UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Supporting Understanding Through Task and Browser Design

Permalink
https://escholarship.org/uc/item/09q8p7gn

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 23(23)

ISSN
1069-7977

Author
Wiley, Jennifer

Publication Date
2001

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Supporting Understanding through Task and Browser Design

Jennifer Wiley (jwiley@uic.edu)
Department of Psychology, 1007 W. Harrison Street (M/C 285)
Chicago, IL 60607 USA

Abstract

While electronic text offers the potential to explain, illustrate, and scaffold understanding in powerful new ways, few studies on educational use of electronic text resources have shown significant learning gains, or even measured learning outcomes in controlled experiments (Chen & Rada, 1996; Dillon & Gabbard, 1998). In a follow-up to previous studies (Wiley & Voss, 1999), the present experiments study the effects of different tasks and browser designs on navigation and reading patterns, as well as on memory and comprehension measures from electronic text. These studies have revealed that only when both the task and the design support integration (such as in a two-windowed browser) and students are explicitly directed how to use the feature, do students take advantage of the flexibility of the multiple-source environment, integrate across sources, and achieve the best level of understanding.

Introduction

One promise of using electronic text in the classroom is the potential for students to search for, access and read multiple forms of information about a topic. Since the search for and navigation of digital documents is student-initiated, requires student interaction, captures the student's interest, proceeds at the student's own pace, and allows for flexible navigation and juxtaposition of multiple sources, a number of theorists have suggested that the web might be a powerful tool for student instruction (Beeman, et al 1987, Spiro & Jehng, 1990). This optimism is consistent with a number of recent cognitive studies demonstrating that activities that require readers to engage in active, constructive and integrative tasks lead to the best understanding of the subject matter (e.g., Chi, de Leeuw, Chiu & LaVancher, 1994; McNamara, et al., 1996).

However, a review of the literature on educational use of electronic text yields two striking conclusions. First, students generally fail to utilize hypertext links and multiple window capabilities effectively, if at all, as they read (Foss, 1989; Gordon, et al., 1988). This is especially true of novice users (Foltz, 1996; Gray & Shasha, 1989; Tombaugh, Lickorish & White, 1987). And second, few studies on educational use of electronic documents, whether from stand-alone hypermedia or the World Wide Web, have actually shown significant learning gains (Chen & Rada, 1996; Dillon & Gabbard, 1998). What is critically needed is for experiments to determine which specific instructional contexts may allow for effective educational use of electronic text.

Although there has been a great deal of evaluation on effective browser design from a Human-Computer Interaction (HCI) standpoint, effectiveness has been measured largely in terms of efficiency of search or ease in information finding (Dillon & Gabbard, 1998). While such fluency measures may be related to some extent to the amount of information a person is able to recall after reading from computer screens, they may not be correlated with whether a person develops an understanding of the text that is being read. A number of studies have shown that conditions that produce the best surface memory for text are not the best conditions for producing the best understanding of text (e.g. McNamara, et al., 1996; Mayer, 1999; Wiley & Voss, 1999). While surface memory may correlate with the fluency or ease of information processing, understanding may depend to some extent on the need to put effort into developing an underlying representation or situation model of the text (Kintsch, 1998). Thus, previous assessments from an HCI perspective cannot reliably indicate which screen layouts will be most effective for promoting understanding from electronic text. Given this educational goal, browser design must be evaluated specifically using measures of conceptual learning.

In a review of the published studies on hypermedia and learning outcomes between 1990 and 1996, Dillon and Gabbard found only 11 studies that performed controlled experiments on hypermedia and learner performance. Of those 11 studies, there were only four results that actually seemed to suggest an advantage for learning from hypermedia over paper. In the majority of studies there was no clear difference between learning from hypermedia and learning in a control (more traditional) setting. While this may be viewed more optimistically as evidence that learning from electronic resources may sometimes be no worse than traditional classroom methods, there is hardly an overwhelming body of evidence that the web can generally be relied upon to provide an enriching instructional experience.

