Comprehension effects of signalling relationships between documents in search engines

Ladislao Salmerón [1], Laura Gil [1], Ivar Bråten [2] & Helge Strømsø [2]

[1] University of Valencia, Spain

[2] University of Oslo, Norway

Contact address:

Ladislao Salmerón

Department of Developmental and Educational Psychology - University of Valencia

Avda. Blasco Ibáñez, 21. 46010 Valencia, Spain

ladislao.salmeron@valencia.edu

Paper draft, please check the original reference. To appear in Computers in Human Behavior.

A key task for students learning about a complex topic from multiple documents on the Web is to establish the existing rhetorical relations between the documents. Traditional search engines such as Google ® display the search results in a listed format, without signalling any relationship between the documents retrieved. New search engines such as Kartoo ® go a step further, displaying the results as a constellation of documents, in which the existing relations between pages are made explicit. This presentation format is based on previous studies of single-text comprehension, which demonstrate that providing a graphical overview of the text contents and their relation boosts readers' comprehension of the topic. We investigated the assumption that graphical overviews can also facilitate multiple-documents comprehension. The present study revealed that undergraduate students reading a set of web pages on climate change comprehended them better when using a search engine that makes explicit the relationships between documents (i.e. Kartoo-like) than when working with a list-like presentation of the same documents (i.e. Google-like). The facilitative effect of a graphical-overview interface was reflected in inter-textual inferential tasks, which required students to integrate key information between documents, even after controlling for readers' topic interest and background knowledge.

Keywords: multiple-documents comprehension; text comprehension; graphical overviews; web search engines.

When students search the Internet to learn about a particular topic, search engines provide them with entries to multiple web documents on the topic. In this learning scenario, students' global comprehension of the topic goes beyond the comprehension of each retrieved web document. As students read through the web documents, they construct a mental representation of the topic that consolidates new information with information that they already know. This process of integrating information is inherent to learning from the Internet, because web documents are usually less complete than their corresponding paper documents (Bhavnani et al., 2003; Britt & Gabrys, 2001). Indeed, web documents usually focus on a particular issue (e.g. origins of a phenomenon, its effects…), and at best they provide links to additional sources that allow students to complete their learning assignment.

Traditional search engines place the burden of the integration process entirely on the student, because they just display a list of web documents without providing any information about the relationships between them (e.g. list-like interface by Google ®). New search engines, however, are exploring the possibility of providing students with rhetorical information about the relations between web documents (e.g. graphical interface by Kartoo ®), in order to help them integrate information from the documents. For example, upon the search query "influenza", the search engine can visually group the retrieved web documents under the labels "symptoms, causes, prevention and treatment". Currently, there is no evidence that this way of displaying web results may influence students' learning. The goal of this study was to provide insights into this issue.

Integration of multiple documents

Perfetti, Rouet, & Britt (1999) proposed a cognitive model to describe how advanced students comprehend and integrate information from multiple documents. In their view, the comprehension of multiple documents at the local level involves the same processes as the comprehension of a single text, i.e. constructing a model of the situation described in the text (Kintsch, 1998). The models that students create for each document can relate in various ways, for example containing overlapping, complementary or even contradictory information. For this reason, in the context of learning from multiple documents, students have to construct an additional representation called the Documents

Model, where they can store additional information reflecting the general situation described across the documents (Situations Model) as well as information about how the documents relate to each other (Intertext Model). The Intertext Model represents relevant information for each document, including main ideas presented in the text, information about the source (author, year, linguistic style), and the rhetorical objectives of the document (audience addressed and purpose of the author), as well as information about the existing relations between the texts.

Research in the context of multiple documents has studied several interface factors that might facilitate students' integration of information. This process can be fostered by displaying the web documents simultaneously in a two-windowed browser (Wiley, 2001), by avoiding using too many embedded links in a text (Britt & Gabrys, 2002), or by showing the web document and the student's writing file at the same time (Olive, Rouet, François & Zampa, 2008). Finally, a potential way of helping students to integrate information from different web documents consists of signalling the rhetorical relations between documents as is done by new search engines such as Kartoo ®. Although this solution has been found to foster the comprehension of related sections from a single document (e.g. Ruddell & Boyle, 1989), efforts to empirically validate this issue in the context of multiple documents are scarce.

Signalling rhetorical relations and multiple-documents comprehension

Perfetti and colleagues' model provides a useful framework for the study of multiple-documents comprehension. However, the model currently operates at a rather descriptive level, and does not provide much insight into the mechanisms that promote integration of documents at the level of the Intertext Model. Further elaboration of this issue may build on two important theoretical models that were originally proposed to account for the effects of signalling relationships with graphical overviews on single-document comprehension. The well-known Assimilation Theory of Mayer (1979) states that graphical overviews enable readers to construct an accurate mental representation of the text, as reflected in the overviews, and thus provide an organizational framework. In other words, a graphical overview increases the salience of the text structure that might

be part of the situation model representation for the text, thereby enhancing memory of the text structure. This effect appears when the content of the text is quite difficult and the students do not possess prior knowledge about the text topic (Lorch & Lorch, 1996; Salmerón, Baccino, Cañas, Madrid & Fajardo, 2009). In the context of multiple-documents comprehension, the Assimilation Theory would imply that when students read a graphical overview, it acts as an initial schema for intertextual organization, allowing the reader to incorporate subsequent information from the isolated documents into an existing representation. Readers with no previous knowledge of the topic, or those faced with difficult texts, will not be overloaded by the need to build an initial frame for the intertextual information. This structural 'preparation' might simplify the operations related to linking related information between documents, which might be reflected in a richer intertextual model. This hypothesis resembles the top-down process of integration proposed by Kurby, Britt and Magliano (2005). Those authors proposed that when students read related documents the content of the already read texts is automatically activated while reading a subsequent document. In order to determine the interrelations between documents, readers must evaluate to what extent the activated knowledge is related to the current document. Kurby et al. (2005) proposed that a potential cue for this process is the title of the document. A label signalling the relations between documents might also serve this purpose.

