Scanning and Deep Processing of Information in Hypertext: An Eye-Tracking and Cued Retrospective Think-Aloud Study

Ladislao Salmerón [1], Johannes Naumann [2], Victoria García [1] & Inmaculada Fajardo [1]

[1] University of Valencia, Spain

[2] Goethe University Frankfurt am Main, Germany

Manuscript in press (25/01/2015). Please check the original reference:

Salmerón, L., Naumann, J., García, V., & Fajardo, I. (2015). Scanning and deep processing of information in hypertext: An eye-tracking and cued retrospective think-aloud study. *Journal of Computer Assisted Learning*.

Abstract

When students solve problems on the Internet, they have to find a balance between quickly scanning large sections of information in web pages and deeply processing those that are relevant for the task. We studied how high school students articulate scanning and deeper processing of information while answering questions using a Wikipedia® document and how their reading comprehension skills and the question type interact with these processes. By analyzing retrospective think-aloud protocols and eye-tracking measures, we found that scanning of information led to poor hypertext comprehension, while deep processing of information produced better performance, especially in location questions. This relationship between scanning, deep processing, and performance was qualified by reading comprehension skills in an unexpected way: scanning led to lower performance, especially for good comprehenders, while the positive effect of deep processing was independent of reading comprehension skills. We discussed the results in light of our current knowledge of Internet problem solving.

Keywords: Internet problem solving, reading comprehension, cued retrospective think-aloud, eye tracking.

Scanning and Deep Processing of Information in Hypertext: An Eye-Tracking and Cued Retrospective Think-Aloud Study

A critical point of information problem solving on the Internet, as in other task-oriented reading activities, is that students are confronted with vast amounts of information in which only a fraction is relevant for their learning goal. Imagine a student reading a Wikipedia® document suggested by her teacher to answer some questions. She has to decide which sections of the page and which links to other articles are relevant for her task (scanning information). Due to the interconnected nature of the Internet, reading the whole collection of linked articles is not a realistic option. After identifying a relevant section or link, she will have to carefully read the article to get an accurate representation of the information (deep processing of information). Determining an optimal trade-off between both processes – scanning and deep processing – during problem solving is essential to warrant a successful learning outcome. If students rely too heavily on scanning, they might cover a vast amount of information at the expense of constructing a rather poor representation of the information. On the contrary, if students rely too much on processing the hypertext deeply, they might end up building an accurate but incomplete representation of only the information they could handle during the time assigned to the task.

In this paper, we studied how high school students articulate scanning and deeper processing of information while answering questions in a Wikipedia® document and how their comprehension skills and the type of question interact with these processes. Understanding the role of those factors on scanning and processing is not only relevant from a theoretical perspective (e.g., how students 'transfer' highly trained printed comprehension skills to electronic environments) but also from a practical one (e.g., which skills and tasks have to be prioritized in reading literacy classes).

Information Problem Solving on the Internet

An influential framework to describe the processes that students undergo when using the Internet for academic activities is the Information Problem Solving on the Internet model (IPS-I) by Brand-Gruwel, Wopereis, and Walraven (2009). The IPS-I-model identifies five main processes (goal definition, search pages, scanning information, deep processing of information, and presentation of information) that are governed by different regulation activities and that are qualified by different conditional skills. Our study focused on the steps of scanning and deep processing of information. Specifically, once students access a web page, they must scan the sections to grasp the gist of the page sections, with the ultimate goal of deciding whether a section contains relevant information for their goal. If so, they might decide to process this particular section in more detail to elaborate a complete representation and to integrate it with their existing knowledge. Whenever students detect that the plan is not being efficient, they might revise it as well as the tactics used to accomplish it (cf. Winne, 2001). The search process relies heavily on conditional skills such as reading comprehension, which is necessary to assist students in linking information from an interconnected set of web pages.