Of the four studies that netted positive learner performance, only two seem to indicate that hypertext may allow students to engage in learning at a more conceptual level. One of these reported that students learned to recognize aircraft more efficiently and effectively when they were able to view the images next to each other (and even overlay two images) in a browser during learning (Psotka, Kerst & Westerman, 1993). Clearly the ability to juxtapose and overlay images gave subjects a better concept of the aircraft prototypes, and this allowed them to perform well on later recognition tasks. In a second study, Jacobson and Spiro (1995) found that a hypermedia environment as opposed to a linear electronic presentation of the same materials allowed for the best performance on a problem-solving essay task. Interestingly, the linear presentation of the same material allowed for the best recall of the facts. Unfortunately, no converging evidence of better comprehension in the hypertext vs. linear condition was obtained. However, the differences in essay quality suggest that students in the hypertext conditions benefitted from the flexibility and ability to jump between sources in the hypertext format, allowing for better synthesis of the material that was presented. This result is consistent with perhaps the most consistently cited study on learning in hypertext (Egan et al, 1989) which found that a multi-window environment called Superbook led to better essay writing than a paper control condition.

Based on their review, and these last two studies in particular, Dillon and Gabbard suggest that hypermedia may afford particular advantages for learners on tasks that require comparison across sources. Notably, when students are not given the ability to view two sources at once in a computer environment, they perform less well than with paper on tasks which require integration across two documents (Wang & Liebscher, 1988). Consistent with the intuition provided by these previous studies, Wiley & Voss (1999) reported that students can show better conceptual learning from a web-like environment when they are provided with multiple windows and are given a task that requires them to integrate information across sources.

In one of few empirical studies evaluating conceptual learning from a web-like environment, Wiley and Voss (1999) demonstrated that reading multiple sources presented in a browser can lead to better understanding of subject matter than reading from a textbook. In this study, students read about the Irish Potato Famine either from several on-line documents in a two-window web site or they read the same information in the form of a textbook chapter. When students were asked to write an argument of "What produced the significant changes in Ireland's population" instead of a narrative, and read the on-line sources instead of the textbook chapter, students gained the best understanding of the material. Understanding was assessed by the causal and integrated nature of their essays, as well as their performance on inference verification and analogy identification tasks. Wiley and Voss (1999) concluded that tasks which require students to construct their own representations of a situation will yield the most conceptual learning in web-like environments. The argument writing task promoted understanding because it required students to integrate information from across the multiple sources as they created support for a thesis. And, the multiple source condition may have promoted understanding by supporting the comparison and integration of the individual sources.

This is an important finding, demonstrating empirically that electronic text can be an effective tool for developing student understanding. However, it is important to note the very specific circumstances under which better understanding may be achieved from web-like environments. For one, only students given a task requiring integration of information across sources showed better learning from a browser. Otherwise, learning from a browser was actually poorer than from a textbook. In fact, there were a number of particular features of the Wiley & Voss (1999) environment, any of which may be important in order for the effect to be obtained:

The site had a small set of documents.
The documents were selected for the user.
The documents were largely relevant.
There were no links embedded in the texts.
Each document fit in a single window.
The task was well-defined and specific.
The task required integration across sources.
The browser had two side-by-side windows.
The overview menu was accessible through an icon.
Images were presented in their own window.
Students were instructed to use both windows.
Students were instructed how to use the menu icon.
Conceptual learning was assessed in the post-test.

The present experiments directly test whether the design of the browser with two side-by-side windows might have been critical for the better learning in the web source/argument writing task condition. Although many computerized tutors and interactive environments use multiple windows, there has been little work on how students use multiple windows or the optimal conditions for multiple window use (von Oostendorp, 1996). Interestingly, the three other studies that suggest that students can gain better conceptual understanding from hypermedia environments (Egan, et al., 1989; Pstoka, et al 1993, Jacobson & Spiro, 1993) all used a multi-window display. However, there have been no studies that have manipulated the number of windows and directly measured comprehension. With converging data from essay tasks, comprehension tests and eyetracking protocols, the present studies address whether a multi-window browser supports better understanding in a web-like environment.