Alternatively, the Active Processing Model (Hofman & van Oostendorp, 1999; Shapiro, 1998) suggests that structured overviews may inhibit the use of comprehension strategies by readers. Although readers generate causal inferences between text ideas as part of the normal course of comprehension (e.g. Graesser, Singer, & Trabasso, 1994), it may be simpler for students who read the overview to perceive the text structure without putting a tremendous amount of thought into discovering the relations between sections (Shapiro, 1998). In contrast, a text without an overview may require a deeper level of processing of the information in order to make sense of the material. In the context of multiple-documents comprehension, the Active Processing Model might stress prior results showing that undergraduate students are normally active integrating information across documents (Kurby et al., 2005). If students are provided with a graphical overview signalling the rhetorical relations between documents, however, they do not need to be active in order to discover

the relations between documents for themselves, which will result in poorer intertextual comprehension.

Currently, based on existing theoretical models for single-text comprehension, it is unclear whether signalling rhetorical relations between documents will enable readers to construct a more adequate Intertext model, inhibit strategic processing of interrelations between documents, or some combination of the two.

To the best of our knowledge, the only attempts to empirically evaluate the impact of signalling rhetorical relations on multiple-documents comprehension have been performed by Britt, Rouet and Perfetti (1996) and by Stadtler and Bromme (2008) (1). Britt et al. (1996) presented undergraduate students with varying experience in history with a hypertext which contained nine conflicting documents on the topic of the history of the Panama Canal. Students read the documents with the goal of 'understanding the controversy' concerning the topic. They were provided with a table of contents and a navigable hierarchical map. An important variable considered was the organization of the table and map. In the 'structured' condition, the table of contents presented the documents grouped under two main topics: 'Planning of the revolution' and 'Execution of the revolution'. The table also included the title of each text, the author's name, the year of publication and the page of the document in the system. The map presented the titles and authors of the documents organized in a manner that corresponded to the hierarchical structure of the information. At a first level it identified the two main topics discussed (i.e. "planning" and "execution" of the Revolution). At a second level in the hierarchy, it signalled rhetorical information regarding conflicting views on the two main topics, such as 'US military intervention not justified' vs. 'US military intervention justified'. Finally, at a third level it displayed primary documents used by authors of the above mentioned texts to back their arguments. In the 'scrambled' condition, the documents were organized in a random order that conveyed no information about how the documents related to each other in a hierarchical structure. Results showed that participants using the 'structured' version recalled more map items (titles and author names), correct page numbers, and arguments and evidence that the author gave to support their claims. These data fit well with the implications of the Assimilation Theory for multiple-documents comprehension: students could have used the structured

map as a schema in which to integrate different information about the relationships between documents.

In sum, the results of Britt et al. (1996) suggest the importance of signalling the existing relationships between texts for effective learning in a multiple-documents scenario. However, this conclusion should be taken cautiously due to two main seasons. First, the 'scrambled' condition used by Britt et al. (1996) may not represent a clear control group in which no rhetorical information is imposed onto the documents, because the 'scrambled' table of contents and hierarchical map may suggest an illogical organization of the documents that could interfere with students' normal comprehension processing. In this case, the students not only had to discover the relationships between documents, as suggested by the Active Processing Model, but also to disentangle the conflicting document relations signalled in the map from the real ones. A condition with no rhetorical information would therefore be a better test for the assumptions of the Active Processing Model. Second, the study by Britt et al. (1996) evaluated participants' memory for arguments from particular documents, but did not address their integration across documents, which is considered an essential component of multiple-documents learning (Perfetti et al., 1999).

Stadtler and Bromme (2008) took a different perspective to explore the impact of signalling rhetorical relations in multiple-documents comprehension. They requested undergraduate students to conduct an Internet search to gather information for a fictitious friend diagnosed with high level of cholesterol. Students had low background knowledge on that topic, and their task was to study the pros and cons of possible medical treatments. The experimental group conducted the study task using the web browser *met.a.ware*, which provided students with a note taking facility including a set of rhetorical classification. Students were instructed to copy and paste relevant information and classify that under the rhetorical categories provided, such as "function of cholesterol", "causes of a high level of cholesterol", or "consequences of a high level of cholesterol". A control group studied the web pages and took notes, but it did not have access to the rhetorical classification. Students from both groups scored similarly in a comprehension test. Nevertheless, participants from the experimental group wrote more structured topic essays, which included most of the rhetorical categories used, and searched for more relevant information, as reflected in more complete notes. Results by Stadtler and

Bromme (2008) add to the importance of using rhetorical relationships while studying on the Web, and support the Assimilation Theory by showing that low knowledge readers are not usually active in processing the rhetorical relations between web documents unless they are provided with a scaffold to do that. However, the extent to which these results apply to study scenarios with regular search engines is currently an open empirical question, which will be addressed in the current study.