Reading Comprehension Skills, Scanning, and Deep Processing

How do the reading comprehension skills of young students shape the way they scan and process texts with the purpose of answering questions? Previous studies have focused on how students read printed texts and subsequently search them to respond to specific questions. Evidence suggest that after an initial reading of a text, good comprehenders quickly scan it and only process deeply (e.g., spend more time rereading) those sections with information relevant to the question (for a review of empirical evidence related to comprehension skills and text relevance, see Authors, 2011a). Along this line, studies using Internet problem solving tasks indicate that undergraduate students are able to efficiently scan information from web pages, as inferred from their eye movements, when they have to quickly process a hypertext (Duggan & Payne, 2009). Studies using think-aloud methods clarify that proficient undergraduate students change from scanning to deep processing whenever they encounter relevant information to answer the questions posed (Zhang & Duke, 2008), especially if they are good learners (Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012). Deeply processing hypertext pages relevant for students' tasks is essential because it is one of the major predictors of hypertext comprehension by undergraduate students in the literature (Authors, 2008a).

Few studies have assessed how middle and high school students scan and process information on hypertext with the goal of answering questions. Those studies have relied exclusively on think-aloud methods. Coiro and Dobler (2007) studied a group of sixth graders that were highly skilled readers, as informed by their teachers and by standardized test scores. Results indicated that those proficient readers scan several sections from web pages and then deeply process a section when they find information relevant to their question. Cromley and Azevedo (2009) described in detail the search behavior of two middle school students that obtained low scores in a series of search tasks in a hypermedia environment. These participants didn't show the above pattern of scanning and deep processing, but rather they over-relied on scanning without noticing the relevant information they encountered. To the best of our knowledge, no previous study has explored the interrelations between reading comprehension skills, scanning, and deep processing in relation to hypertext comprehension in high school students.

In sum, previous research suggests that good comprehenders in the print medium are more efficient at changing from scanning to deep processing whenever they detect a hypertext section relevant to their goal (Coiro & Dobler, 2007; Zhang & Duke, 2008), and as a consequence, their comprehension of the hypertext improves (Authors, 2008b). Nevertheless, those results should be taken cautiously because previous studies have mostly focused on undergraduate students or on adolescents with high reading comprehension skills. In our study, we assessed high school students with a broader range of reading comprehension levels, which allowed us to extend prior evidence to less proficient comprehenders.

Question Type, Scanning, and Deep Processing

A different aspect of Internet problem solving that may be related to students' strategic behavior is question type. According to well-established frameworks of electronic reading literacy, such as OECD's PISA, two of the main tasks in the digital domain require students to either locate or integrate information (OECD, 2011, Chapter 1). Whereas to solve location questions students have to find information displayed in a particular hypertext section, to solve integrate questions they not only have to locate relevant pieces of information, usually from different sections, but also to relate them in a coherent way. Previous studies with printed texts (Cerdán & Vidal-Abarca, 2008; Cerdán, Vidal-Abarca, Martínez, Gilabert, & Gil, 2009; Rouet, Vidal-Abarca, Bert-Erboul, & Millogo, 2001) have shown that undergraduate students solved location questions by quickly scanning several text sections until they found one that may contain a response. To solve integrate questions, by contrast, students repeated cycles of scanning and deep processing of relevant sections. In addition, to solve integrate questions students not only need to process relevant sections but also must relate the information from those sections – for example, by means of inferences. In one of the few studies exploring question type and hypertext navigation, Rouet (2003) found that in location questions students inspected a low percentage of irrelevant pages, as indicated by log files, as compared to integrate questions. Those studies used log files to analyze students' inspection patterns, which only provide global measures of navigation between pages, but they didn't provide information about the inspection within a web page, just the time the participant spent reading it. Actually, the same reading times from the log files of two participants may mean very different things – for example, a participant could have just scanned the whole web page, while another could have read in detail one paragraph of the web page. By analyzing students' eye movements in our study, we aimed to shed more light on students' inspection patterns of web pages.