Experiment 1

The first experiment investigated the effect of two-window browsers on learning historical subject matter from a web site. This experiment tested the hypothesis that the design of the browser had an impact on students' understanding. Thirty undergraduates were asked to read 10 pages from a web site about the Irish Potato Famine in order to write either a narrative or an argument of what produced the significant changes in Ireland's population. (The pages contained 5 texts (about 1500 words), 4 tables, 1 graph and 1 map.) In addition, students either read the information from a single-window browser, a two-window browser, or from a two-window browser with specific instructions about why they were being given two windows. Further in this third condition the list of documents was split across the two windows, so that in order to read all of the information readers had to use both windows.

Student learning was assessed with a number of learning measures taken from Wiley and Voss (1999). Of most interest are three measures thought to reflect understanding: the proportion of sentences in student essays which represented an integration or "transformation" of the presented information (as opposed to simply copying the presented information), an inference verification task and an analogy task. The inference verification task (IVT) contained 10 inferences that could be generated by integrating information across two sources, such as "As rent costs increased, emigration from Ireland increased." as well as 10 distractors. The analogy task consisted of short descriptions of potentially analogous events. Students were asked to rate the similarity of the causes of each event with the population decline in Ireland. These analogies were intended to vary in surface and deep similarity, and the critical analogy, the institution of a Poll Tax in the U.S. South after the Civil War, was intended to be similar only on a deep level (as it was related to sociopolitical inequities and class power struggles, but there was no large-scale loss of life). Thus, recognition of the Poll Tax as causally similar to the changes in Ireland's population indicates a particularly good understanding of the text. This rating serves as the critical analogy task (CAT).

The means for each condition on each measure are presented in Figure 1. Performance on all measures was better when students wrote arguments as opposed to narratives (TRSENT: $\underline{F}(1,24)=14.9$, $\underline{p}<.01$; IVT: $\underline{F}(1,24)=10.9$, $\underline{p}<.01$; CAT: $\underline{F}(1,24)=11.23$, $\underline{p}<.01$). The main effect for number of windows only reached significance for the inference task, $\underline{F}(2,24)=5.3$, $\underline{p}<.01$, as did the interaction, $\underline{F}(2,24)=7.89$, $\underline{p}<.01$. However, the interaction approached significance for the proportion of transformed sentences ($\underline{p}<.15$) as did the main effect for number of windows on the critical analogy task ($\underline{p}<.13$). Pairwise comparisons based on a priori hypotheses revealed that the two-window/ argument writing condition significantly outperformed

