in Figure 5. However, the effect size indicated only a small interaction effect.

Influence of Type of Documents

The type of available documents had no influence on measures regarding the mental representation of text content.

DISCUSSION

The aim of the present study was to investigate factors having an influence on learning with multiple hypertexts. In this study, these factors were the type of task, presentation format, and type of document sources. The theoretical base of these assumptions was the documents model of Britt and colleagues (Britt et al., 1999; Perfetti et al., 1999).

Influence of Task

On the basis of previous research (Hemmerich & Wiley, 2002; Wiley, 2001) we assumed that an argumentative task requires active processing and therefore leads to better comprehension of text material and a more causally structured mental representation than a narrative task. The results predominantly confirmed our assumption. An argumentative task resulted in more causal connections,

Table 3
Means of Correct Answers in the Sentence Verification Task, by Task and Presentation Format

Task	Presentation Format	Sentences in Text		Sentences Not in Text	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Argumentative	Active	3.83	1.07	3.90	1.14
	Static	3.89	0.83	3.21	0.99
Narrative	Active	3.47	1.17	2.97	1.13
	Static	4.09	0.93	3.06	1.11

Figure 5. Means of correct multiple-choice answers, by task and presentation format.

more essays with a nonlinear and analytical structure, and more additional information integrated into the essays. It seems that participants in the argumentative group integrated more world knowledge in their mental representation, which in turn led to a more elaborated situation model. Contrary to our assumptions, the argumentative task also resulted in more correctly recognized sentences. This could mean that also for argumentative tasks, a correct mental representation of the surface code is built. However, the result could also be explained by a ceiling effect, because both groups performed very well. For factual knowledge, there are only marginal differences between the narrative and the argumentative groups. This is in line with the current state of research: The acquisition of factual knowledge is comparable or only slightly better with narrative tasks than with argumentative tasks (Hemmerich & Wiley, 2002; Stahl et al., 1996). The results of our study show that the advantage of argumentative tasks lies primarily in better source processing. This supports our hypothesis (based on Britt et al., 1999, and Perfetti et al., 1999) that the construction of an intertextual model for the mental representation of text information works better for argumentative tasks than for narrative ones.

Influence of Task and Presentation Format

The hypothesis that argumentative tasks profit from active presentation formats and narrative tasks from static presentation formats was supported, even though the effects were not as strong as in previous studies (Voss & Wiley, 2000; Wiley, 2001). For the combination of argumentative task and active hypertext, mental representations showed the most causal connections; the combination with static presentation format showed the least. Performance in sentence recognition was, according to our assumptions, high for the combinations argumentative/active and narrative/static. Also in line with the hypothesis, for factual knowledge (multiple-choice questions), the best performance was achieved with the combination narrative/static. However, regarding factual knowledge, the small effect size indicated that the practical concern of this result can be put into question.

Influence of Document Sources

Contrary to our assumptions, the construction of a mental representation was not affected by the type of available source material. No advantage for primary sources was shown. One explanation could be the participants' lack of experience with primary sources. They might have rated both types of sources as equivalent and might not have seen the advantage of primary sources. This is supported by the result that specific source information was only minimally considered in the essays.

Future Prospects

Our results support the assumption that argumentative tasks help in constructing a mental representation. However, the uncertainty remains that essay structure might be influenced much too heavily by the task. Participants in the narrative condition wrote only a listing of

facts because that complied with their instructions. But it does not necessarily reflect their mental representation. The design of a follow-up study should take this problem into account. Furthermore, regarding the consideration of source information, additional measurements are possible. For example, the construction of an accurate intertextual model could be tested with a task in which statements or text content need to be assigned to their authors. Also, the present study examined only purely text-based documents. Further studies could investigate how the results can be replicated for hypermedia documents with additional multimedia elements.

AUTHOR NOTE

This article was presented at the Society for Computers in Psychology conference in Long Beach, CA, on November 15, 2007. We thank Tobias Winkler for programming the DEWEX tool and Jacqueline Wanick and Angela Brunstein for their help with the project "User Oriented Presentation of Information on the Internet." We also thank two anonymous reviewers for their helpful comments. Correspondence concerning this article should be addressed to A. B. Naumann, Deutsche Telekom Laboratories, Ernst-Reuter-Platz 7, D-10587 Berlin, Germany (e-mail: anja.naumann@telekom.de).