In order to test the predictions of the Assimilation Theory and the Active Processing Model, respectively, regarding the role of rhetorical information on the integration of information across multiple documents, we performed an experiment in which students with varying prior knowledge about the documents' topic were provided with the results page for the query "Climate change" in two search engines: one providing information about the relationships between documents (a mocked version of www.kartoo.com), and another without such information (a mocked version of www.google.com). After reading the documents, students were required to complete a series of comprehension tasks, including a verification task aimed at assessing intertextual (i.e., cross-document) comprehension. As prior knowledge has been demonstrated to affect students' reading of multiple texts in several studies (Bråten & Strømsø, 2006a; Gil, Bråten, Vidal-Abarca & Strømsø, in press; Le Bigot & Rouet, 2007; Pieschl, Stahl, & Bromme, 2008; Rouet, Favart, Britt, & Perfetti, 1997; Strømsø, Bråten, & Samuelstuen, 2008), we decided to include a measure of participants' topic knowledge in order to control for this variable. In addition, we included a measure of participants' topic interest. Although there is, to date, at least to our knowledge, no empirical evidence that topic interest is related to multiple-document comprehension, it seems plausible that students' interest will play a role when they face the demanding task of comprehending and integrating content from multiple sources about a complex topic. It has been documented that well-developed personal interest in an area facilitates single-text comprehension (Hidi, 2001), and there is also some indication that personal interest may affect the quality of text-based learning by leading to more elaborate and deeper text processing (Krapp, 1999; Schiefele, 1999); such qualities are often needed when a reader tries to integrate information across different sources (Bråten & Strømsø, 2006b). As multiple-text comprehension will require students both to integrate information and to show persistence in reading several texts about the same topic, we decided to include topic interest as a control variable.

Experiment

Method

Participants. Seventy-four undergraduate education majors from the University of Valencia participated to fulfil class-credit (Mean age= 20.4, *SD*= 3.7; 88% female; mostly Caucasian). Six participants did not complete all experimental tasks and were therefore dropped from the analyses.

Apparatus. We recorded participants' on-line reading behaviour using an ad-hoc program written in Microsoft Visual Basic ® simulating a well-known web browser. The program recorded participants' clicks on hyperlinks or browser buttons, as well as reading times with reliable millisecond timing.

Materials. We adapted seven authentic texts from different sources on the topic of climate change to single-node web sites. Sources varied from well-known national newspapers (n=2), international universities (n=2), the Spanish Minister of Environment, a Spanish publishing company, and a Spanish-Argentinean gas company. Apart from a more neutral textbook excerpt, the six other pages contained partly conflicting information, with two pages presenting different views on the causes of global warming (manmade versus natural), two pages presenting different views on the consequences of global warming (negative versus positive), and two pages presenting different views on the solutions to global warming (international cooperation versus new technology). Page length ranged from 299 to 464 words (*M*= 376, *SD*=56). These materials have been successfully used in previous studies (e.g. Gil et al., in press).

We designed two types of results pages: a list-like and a graphical-overview interface. The list-like results page presented the seven pages as a list of results in the following order: First the introductory textbook excerpt; second and third, two pages dealing with causes of climate change; fourth and fifth, two pages on possible consequences of climate change; and sixth and seventh, two pages on possible solutions to reduce climate change. The results page included the title of each page, the publisher of the page, and the year of publication. The interface mimicked that of www.Google.com (Figure 1). The graphical-overview interface presented that same information as a constellation of page-links around the screen. Specifically, pages were grouped according to their rhetorical cause-effect relations

in four labelled chunks: (1) definition, (2) causes, (3) consequences and (4) solutions. This type of rhetorical relations is common to many learning situations involving scientific documents. The interface was inspired by that of www.Kartoo.com (Figure 2).

--- Insert figures 1 & 2 here ---

Prior knowledge measure. We used a multiple-choice test consisting of 17 items with five answer alternatives. Diverse aspects of the topic were assessed, with items referring to both scientific (e.g., the greenhouse effect) and political (e.g., the Kyoto Protocol) issues.

Topic interest measure. We used a 12-item test to measure participants' personal interest and engagement in issues and activities concerning climate change. Participants indicated their level of interest or engagement by rating each item on a 10-point Likert-type scale (from 1 = 'not at all true of me' to 10 = 'very true of me'). Half of the items allowed participants to express their interest in the topic without reporting any active engagement or involvement in addressing the problem of climate change (sample items: 'I am interested in what conditions influence the Earth's climate'; 'Global warming is an issue that interests me'). The other items focused more on participants' active engagement and involvement in the issue, thus reflecting their willingness to act for the benefit of the Earth's climate (sample items: 'I am concerned with how I myself can contribute to the reduction of environmental pollution'; 'I try to convince others that we must reduce the discharges of climate gases').

Text understanding measures. Readers' memory for text content was measured with a sentence verification task (SVT) following Royer, Carlo, Defresne, and Mestre's (1996) procedure. This task included 34 sentences (10 originals, 9 paraphrases, 7 meaning changes, and 8 distractors), and participants were instructed to mark a sentence yes if it had the same meaning as a sentence in one of the texts (i.e., an original or a paraphrase) and no if it had a different meaning (i.e., a meaning change or a distractor). Participants' score on this measure was the number of correct responses out of the 34 items.

To assess participants' deeper understanding of each single text we used an intratextual inference verification task (IntraVT), again following Royer el al. (1996). The IntraVT consisted of 29 items that were constructed by combining information from different sentences within one of the texts to form either a valid or an invalid inference. There were 16 valid and 13 invalid inferences, and participants were instructed to mark those sentences *yes* that could be inferred from material presented in one of the texts and those sentences *no* that could not be inferred from material presented in one of the texts. Participants' score on this task was the number of correct responses out of the 29 items.