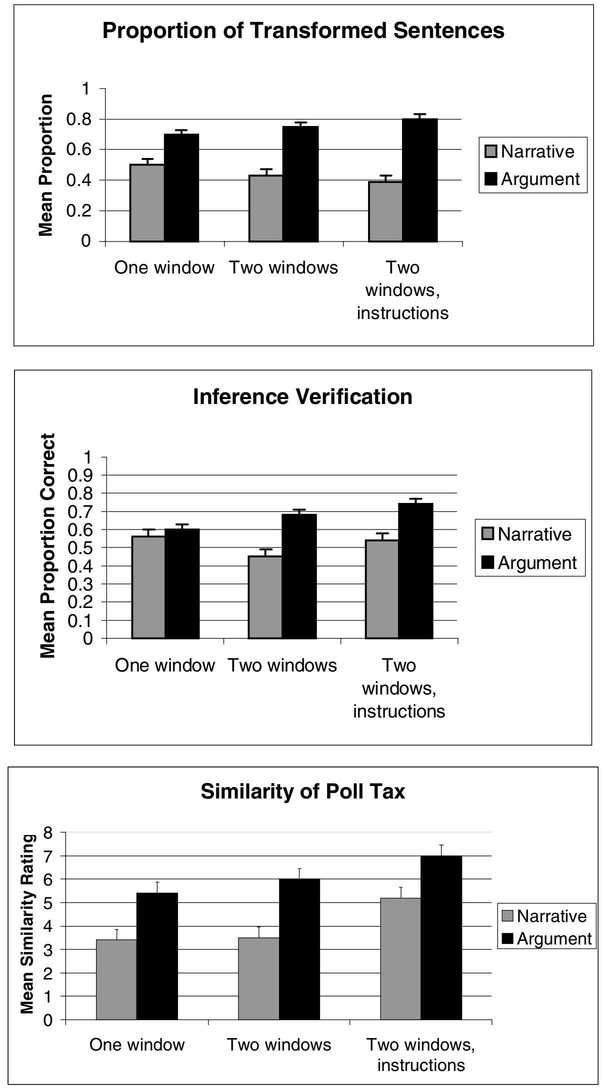

Figure 1: Experiment 1 Results.

the single window/narrative writing conditions on all measures ($\underline{p}s<.05$), but only when students were given specific instructions.

Two other sources of data may be of interest here. The first is anecdotal evidence from a failed pilot study showing that students are quite reluctant to use a two-window browser. In the first design of the two-window browser, I failed to set the frame so that it could not be removed. The first thing that students did when they were seated in front of a computer with the two-window browser was click and drag that frame off the screen so they had one big window. It is important to note that I designed all my materials so that they would fit in a single side-by-side window. Thus, it was not that students could not see all of the documents that prompted them to want a bigger screen. Their initial preference was for a single window. Further, even once the two-window browser was set so students couldn't remove the frame, many of them still failed to ever use the second window. Instead, many students kept the graph in the right window and read through the sources on the left. Only when students were given the specific instructions and forced to use both windows did all of the students actually use both windows.

Previous work with multiple window environments is consistent with these findings. Foss (1989) found that about a third of people generally prefer simple screens and only having a single window open. This prevents students from making important connections across sources. Another third of readers did open more than one window at once, but they did not use the multiple windows effectively. Their screens quickly became messy and busy. Only a third of Foss' subjects were able to use multiple windows effectively for a search and decision making task. Tombaugh et al (1987), also found a general preference for and better search in a single window environment, and found that only with practice in a multiple window environment were participants able to use the overlapping multiple windows efficiently. Part of these results may be due to the overlapping window environment that was used, however, as Instone, et al (1996) found that participants were able to use a tiled multiple-window environment more efficiently than an overlapping window display. Unfortunately, all of these previous experiments measured the effects of the windows in terms of speed and accuracy of search for information, and not in terms of comprehension. Only the present study has manipulated the number of windows and investigated how the use of multiple windows can lead to a better conceptual understanding of the subject matter, as it allows for the concurrent presentation of related concepts from different sources.

A second additional source of information about how readers behaved in the different conditions comes from analysis of browsing logs and eyetracking data. These sources indicated that both general instruction about how to use web sources and specific writing instruction yielded different navigation and reading patterns. A pilot eyetracking study on 4 students in the 2-window/2-list condition gives us a better idea of exactly how the students used these windows. Two of these students were told to read with the purpose of writing a narrative, and the other two were told the argument instruction. Like all other students, these readers tended to begin their task by reading through each document, one at a time, in order. Students usually simply went down the documents in the list on the left side of the screen and then down the documents on the right side of the screen. During this initial reading phase, the eyes rarely left the document that was being read. At this point, both readers in the narrative condition stopped reading and declared they were ready to write their essay. On the other hand, both students in the argument condition moved on to a second phase of reading.