To the best of our knowledge, no previous study has explored how the different text inspection patterns relate to students' performance. We may expect that to answer a location question correctly students just need to identify and process the particular hypertext section containing relevant information. In integrate question, however, just locating and processing relevant pieces of text may not be enough to provide a correct response because students must also infer the relationships between those pieces of information (Cerdán & Vidal-Abarca, 2008). Students' reading comprehension skills may be necessary to perform this inferential processing. To test these hypotheses, we ran an eye-tracking study combined with a cued retrospective think-aloud protocol (Van Gog, Paas, Van Merriënboer, & Witte, 2005). The method allowed us to capture several indicators of students' processing of the hypertext while answering questions. With this methodology, we aimed to overcome a major concern of previous works that used different forms of concurrent think-aloud protocols (Coiro & Dobler, 2007; Cromley & Azevedo, 2009; Zhang & Duke, 2008), which may limit students' ability to report their thoughts under highly demanding situations, especially in the case of less skilled students (Van Gog, Paas, Van

Merriënboer, & Witte, 2005). Nevertheless, retrospective think-aloud protocols may involve a certain amount of interpretation on the part of the participants (Ericsson & Simon, 1984). For this reason, we also analyzed a more objective on-line measure, such as students' eye movements while navigating hypertext.

Hypotheses

Based on the previous findings reported above, we outlined a series of hypotheses that aimed at extending prior research regarding the relationship between adolescents' reading comprehension skills, task type, scanning and deep processing, and hypertext comprehension.

Hypothesis 1 predicts a negative relation between students' reading comprehension skills and the processing of irrelevant hypertext sections and a positive relation between skills and the processing of relevant sections.

Hypothesis 2 predicts that with location questions, students scan fewer number of sections than with integrate questions.

Hypothesis 3 states that with location questions, deep processing of relevant sections relates to higher hypertext comprehension regardless of students' reading comprehension skills, whereas with integrate questions, deep processing of relevant sections relates to higher hypertext comprehension only for students with high reading comprehension skills.

Method

Participants

Twenty-seven ninth and tenth grade students (*M*= 15.44 years old, *SD*: 1.02) from a high school in Valencia participated in the study. Only students with complete data (*N*= 21) were included in the analyses. Of the sample, 74% were female students, and they were native Spanish speakers with normal or corrected to normal visual acuity. On average, the students had been using computers for 5.6 years (*SD*= 1.8). On a five-point scale from “Never” to “Almost every day,” they reported that they read articles from Wikipedia® “Once or twice a month.” Finally, students had little background knowledge about the hypertext topic. Their scores in a questionnaire of 10 questions assessing their prior knowledge of the French Revolution (α = .66) were not different from chance level: $t(20) = 1.41$, $p = .17$. Their reading comprehension skills, as measured by a standardized test (see below), corresponded to their age-equivalent level (*Mdn* = 50, *IQR* = 20-80). Parents were informed about the goals of the study, and they signed a consent form prior to the students’ participation.

Materials

Hypertext system. Students accessed a document about the French Revolution adapted from a hypermedia distributed by the Spanish Ministry of Education through its on-line educational repository (Tapia, 2004). The presentation of the material was slightly modified, while maintaining the content of the original material. Specifically, we included a navigable table of contents at the beginning of the main hypertext to mimic the interface of Wikipedia® (Figure 1). In addition, we increased the distance between text lines to improve the accuracy of the eye-tracking system.

- Insert Figure 1 about here -

The main document contained 1,878 words, four sections distributed across 13 subsections, and 48 embedded links. Links varied in their relevance for each particular question. We considered relevant links those that included information useful to answer a particular question. On the contrary, neither irrelevant or distractor links conveyed information useful to answer a particular question. However, whereas distractor links included at least one word that also appeared in the question, irrelevant links didn't share any word with those in the question (cf. Cerdán et al., 2011). To corroborate our manipulation, in a pilot study 45, undergraduate students rated links' usefulness in answering each particular question on a scale from 0 (totally sure it doesn't include useful information) to 5 (totally sure it does include useful information). Results indicated that relevant links were rated as more useful (*M* = 3.66, *SD* = 1.59) than distractor links (*M* = 2.59, *SD* = 1.80), which in turn were rated as more useful than irrelevant links (*M* = 1.32, *SD* = 1.48).

Hypertext comprehension questions. We constructed six retrieve and six integrate open-ended questions to assess students' comprehension of the hypertext, following the scheme proposed by OECD (2009). Retrieve questions demanded readers to select specific pieces of information from a relevant linked page, and integrate questions required them to connect pieces of information through inferences within relevant linked pages (or within the main hypermedia page and linked pages). The order of presentation of the questions was randomized, and this random sequence was used for all participants. Questions were corrected using a global score of 0 (incorrect) or 1 (correct). Interrater reliability was acceptable (Cohen's κ = .82).