Finally, we used an intertextual inference verification task (InterVT) to measure participants' ability to draw inferences across texts. The task consisted of 26 statements, 14 of which could be inferred by combining information from at least two of the texts (i.e., valid inferences), and 12 of which could not be inferred by combining information from at least two of the texts (i.e., invalid inferences). Most valid inferences linked information from texts of different categories, as defined in the Kartoo interface. For example, in order to correctly verify the valid inference 'Countries that have signed the Kyoto Protocol do not agree with theories that consider astronomic conditions the unique cause of climate change', students have to link the idea that the Kyoto Protocol aims at reducing manmade discharges of climate changes (text from the Solutions category) with the view that astronomic conditions natural cause climate change (text from the Causes category). The participants were instructed to mark the valid inferences *yes* and the invalid inferences *no*. Participants' score on the InterVT was the number of correct responses out of the 26 items.

Navigation path measures. Cohesion of the navigation path was analyzed with the mathematic technique of Latent Semantic Analysis (LSA), using a Spanish general corpus. The text of all the pages was analyzed with the matrix analysis contrast (document to document comparison) that compares the content of each page with the content of every other page. LSA cosines provided a measure of the degree of argument overlap between texts, which is assumed to reflect the level of cohesion between them (Foltz, Kintsch & Landauer, 1998). The rationale for this approach is that when two propositions are actually related semantically, there frequently exists a shared argument between them (Kintsch, 1992). For each participant, we computed LSA cosines between transited pages. We used the mean LSA cosine of all the transitions as a measure of the cohesion of the

participants' navigation path. For example, the page titled 'Could climate change be due to natural causes?' (from the category 'causes') had a LSA cosine of .45 with the page 'Manmade greenhouse effect ' (category 'causes'), and a cosine of .30 with the page 'New technology reducing discharges of climate gases' (category 'solutions'). Thus, a participant will have a more cohesive navigation path if going from the first-mentioned page to the second one than if going from the first to the third one. Prior research on single web document comprehension has shown that a higher cohesion between transited nodes of a hypertext boosts comprehension (Madrid, van Oostendorp, & Puerta Melguizo, 2009; Salmerón, Cañas, Kintsch & Fajardo, 2005).

In addition, we measured the homogeneity of participants' navigation path by means of the Levenshtein algorithm, a string-edit method that produces an index of the similarity between any two sequences (e.g. navigation paths). This index, called the Levenshtein distance, is computed as the smallest possible cost of elementary operations of insertion, substitution and deletion of pages required to align or transform one navigation path into another (Sankoff & Kruskal, 1983). Similar navigation paths have smaller Levenshtein distances. We computed the Levenshtein distance for all participants' navigation path in each condition separately (i.e. list and graphical-overview interfaces). Mean distances for each participant were used to analyze the homogeneity of navigation paths in the two types of interfaces.

Procedure. In one session, participants were tested with respect to their prior knowledge and interest concerning the topic. In a subsequent session, participants read the seven web pages with instructions to read them carefully as if they were preparing a test. They were also told that the information in the search results page might be helpful for comprehending the texts. However, no instructions were given on how to examine the overview. Participants were first presented with the search results, and afterwards they could click on any link leading to any of the seven web pages. No links were included in the pages themselves, except for a "going back" link at the bottom of the page that lead participants back to the search results page. When participants decided that they were ready to answer the questions, they were administered the text understanding measures in the following order: they all first answered the SVT; then, half of them completed the IntraVT and after the InterVT, whereas the other half completed first the InterVT and after the IntraVT.

Results

To test the hypotheses we performed separate ANCOVAs for each text understanding measure (SVT, IntraVT and InterVT), reading times and navigation paths followed by students. Prior knowledge and topic interest were included as covariates, and type of results page (list-like or graphical-overview) as independent variable (data are summarized in Table 1).

Comprehension outcomes

Participants' memory for text content, as measured by the SVT task, did not vary based on type of interface, $F(1, 64)= 0.01$, $p > .95$. The covariates did not explain either participants' differences on this variable, with $F(1, 64) = 0.02$, $p < .95$ and $F(1, 64)= 0.77$, $p < .40$, for prior knowledge and topic interest, respectively.

Similarly, inferential understanding of each single text, as measured by the IntraVT task, did not differ for the two interfaces, $F(1, 64)= 0.14$, $p < .75$. Regarding the covariates, topic interest did not influence students' performance, $F(1, 64) = 0.05$, $p < .85$, but prior knowledge did, $F(1, 64)= 6.35$, $p < .05$, partial $\eta^2 = .09$. As prior knowledge increased, so did scores on the IntraVT task.

Finally, participant's ability to draw inferences across documents varied by type of interface, $F(1,64)= 3.93$, $p < .05$, partial $\eta^2 = .05$. After controlling for the effect or prior knowledge and topic interest, participants using the graphical-overview interface scored higher than those using the list interface (see means in Table 1). With this dependent variable, prior knowledge did not influence students' performance, $F(1, 64) = 1.22$, $p < .30$, but topic interest did, $F(1, 64)= 4.30$, $p < .05$, partial $\eta^2 = .06$. As topic interest increased, so did scores on the InterVT task.