One student started over again from the beginning, starting at the top of the left window list and skimmed the documents in order. But, from time to time, she would call up documents in the other window, or skip to another document on the same list and look for particular sentences. The other student in the argument condition began the second phase by calling up pairs of documents and alternated reading between the two. Importantly, the sentences that these students tended to re-read were important for inferences about the causes of the Potato Famine. Eyetracking data revealed that students in argument condition spent more total time on sentences important for inferences. Thus, the selection patterns and eyetracking evidence suggest that under some conditions web sites can promote more active reading patterns, suggesting more active integration of the text at a conceptual level. Further, this second phase of reading, or review of documents, seems to be particularly critical for understanding. It is what students do when they are reading from paper documents, and what Dee-Lucas and Larkin (1995) found when students effectively used structured overviews in electronic text.

Taken together, these sources of data demonstrate that students need to have both a task and an environment that forces them to be more active in order for students to gain the benefits of web resources. Only when both the task and the design support integration, and students are explicitly directed how to use the feature, do students take advantage of the flexibility of the multiple source environment. Only then do students integrate across sources, selectively re-read sources, and achieve the best level of understanding.

The second experiment is an important extension of this work using scientific texts as content. Of particular interest is whether the ability to juxtapose two documents, while performing a task that supports integration, will allow for better understanding of scientific concepts as well.

Experiment 2

Although there may be some differences between reading from history and science-related text, when readers must connect information across documents, in order to make inferences or construct global models of causality, then simultaneous presentation of the sources that need to be linked should help regardless of the subject matter. Thus, the second experiment investigated the effect of two-window browsers on learning from a scientific web site. This experiment tested the hypothesis that the design of the browser has an impact on student's scientific understanding. Forty undergraduates were asked to read 16 pages from a web site about Earthquakes and Volcanoes (based on sources from the USGS web site) in order to write either an essay or an argument of "What caused the explosion of Mt. St. Helens?" (The pages contained 10 texts (about 3000 words), 4 diagrams, 3 maps, and 2 photographs.) In addition, students either read multiple sources in a single-window browser, or multiple sources in a two-window browser with specific instructions about why they were being given two windows. Further in the two-window condition the list of documents was split across both windows, so that in order to read all of the information readers had to use both windows. This yielded a 2x2 (writing task x presentation format) design with 10 students in each of the conditions.

Student learning was assessed with a number of learning measures similar to those used in Wiley & Voss (1999) and in Experiment 1. The same 3 measures of understanding are reported as in Experiment 1: proportion of transformed sentences (TRSENT), the inference verification and analogy tasks. The inference verification task (IVT) contained 10 inferences that could be generated by integrating information across two sources, such as “Volcanoes are likely to develop where continents collide” as well as 10 distractors. The critical analogy task (CAT) asked students to rate the similarity of the causes of the Kobe earthquake with the causes of the Mt. St. Helens eruption. The Kobe earthquake was intended to be similar only on a deep level (as it was related to disturbance due to subduction of a tectonic plate, but there was no volcanic activity). Thus, recognition of Kobe as causally similar to Mt. St. Helens indicates a particularly good understanding of the text, and the relation between plate tectonics and volcanic activity.

The means for each condition are presented in Figure 2 for the 3 tasks. As in Experiment 1, performance on all tasks was better when students wrote arguments (TRSENT: $F(1,36)=8.96$, $p<.01$; IVT: $F(1,36)=9.06$, $p<..01$; CAT: $F(1,36)=8.81$, $p<.01$). The main effect for number of windows only reached significance for the proportion of transformed sentences, $F(1,36)=6.82$, $p<.01$.

While none of the interactions were significant, based on previous results and a priori hypotheses, pairwise comparisons were performed and revealed that the two-window/argument condition outperformed the essay conditions on all measures ($ps<.08$). However, it is notable that although there was a trend toward a main effect for number of windows on the inference task ($p<.14$), there were no trends in either the windows effect or the interaction on the critical analogy task. This suggests that while the best essay writing may have occurred due to combination of writing task and two-window browser, for the learning measures, the writing task alone was responsible for better understanding.