REFERENCES

- BRITT, M. A., PERFETTI, C. A., SANDAK, R., & ROUET, J.-F. (1999). Content integration and source separation in learning from multiple texts. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 209-233). Mahwah, NJ: Erlbaum.
- BRUNSTEIN, A., NAUMANN, A. [B.], & KREMS, J. F. (2005). The Chemnitz LogAnalyzer: A tool for analyzing data from hypertext navigation research. *Behavior Research Methods*, 37, 232-239.
- CONKLIN, J. (1987). Hypertext: An introduction and survey. *Computer*, 20, 17-41. doi:10.1109/MC.1987.1663693
- EINSTEIN, G. O., MCDANIEL, M. A., OWEN, P. D., & COTÉ, N. C. (1990). Encoding and recall of texts: The importance of material appropriate processing. *Journal of Memory & Language*, 29, 566-581. doi:10.1016/0749-596X(90)90052-2
- GOLDMAN, S. R. (2004). Cognitive aspects of constructing meaning through and across multiple texts. In N. Shuart-Faris & D. Bloome (Eds.), *Uses of intertextuality in classroom and educational research* (pp. 317-351). Greenwich, CT: Information Age Publishing.
- GRAESSER, A. C., MILLIS, K. K., & ZWAAN, R. A. (1997). Discourse comprehension. *Annual Review of Psychology*, 48, 163-189. doi:10.1146/annurev.psych.48.1.163
- HEMMERICH, J., & WILEY, J. (2002). Do argumentation tasks promote conceptual change about volcanoes? In *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society* (pp. 453-458). Hillsdale, NJ: Erlbaum.
- JOHNSON-LAIRD, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
- KINTSCH, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- KUBINGER, K. D. (1986). A note on non-parametric tests for the interaction in two-way layouts. *Biometrical Journal*, 28, 67-72. doi:10.1002/bimj.4710280113
- MCDONALD, S., & STEVENSON, R. J. (1998). Effects of text structure and prior knowledge of the learner on navigation in hypertext. *Human Factors*, 40, 18-27. doi:10.1518/001872098779480541
- MCMANARA, D. S., KINTSCH, E., SONGER, N. B., & KINTSCH, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition & Instruction*, 14, 1-43. doi:10.1207/s1532690xci1401_1
- NAUMANN, A. [B.] (2003). *Wissenserwerb und Informationssuche mit Hypertexten: Die Bedeutung von Strukturierung, Navigationshilfen und Arbeitsgedächtnisbelastung* [Knowledge acquisition and informa-

- tion retrieval with hypertext: The impact of structure, navigation aids, and working memory load]. Unpublished dissertation, TU Chemnitz. <http://archiv.tu-chemnitz.de/pub/2004/0117>.
- NAUMANN, A. [B.], BRUNSTEIN, A., & KREMS, J. F. (2007). DEWEX: A system for designing and conducting Web-based experiments. *Behavior Research Methods*, **39**, 248-258.
- NAUMANN, A. [B.], WANIEK, J., & KREMS, J. F. (2001). Knowledge acquisition, navigation and eye movements from text and hypertext. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet science* (pp. 293-304). Lengerich: Pabst Science.
- PASCH, R., BRAUßE, U., BREINDL, E., & WAßNER, U. H. (2003). *Handbuch der deutschen Konnektoren* [Handbook of connectors in the German language]. Berlin: de Gruyter.
- PERFETTI, C. A., BRITT, M. A., & GEORGI, M. C. (1995). *Text-based learning and reasoning: Studies in history*. Hillsdale, NJ: Erlbaum.
- PERFETTI, C. A., ROUET, J.-F., & BRITT, M. A. (1999). Toward a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 99-122). Mahwah, NJ: Erlbaum.
- ROUET, J.-F., BRITT, M. A., MASON, R. A., & PERFETTI, C. A. (1996). Using multiple sources of evidence to reason about history. *Journal of Educational Psychology*, **88**, 478-493. doi:10.1037/0022-0663.88.3.478
- ROUET, J.-F., FAVART, M., BRITT, M. A., & PERFETTI, C. A. (1997). Studying and using multiple documents in history: Effects of discipline expertise. *Cognition & Instruction*, **15**, 85-106. doi:10.1207/s1532690xci1501_3
- SINGER, M., GRAESSER, A. C., & TRABASSO, T. (1994). Minimal or global inference during reading. *Journal of Memory & Language*, **33**, 421-441. doi:10.1006/jmla.1994.1020
- STADTLER, M., & BROMME, R. (2005). Dealing with multiple documents on the WWW: The role of metacognition in the formation of documents models. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 2068-2073). Mahwah, NJ: Erlbaum.
- STAHL, S. A., HYND, C. R., BRITTON, B. K., MCNISH, M. M., & BOSQUET, D. (1996). What happens when students read multiple source documents in history? *Reading Research Quarterly*, **31**, 430-456. doi:10.1598/RRQ.31.4.5
- VANSLEDRIGHT, B. [A.] (2002). *In search of America's past: Learning to read history in elementary school*. New York: Columbia University, Teachers College Press.
- VOSS, J. F., & WILEY, J. (2000). A case study of developing historical understanding via instruction: The importance of integrating text components and constructing arguments. In P. N. Stearns, P. Seixas, & S. Wineburg (Eds.), *Knowing, teaching, and learning history: National and international perspectives* (pp. 375-389). New York: New York University Press.
- WILEY, J. (2001). Supporting understanding through task and browser design. In *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 1136-1143). Hillsdale, NJ: Erlbaum.
- WILEY, J., & VOSS, J. F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology*, **91**, 301-311. doi:10.1037/0022-0663.91.2.301
- WINEBURG, S. S. (1994). The cognitive representation of historical texts. In G. Leinhardt, I. L. Beck, & C. Stainton (Eds.), *Teaching and learning in history* (pp. 85-135). Hillsdale, NJ: Erlbaum.
- WINEBURG, S. [S.] (2000). Making historical sense. In P. N. Stearns, P. Seixas, & S. Wineburg (Eds.), *Knowing, teaching, and learning history: National and international perspectives* (pp. 306-325). New York: New York University Press.
- ZWAAN, R. A., & RADVANSKY, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, **123**, 162-185. doi:10.1037/0033-2909.123.2.162

(Manuscript received November 8, 2007;
revision accepted for publication July 21, 2008.)