Reading times

We measured participants' reading times both while they were reading the results interface and while they were reading the pages. For the interface reading times, we first analyzed reading times for every visit to the interface. Total reading times were weighted on the basis of the number of words in the interface. For the pages reading times, we distinguished between reading times for the

first visit to the page, and times for the subsequent visits to a page (i.e. rereading). Reading times were weighted on the basis of the number of words in each page. Finally, reading times for each page were averaged to get a single text mean reading time value.

Reading times of the results interface varied between condition, $F(1, 64)= 5.33$, $p < .05$, partial $\eta^2 =.08$. Participants in the graphical-overview interface spent a longer time reading the results page than those in the list interface condition. Covariates did not explain differences on this variable: $F(1, 64)= 0.05$, $p < 0.95$ and $F(1, 64)= 0.61$, $p < .45$, for prior knowledge and topic interest, respectively. In addition, students' reading times for the pages during the initial visit differed between interfaces, $F(1, 64)= 3.90$, $p = .05$, partial $\eta^2 =.06$. Participants using the list interface spent more time reading the texts than those using the graphical-overview interface. Again, neither prior knowledge, $F(1, 64)= 0.02$, $p < 0.90$, nor topic interest, $F(1, 64)= 0.08$, $p < 0.08$, was related to reading times (2). Finally, texts' reading times for rereadings did not differ between interfaces, $F(1, 64) = 0.11$, $p < .75$. Prior knowledge did not explain differences between participants, $F(1, 64) = .94$, $p < .35$, but topic interest did, $F(1, 64) = 6.92$, $p < 0.05$, partial $\eta^2 = .10$. The higher the participants' topic interest in climate change, the longer they spent rereading documents.

Navigation path

Participants followed a similarly cohesive navigation path, as measured by the mean LSA cosines of participants' transitions, $F(1, 64)= 2.62$, $p < .15$. Covariates did also not affect the cohesion of the path, $F(1, 64)= 0.51$, $p < .50$ and $F(1, 64)= 0.08$, $p < .80$, for prior knowledge and topic interest, respectively. Nevertheless, participants differed by interface on the homogeneity of their navigation path, as measured by the string distance provided by the Levenshtein algorithm, $F(1, 64) = 159.13$, $p < .01$, partial $\eta^2 = .71$. Navigation paths of participants in the list interface condition were more homogeneous than were those in the graphical-overview interface. Covariates did not have an effect on this variable, $F(1, 64)= 0.14$, $p < .75$ and $F(1, 64)= 1.70$, $p < .20$, for prior knowledge and topic interest, respectively (3).

Discussion

Results clearly indicate that signalling the rhetorical relations between documents, as done by the search engine www.kartoo.com, benefits students' learning in terms of comprehension, reading times' efficiency and navigation. Results converge in the idea that students may have relied on the structural information provided by the graphical-overview interface to establish links between ideas from the different documents, as suggested by the Assimilation Theory (Mayer, 1979).

Regarding comprehension outcomes, results indicate that signalling documents' rhetorical relations improves students' integration of ideas across documents, compared to a mere list of the documents without explicit signalling. Interestingly, this effect does not arise for surface and deep comprehension of single documents. This is congruent with the view that signalling relationships between documents only tapped comprehension processes related to construction of the Intertext model, which is supplementary to the comprehension of each individual document (Britt et al., 1996). Students could have used the overview information to form an initial frame of the existing relationships between documents, as proposed by the Assimilation Theory (Mayer, 1979). In turn, students might have relied on this initial frame while reading the documents to evaluate the extent to which their knowledge of previous documents relates to subsequent texts (Kurby et al., 2005). The facilitative effect of the interfaces signalling documents' rhetorical relations is not mediated by readers' background knowledge level, which do not concurs with prior results on single document comprehension (Lorch & Lorch, 1996; Salmerón et al., 2009). Thus, whereas high prior knowledge may serve as a schema to guide comprehension of single documents, it does not seem to warrant an optimal processing of the relationships across web documents.

The Active Processing Model, in contrast, cannot easily account for the facilitative effect of signalling the relationships between documents. From the viewpoint of this model, under those circumstances students are less active in evaluating relationships between information, which should lead to poorer intertextual understanding. Similarly, data do not support the proposal by Kurby et al. (2005) suggesting that titles may be a potential cue for a "top-down" process of document integration. Students' evaluation of the extent to which activated information from a recently read document is related to the current document seems to be better supported by a combination of the document title and a label signalling its rhetorical relationship with other pages.

Students using the graphical overview read the interface for a longer time but read the text for a shorter time than those using the list interface. This pattern of results suggests that participants in the graphical overview condition process intertextual relations between documents after reading each web page, with the visual support of the interface. In the list overview condition, students may just disregard the interface for the purpose of elaborating the relations between documents, and focus instead on their integration while reading the documents, which may be a harder task. Nevertheless, this interpretation of reading time data is rather tentative, and should be supported in future research with methodologies able to capture readers' on-line processes, such as the think-aloud methodology.

Finally, data from navigation patterns reveal that the graphical-overview interface supported a more free navigation across documents (cf. Stadtler & Bromme, 2008). Students learning with this interface may thus be more eager to explore the documents in their own fashion, according to their learning needs or interests. Interestingly, this free navigation pattern is as semantically cohesive as the mean pattern in the list interface condition. Thus, an open navigation in the graphical interface condition does not come with the cost of a less cohesive - and thus more difficult to learn - navigation pattern (Madrid et al., 2009; Salmerón et al., 2005). Participants in the list interface condition, in contrast, are quite homogeneous in their navigation, sticking to a linear reading of the documents from the top of the list to the bottom. A list interface heavily imposes a strict order in which to read the documents, which may inhibit students' free exploration of the documents (Pan, Hembrooke, Joachims, Lorigo, Gay & Granka, 2007).