Conclusions and Implications

In both experiments, when both the task and the browser design supported integration, and there was explicit instruction how to use the two-window browser, students were able to write more integrated essays in a multiple window environment. This condition also led to the best conceptual learning in both experiments. Although in Experiment 1 it seemed

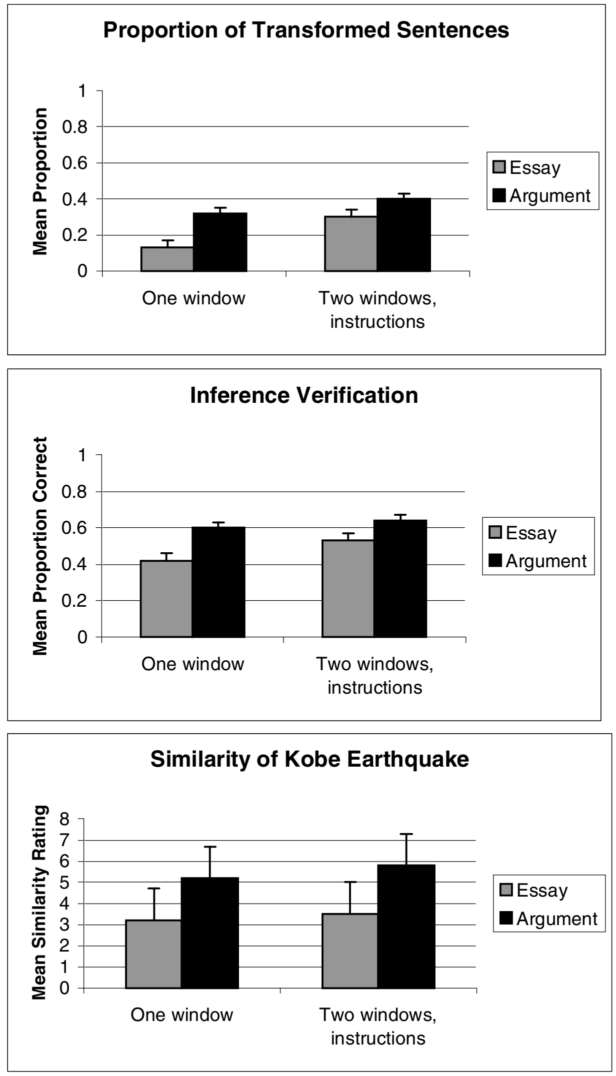

Figure 2: Experiment 2 Results.

both the writing task and the number of windows contributed to the effect, in Experiment 2, the effect of the two-window browser on the understanding of the subject matter was not as evident. The argument writing task, however, did support better understanding in both history and science.

There are many possible reasons for the differential effects of the browser design in these two studies. The most obvious are the differences in the materials that students were presented with. More information was presented in the scientific site. Further, looking at the single window/narrative essay condition as a baseline, we can see that the scientific tests were more difficult for students. Students accurately recognized 55% of inferences in the history test, whereas they recognized only 40% of inferences correctly in the science site. Further, different kinds of images were used in the history and science sites. These findings have led to several interesting questions that we are currently pursuing (Wiley, Ash, Brodhead & Sanchez, 2001). Do students simply respond to history and science subject matter differently? Are images processed differently in the two domains? Did the different types of images in the two sites lead to learning differences? Or, is the difficulty of the subject matter driving these processing differences?

Even though we are still looking for the best task/environment combination for conceptual learning from scientific web sites, taken together, the present studies demonstrate that specific conditions are necessary for effective educational use of electronic texts. In order for conceptual learning to occur, readers of electronic text may need a multimedia environment that promotes integration of the presented information and certainly need a task that prompts integration across sources. Only through the specification and demonstration of which computerized learning environments lead to better understanding, may we begin to realize some of the educational potential of electronic text.

Acknowledgments

This work was supported by grants from the Office of Naval Research and the Paul G. Allen Virtual Education Foundation. Thanks to Linda Gentry, Ed Kearney, and Rebecca Schrader for their research assistance.