Reading comprehension test. We used the Test of Comprehension Processes (TPC; Vidal-Abarca, et al., 2007) to assess students' comprehension skills. TPC is a standardized paper-and-pencil test (Cronbach's $\alpha = .80$) composed of two expository texts and 10 multiple-choice questions per text. These questions target different comprehension processes as proposed by Kintsch (1998). In the analyses, we used the composite score corresponding to the percentage of correct responses because this test didn't provide individual values of the different factors involved in reading comprehension.

Apparatus

We tracked students' eye movements using a remote eye-tracking system (SMI RED250). Data were registered binocularly at a sampling rate of 250 Hz. Participants were seated in a quiet room located at their school, at approximately 60 cm from the screen. We performed a calibration before starting questions 1, 5, and 9.

Procedure

In the first session (approximately 50 minutes), students completed the reading comprehension test, the prior knowledge test, and a demographic questionnaire. In the second session (approximately 90 minutes), they first practiced the question-answering task with a Wikipedia® page about the solar system. Once they felt confident with the procedure, the experimental task began. Students worked with one question at a time. After reading a question, they were encouraged to navigate through the hypertext on the French Revolution to provide an answer. The question remained visible on top of the web pages (see Figure 1, box on top). Once the student responded, he/she was provided with a different question. After working with the 12 questions, students completed the cued retrospective think-aloud protocol (Van Gog et al., 2005). Specifically, students watched a screen recording video of their learning session that included one dot representing their gaze as they inspected the web pages. Students had to recall what they were thinking during the learning session using the video as a cue. Due to time constraints, students performed the cued retrospective think-aloud task for the first six questions only (three retrieve and three integrate questions).

Coding of Think-Aloud Data

We coded the data from the cued retrospective think-aloud task distinguishing between utterances referring to navigation or to text reading. First, to code navigation utterances, we followed the coding scheme proposed by Coiro and Dobler (2007), which included inferential predictions of a section's relevance, activation of prior knowledge, and self-regulation (setting a plan, monitoring the current understanding, and revisiting the plan when it is deemed inadequate). Second, to code the utterances referring to text reading, we differentiated between those reflecting scanning (e.g., "I skimmed looking for it [the information requested in the question]," "I looked quickly at that text," or deep processing (e.g., "I carefully read all the information [from a particular section]," "I was reading slowly to understand [the text]"). We counted utterances separately for each of the six questions analyzed.

The third and fourth authors coded the utterances of three different students, reaching a reliability of .65 (Holsti's CR). They discussed all disagreements, and subsequently the third author coded the remaining protocols.

Coding of Eye Movement Data

To analyze eye movement data, we designated paragraphs as our main areas of interest (AOI). As a reference, the main Wikipedia® document included one to three paragraphs per section, ranging from one to seven lines of text. AOIs were defined as relevant, distractor, or irrelevant (see the Hypertext System section above). In addition, we defined an AOI corresponding to the table of contents of the main Wikipedia® document. Then, we computed a series of measures based on the eye-tracking data (cf. Kaakinen & Hyönä, 2011). The number of dwells was computed as the number of times a participant visited an AOI (i.e., entered an AOI and exited it). Run dwell time was computed as the sum of all fixation times within an AOI that occurred during the first dwell on an AOI (first run dwell time) or during the second and subsequent dwells on that AOI (second run dwell time).

Results

We ran a series of analyses using question type (locate and integrate) as the independent variable, scores on the reading comprehension skills test as the covariate, and several indexes of scanning and deep processing of text as the dependent variables. Analyses with eye-movement data also included hypertext paragraphs (relevant, irrelevant, and distractors) as independent variables.