Limitations and Future Research

Even though we believe that this experiment contributes to the literature on multiple-documents comprehension by clarifying the beneficial effect of signalling rhetorical relationships between pages, it should be acknowledged that our investigation comes with certain limitations. First of all, the type of rhetorical relations signalled in this experiment corresponds to a case of cause, effect and solution relationships, which are common in learning from scientific documents. Nevertheless, further research should evaluate the effectiveness of signalling other types of rhetorical relations between documents, such as comparison and contrast, providing evidence, antithesis … (Mann, & Thompson, 1988). The

Rhetorical Structure Theory (RST) may be considered as a starting framework to test possible web document relationships (Mann, & Thompson, 1988).

In addition, future studies may explore the impact of the degree of explicitness of the signalling. Providing too much information about the relations between documents could indeed be detrimental for the comprehension of less skilled readers, as suggested by recent studies of single-document comprehension (Scott & Schwartz, 2007).

Furthermore, future research may study how signalling rhetorical relationships between web pages interacts with other types of document signals, such as source of information. Recent studies suggest that signalling the type of source (e.g. commercial, university, …) in list-like search interfaces may influence readers' navigation and performance (Kammerer, Wollny, Gerjets & Scheiter, 2009). The multiple-documents model (Perfetti et al., 1999) suggests that providing both rhetorical and source web page information might have a stronger impact on students' construction of a complete intertextual model.

Finally, we should note that the graphical overview interface used included not only an explicit signalling of the relationships between web documents, but also a particular arrangement of the items that facilitated a visual grouping of the related pages. Although that was done in order to maximize the ecological validity of the interface, which resembled the one used by www.kartoo.com, it does not allow considering the isolated effects of signalling and grouping. In order to disentangle the possible effects of both design factors, signalling and grouping, future studies should explore the possibility of signalling rhetorical relationships in a list-like interface.

Implications for Instruction and Design

From a practical point of view, the results of the experiment suggest that working with web pages' rhetorical signals may be beneficial for learners. Currently the automatic tagging of these relationships is limited, to the best of our knowledge, to the meta search engine Kartoo (www.kartoo.com), and to a lesser extent, to Grokker (www.groxis.com) (Turetken & Sharda, 2007). Nevertheless, the quality of the automatic tagging and grouping varies depending on the language and

topic. Thus, the engine could probably not be useful in its current version to support a big range of learning situations.

Also, other instructional approaches can be taken in order to take advantage of working with web pages' rhetorical relationships, with the goal of engaging students in an active processing of intertextual information. For example, students can be instructed to actively tag some rhetorical relations between pages provided by the instructor (or selected by them), as supported by the web browser *met.a.ware* (Stadtler & Bromme, 2008). This activity resembles the manual "tagging" of documents (e.g. photographs, videos, blog entries…) which has become a common activity for many users of the Web 2.0 (Alexander, 2006). Another possibility consists of instructing students to create their own "rhetorical graphical overviews", by tagging the relationships between visited pages during a particular learning session. Currently, some web browsers, such as *Nestor* (Zeiliger, & Esnault, 2008), automatically construct navigation maps from the students' learning sessions, and allow them to tag both each page and the links between the pages visited. Further research will be required to empirically validate the effect of these instructional activities on students' learning.

References

Alexander, B. (2006). Web 2.0: A new wave of innovation for teaching and learning? *EDUCAUSE Review, 41*, 32–44.

Bhavnani, S. K., Jacob, R. T., Nardine, J., & Peck, F. A. (2003). Exploring the distribution of online healthcare information. *CHI '03 extended abstracts*, 816-817.

Bråten, I., & Strømsø, H.I. (2006a). Effects of personal epistemology on the understanding of multiple texts. *Reading Psychology, 27*, 457–484.

Bråten, I., & Strømsø, H.I. (2006b). Constructing meaning from multiple information sources as a function of personal epistemology: The role of text-processing strategies. *Information Design Journal, 14*, 56–67.

Britt, M.A., & Gabrys, G. (2001). Teaching advanced literacy skills for the World Wide Web. In Wolfe, Christopher R. (Ed). *Learning and teaching on the World Wide Web* (pp. 73-90). San Diego, CA, US: Academic Press.

Britt, M. A., & Gabrys, G. (2002). Implications of document-level literacy skills for Web-site design. *Behavior Research Methods, Instruments, & Computers, 34,* 170-176.

Britt, M.A., Rouet, J.-F., & Perfetti, C.A. (1996). Using hypertext to study and reason about historical evidence. In J.-F. Rouet, J. J. Levonen, A. Dillon & R. J. Spiro (Eds.), *Hypertext and cognition* (pp. 43-72). Mahwah, NJ: Lawrence Erlbaum Associates.

Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). The measurement of textual coherence with Latent Semantic Analysis. *Discourse Processes, 25,* 285-307.

Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H.I. (in press). Understanding and integrating multiple science texts: Summary tasks are sometimes better than argument tasks. *Reading Psychology*.

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

Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational Psychology Review, 13,* 191–209.

Hofman, R., & van Oostendorp, H. (1999). Cognitive effects of a structural overview in a hypertext. *British Journal of Educational Technology, 30,* 129-140.