References

Beeman, W., Anderson, K., Bader, G., Larkin, J. McClard, A., McQuillian, M. & Sheilds, M. (1987) Hypertext and pluralism: from lineal to nonlineal thinking. In *Proceedings of Hypertext 87*, pp. 67-88, University of North Carolina, Chapel Hill.

Chen, C. & Rada, R. (1996) Interacting with hypertext: A meta-analysis of experimental studies. *Human Computer Interaction, 11*, 125-156.

Chi, M., de Leeuw, N., Chiu, M. & LaVancher, C., (1994). Eliciting self-explanations: improves learning. *Cognitive Science, 18*, 439-478.

Dee-Lucas, D. & Larkin, J..(1995). Learning from electronic texts: Effects of overviews for information access. *Cognition & Instruction, 13*, 431-468.

Dillon, A. & Gabbard, R. (1998). Hypermedia as educational technology: A review of the quantitative research literature on learner comprehension, control and style. *Review of Educational Research, 68*, 322-349.

Egan, D., Remde, J., Landauer, T., Lochbaum, C. & Gomez, L. (1989) Behavioral evaluation and analysis of a hypertext browser. *Proceedings of CHI 89*, 205-210.

Foss, C. (1989). Detecting lost users: Empirical studies on browsing hypertext. INRIA Tech Report 973. Valbonne, France: L'Institut National de Recherche en Informatique et en Automatique.

Foltz, P. (1996). Comprehension, coherence and strategies in text and hypertext. In Rouet, J.F. et. al. (Eds) *Hypertext and Cognition.* Mahwah, NJ: Erlbaum.

Gordon, S., Gustavel, J., Moore, J. & Hankey, J. (1988). The effects of hypertext on reader knowledge representation. *Proceedings of the Human Factors Society 32nd Annual Meeting*, 296-300.

Gray, S. & Shasha, D. (1989) To link or not to link *Behavior Research, Methods, Instruments, and Computers, 21*, 326-333.

Instone, K., Teasley, B. & Leventhal, L. (1996) Lessons learned from redesigning hypertext user interfaces. In van Oostendorp & de Mul (Eds*) Cognitive aspects of electronic text processing*. Norwood, NJ: Ablex.

Jacobson, M. & Spiro, R. (1995) Hypertext learning environments, cognitive flexibility and the transfer of complex knowledge. *Journal of Educational Computing Research*, *12*, 301-303.

Kintsch, W. (1998*) Comprehension: A paradigm for cognition*. Cambridge: Cambridge.

Lehto, M., Zhu, W. & Carpenter, B. (1995) The relative effectiveness of hypertext and text. *International Journal of Human Computer Interaction, 7*, 293-313.

McNamara, D. Kintsch, E., Songer, N., & Kintsch, W. (1996) Are good texts always better*? Cognition & Instruction, 14*, 1-43.

Psotka, J., Kerst, S. & Westerman, T. (1993) The use of hypertext and sensory-level supports in visual learning. *Behavior Research Methods, 25*, 168-172.

Spiro, R. & Jehng, J. (1990) Cognitive flexibility and hypertext. In D. Nix & R. Spiro (Eds*) Cognition, Education and Multimedia.* Hillsdale, NJ: Erlbaum.

Tombaugh, J., Lickorish, A. & Wright, P. (1987) Multi-window displays for readers of lengthy texts. *International Journal of Man-Machine Studies, 26*, 597-615.

van Oostendorp, H. (1996) Studying and annotating electronic text. In J.F. Rouet, et al (Eds.*) Hypertext and Cognition*. Mahwah, NJ: Erlbaum.

Wang, X. & Liebscher, P. (1988) Information seeking in hypertext: Effects of physical format and search strategy. *Proceedings of the ASIS Annual Meeting, 25*, 20-205.

Wiley, J., Ash, I., Brodhead, A. & Sanchez, C. (2001, May) *The impact of images in learning from web pages on science and history*. Paper presented at the Midwestern Psychological Association, Chicago, IL.

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*, 1-11.