Hypotheses 1 & 2: Effects of Reading Comprehension Skills, Task Type, and Scanning on Deep Processing

We first analyzed the effects of reading comprehension skills (Hypothesis 1) and task type (Hypothesis 2) on scanning and deep processing of the hypertext. In the first analysis using data from the verbal protocols, we ran an MANCOVA with question type (locate and integrate) as the independent variable, scores on the reading comprehension skills test as the covariate, and verbal utterances referring to the scanning and deep processing of text (for each question type) as dependent variables. In subsequent analyses, we also included the only verbal utterance related to navigation that was related significantly to hypertext comprehension (as indicated by a significant Spearman correlation) – that is, inferential predictions of paragraph relevance based on textual context. The MANCOVA showed a non-significant multivariate effect of question type for the three variables as a group, $V = .10$, $F(3, 17) = 0.66$, $p = .59$. Similarly, neither the covariate reading comprehension skills nor the interaction with question type reached significance levels ($V = .18$, $F(3, 17) = 1.26$, $p = .32$; $V = .04$, $F(3, 17) = 0.88$, $p = .88$, respectively). Univariate analyses for the effect of each dependent variable showed a similar pattern (all $p > .15$).

In a second analysis using data from eye movements, we ran an ANCOVA with task type (locate and integrate) as the independent variable and hypertext paragraphs (relevant, irrelevant, and distractors) as independent variables, scores on the reading comprehension skills test as the covariate, and number of dwells as the dependent variable. Main effects were not significant for question type or reading comprehension skills (both $p > .10$), but it was significant for paragraph type – $F(1, 19) = 22.31$, $p < .01$. Post-hoc analyses with Bonferroni correction showed that students visit more irrelevant ($M = 5.51$, $SD = 2.55$) than relevant paragraphs ($M = 2.32$, $SD = .81$) ($p < .01$) and fewer distractor paragraphs ($M = .51$, $SD = .38$) (in both cases $p < .01$). This effect was qualified by a significant interaction between paragraph type and reading comprehension skills – $F(1, 19) = 6.12$, $p < .05$. Specifically, there was a negative significant relationship between reading comprehension skills and the number of dwells on irrelevant paragraphs ($r = -.50$, $p < .05$), whereas the correlation with relevant or distractor paragraphs was not significant (both $p > .55$). Finally, no other interaction resulted in significant differences (all $p > .14$).

In the third analysis with dwell times, we ran an ANCOVA with task type (locate and integrate) as the independent variable, and hypertext paragraphs (relevant, irrelevant, and distractors) as independent variables, scores on the reading comprehension skills test as the covariate, and dwell times (first and second run) as the dependent variable. Results showed no significant effects of task type, dwell run, or reading comprehension skills (all $p > .19$) but a significant effect of paragraph – $F(1, 19)= 5.27$, $p= .01$. Post-hoc analyses with Bonferroni correction showed that students spent a similar amount of time reading relevant (M= 4480 *msc*, SD= 2567) and irrelevant paragraphs (M = 3444, SD = 2066) (p =.22), and compared to those, they spent less time reading distractor paragraphs (M = 770, SD = 861) (in both cases p <.01). Those reading times corresponded to high reading speed values (37, 29 and 7 *msc* per word, for relevant, irrelevant, and distractor paragraphs, respectively) (cf. Rainer, 2009), which indicated that, overall, students tended to scan most of the hypertext and didn't perform a complete reading of the paragraphs, not even of relevant ones.

The only significant two-way interaction was that of paragraph and dwell run – $F(2, 38)= 7.01$, $p< .01$ (all other two-way interactions $p> .10$), which was qualified by a three-way interaction between paragraph, dwell run, and reading comprehension skills – $F(2, 38)= 3.83$, $p< .05$ (all other three-way interactions $p> .24$). The interaction was explained by two negative correlations between reading comprehension skills and reading time of irrelevant paragraphs while answering integrate questions: marginal during the first run (r=-.41, p= .07) and significant during the second run (r=-.44, p< .05). No other correlation approached significance (all p > .21).

In sum, data from verbal protocols and eye movements only partially supported Hypothesis 1 in that there was a negative relation between students' reading comprehension skills and the scanning of irrelevant hypertext sections, as revealed by the eye-tracking data. However, the expected positive relation between reading comprehension skills and the deep processing of relevant sections was not found. In the same vein, the data did not support Hypothesis 2: there was no difference between location and integrate questions in terms of scanning irrelevant sections. The eye movement data revealed that the negative relation between reading comprehension skills and the processing of irrelevant paragraphs was only evident with integrate questions, not with locate questions.