Kammerer, Y., Wollny, E., Gerjets, P., & Scheiter, K. (2009). How authority-related epistemological beliefs and salience of source information influence the evaluation of web search results – an eye tracking study. *Proceedings of the Thirty-first Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Kintsch, W. (1992). How readers construct situation models for stories: The role of syntactic cues and causal inferences. In A. F. Healy, S. N. Kosslyn & R. M. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (pp. 216-278). Hillsdale, N.J.: Lawrence Erlbaum Associates.

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

Krapp, A. (1999). Interest, motivation, and learning: An educational-psychological perspective. *European Journal of Psychology of Education, 14,* 23–40.

Kurby, C. A., Britt, M. A. and Magliano, J. P. (2005). The role of top-down and bottom-up processes in between-text integration. *Reading Psychology, 26*, 335-362.

Le Bigot, L., & Rouet, J.F. (2007). The impact of presentation format, task assignment, and prior knowledge on students' comprehension of multiple online documents. *Journal of Literacy Research, 39*, 445–470.

Lorch, R. F., & Lorch, P.E., (1996). Effects of organizational signals on free recall of expository text. *Journal of Educational Psychology, 88*, 38–48.

Madrid, R. I., van Oostendorp, H., & Puerta Melguizo, M. C. (2009). The effects of the number of links and navigation support on cognitive load and learning with hypertext: The mediating role of reading order. *Computers and Human Behavior, 25*, 66-75.

Mann, W.C. & Thompson, S. (1988). Rhetorical structure theory: toward a functional theory of text organization. *Text, 8*, 243-281.

Mayer, R. E. (1979). Twenty years of research on advance organizers: Assimilation theory is still the best predictor of results. *Instructional Science, 8*, 133-167.

Olive, T., Rouet, J.-F., François, E., Zampa, V. (2008). Summarizing digital documents : effects of alternate or simultaneous window display. *Applied Cognitive Psychology, 22*, 541-558.

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust: users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication, 12*, article 3. http://jcmc.indiana.edu/vol12/issue3/pan.html

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, N.J.: Lawrence Erlbaum.

Pieschl, S., Stahl, E., & Bromme, R. (2008). Epistemological beliefs and self-regulated learning with hypertext. *Metacognition and Learning, 3*, 17–37.

Royer, J.M., Carlo, M.S., Dufresne, R., & Mestre, J. (1996). The assessment of levels of domain expertise while reading. *Cognition and Instruction, 14*, 373–408.

Rouet, J.-F., Favart, M., Britt, M.A., & Perfetti, C.A. (1997). Studying and using multiple document in history: Effects of discipline expertise. *Cognition and Instruction, 15*, 85–106.

Ruddell, R. B., & Boyle, O. F. (1989). A study of cognitive mapping as a means to improve summarization and comprehension of expository text. *Reading Research and Instruction, 29*, 12-22.

Salmerón, L., Baccino, T., Cañas, J. J., Madrid, R. I., & Fajardo, I. (2009). Do graphical overviews facilitate or hinder comprehension in hypertext? *Computers & Education. 53,* 1308-1319.

Salmerón, L., Cañas, J. J., Kintsch, W. & Fajardo, I. (2005). Reading strategies and hypertext comprehension. *Discourse Processes, 40,* 171-191.

Sankoff, D. & Kruskal, J. B. (Eds.) (1983). *Time warps, string edits, and macromolecules: The theory and practice of sequence comparison*. Reading, MA: Addison-Wesley.

Schiefele, U. (1999). Interest and learning from text. *Scientific Studies of Reading, 3,* 257–279.

Scott, B. M., & Schwartz, N. H. (2007). Navigational spatial displays: The role of metacognition as cognitive load. *Learning and Instruction, 17,* 89-105.

Shapiro, A. M. (1998). Promoting active learning: The role of system structure in learning from hypertext. *Human-Computer Interaction, 13*, 1-35.

Stadtler, M., & Bromme, R. (2008). Effects of the metacognitive computer-tool met.a.ware on the web search of laypersons. *Computers in Human Behavior, 24*, 716-737.

Strømsø, H.I., Bråten, I., & Samuelstuen, M.S. (2008). Dimensions of topic-specific epistemological beliefs as predictors of multiple text understanding. *Learning and Instruction, 18*, 513–527.

Turetken, O., & Sharda, R. (2007). Visualization of web spaces: state of the art and future directions. *The DATA BASE for Advances in Information Systems, 38*, 51-81.

Wiley, J. (2001). Supporting understanding through task and browser design. *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Zeiliger, R., & Esnault, L. (2008). The constructivist mapping of internet information at work with Nestor, in A. Okada, S.B. Shum, & T. Sherborne (Eds.). *Knowledge Cartography* (pp. 89-111). London: Springer.

Notes

1. Prior studies have extensively explored the comprehension effects of signalling the structure of hypertext nodes by means of graphical overviews (for a revision see Salmerón et al., 2005). The typical interface used in these studies depicts the structure of a single hypertext by means of arrows linking the node tittles. In a hierarchical overview a node linked to an upper section likely corresponds to a subsection of the second. In a network-like overview, by contrast, a link may suggest just a potential relation between two hypertext sections, but it does not specify the type of the existing relationship. Our study deals with interfaces signalling explicitly the relationship between pages, e.g. using the tag “consequences” linking two documents about “climate change” denotes that both documents deal with the consequences of the phenomenon.

2. In light of the significant differences in reading times between interfaces, we reanalyzed the ANOVAs with the comprehension outcomes as dependent variables, including reading times for both the results interface and for the pages as covariates. The reported ANOVAs did not varied significantly after the inclusion of the new covariates.

3. In light of the significant differences in the similarity of navigation paths between interfaces, we reanalyzed the ANOVAs with the comprehension outcomes as dependent variables, including the Levenshtein distance of students’ navigation paths as covariates. The reported ANOVAs did not varied significantly after the inclusion of the new covariate.

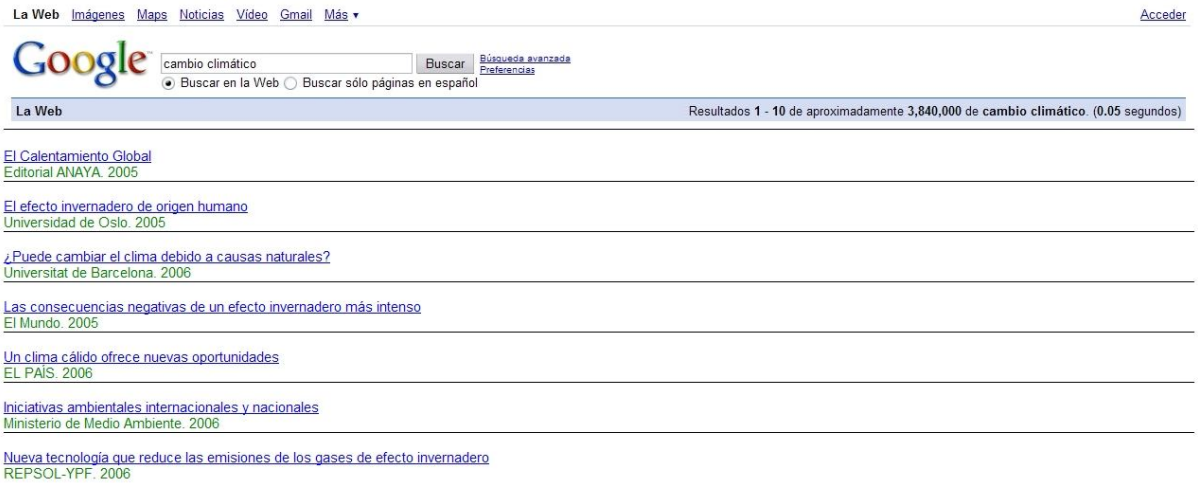
Author notes

This research was funded by the Spanish Minister of Science and Technology (Projects SEJ2005-04500/EDUC and SEJ2008-03072/EDUC). We thank Eduardo Vidal-Abarca for his helpful comments on a previous draft of this paper.

Figure Caption

Figure 1. Screen capture from the list-like interface used in the experiment. The interface mimicked www.google.com, and provided a list with the tittles of seven web pages.

Figure 2. Screen capture from the graphical-overview interface used in the experiment. The interface mimicked www.kartoo.com, and signalled the rhetorical relations across seven web pages (i.e. effects, causes, and consequences).


La Web Imágenes Maps Noticias Vídeo Gmail Más ▾
Acceder
Google
cambio climático
Buscar
Búsqueda avanzada
Preferencias
Buscar en la Web
Buscar sólo páginas en español
La Web
Resultados 1 - 10 de aproximadamente 3,840,000 de cambio climático. (0.05 segundos)
El Calentamiento Global
Editorial ANAYA. 2005
El efecto invernadero de origen humano
Universidad de Oslo. 2005
¿Puede cambiar el clima debido a causas naturales?
Universitat de Barcelona. 2006
Las consecuencias negativas de un efecto invernadero más intenso
El Mundo. 2005
Un clima cálido ofrece nuevas oportunidades
EL PAIS. 2006
Iniciativas ambientales internacionales y nacionales
Ministerio de Medio Ambiente. 2006
Nueva tecnología que reduce las emisiones de los gases de efecto invernadero
REPSOL-YPF. 2006

Definición

El Calentamiento Global
Editorial ANAYA. 2005

Causas

El efecto invernadero de origen humano
Universidad de Oslo. 2005

¿Puede cambiar el clima debido a causas naturales?
Universitat de Barcelona. 2006

Consecuencias

Las consecuencias negativas de un efecto invernadero más intenso
El Mundo. 2005

Un clima cálido ofrece nuevas oportunidades
EL PAÍS. 2006

Soluciones

Iniciativas ambientales internacionales y nacionales
Ministerio de Medio Ambiente. 2006

Nueva tecnología que reduce las emisiones de los gases de efecto invernadero
REPSOL-YPF, 2006

Table 1

Summary of ANCOVAS for the effect of Type of Interface (list-like and graphical-overview), with Prior Knowledge and Topic Interest as covariates.

	List interface (No signaling)	Graphical-overview interface (Explicit signaling)
SVT	0.67 (0.15)	0.67 (0.16)
IntraVT	0.60 (0.15)	0.62 (0.16)
InterVT	0.67 (0.15)	0.74 (0.15)
Results interface reading time	176.58 (225.36)	303.79 (232.19)
Text reading time (initial visit)	420.36 (63.76)	390.67 (65.69)
Text reading time (revisits)	42.79 (55.21)	47.31 (56.89)
Cohesion of navigation paths (LSA cosine)	0.51 (0.02)	0.50 (0.02)
Levensthein distance of navigation paths	2.18 (1.54)	6.92 (1.58)

Note: Adjusted means are presented together with its standard deviation (in brackets).