Hypothesis 3: Effects of Reading Skills, Question Type and Scanning and Deep Processing on Hypertext Comprehension

Next, we analyzed the effects of reading comprehension skills, task type, and scanning and deep processing on hypertext comprehension (Hypothesis 3). To this aim, we ran three different hierarchical regression analyses for each set of predictors (verbal protocols or eye movements). In the three models, we included as z-standardized values the scores on the reading comprehension test at the student level and type of task (locate and integrate) at the item level. For each of the three models, we varied the different indicators of scanning and deep processing (see below) that were included at the students' level. As the dependent variable, we included the percentage of the correct responses to the hypertext comprehension questions. On average, students' responded correctly on 49.21% (SD = 24.98) of location questions and on 47.62% (SD= 35.85) of integrate questions.

In a first hierarchical regression analysis, we added students' verbal utterances referring to scanning, deep processing, and inferential predictions of paragraph relevance based on textual context as predictors. The number of verbal utterances referring to scanning was negatively related to hypertext comprehension ($t(20)= -3.08$, $B=-0.94$, $SEB=0.31$, $p<.01$). Although verbalizations reflecting the deep processing of information were not related to hypertext comprehension ($t< 1$), the use of textual context to infer sections' relevance (an indirect indicator of deep processing) was marginally related ($t(20)= 1.86$, $B=0.52$, $SEB =0.28$, $p=.06$). Reading comprehension skills were only marginally related to performance ($t(20)= 1.63$, $B=0.45$, $SEB = 0.27$, $p=.10$), while the effect of question type was not significant ($p< .74$). In addition, we analyzed the interaction between the use of scanning and students' reading comprehension skills, which turned out to be significant: $t(20) = -1.90$, $B=-0.44$, $SEB = 0.24$, $p<.05$. A standard way to interpret interactions in multiple regressions is to test for simple slopes (Aiken & West, 1991). Readers more familiar with ANOVA can think of simple slopes as the equivalent of simple effects in ANOVA. To test a significant interaction in multiple regressions through simple slopes requires that the original equation be recast as the regression of the criterion for one predictor. Usually two equations are run, in which the slope (in our case, the relation between scanning and hypertext comprehension) is tested at the value of one standard deviation above or below the predictor mean (in our case, the group mean on the reading comprehension skills test). For poor readers (1 *SD* below the group mean), scanning thorough the hypertext didn't affect their hypertext comprehension ($p= .16$). By contrast, for skilled readers (1 *SD* above the group mean), the use of scanning led to a strong decrease in hypertext comprehension ($p< .01$) (Figure 2). No other interactions were significant.

- Insert Figure 2 about here -

In a second hierarchical regression analysis we used the number of dwells to paragraphs (relevant, distractor, irrelevant) at the student level as predictors. Hypertext comprehension was negatively related to the number of dwells on irrelevant paragraphs ($t(20)= -2.67$, $B=-1.09$, $SEB=0.41$, $p<.01$), positively related to the number of dwells on relevant paragraphs ($t(20)= 3.93$, $B=0.85$, $SEB = 0.22$, $p<.01$), and not related to the number of dwells on distractors ($p= .33$). In this model, reading comprehension skills were positively related to hypertext comprehension ($t(20)= 2.19$, $B=0.59$, $SEB = 0.20$, $p<.05$). Finally, the effect of question type was not significant ($p= .47$). Furthermore, we analyzed the interaction terms of the hierarchical linear model. The interaction between number of dwells on irrelevant paragraphs and reading comprehension skills was significant: $t(20) = -1.95$, $B=-0.41$, $SEB = 0.22$, $p<.05$. For poor comprehenders (1 *SD* below the mean), visits to irrelevant paragraphs didn't affect their final performance ($p=.10$), but for skilled comprehenders (1 *SD* above the mean), those visits had a detrimental effect ($p<.01$). In addition, the interaction between question type and number of dwells on relevant paragraphs turned out to be significant as well: $t(20) = 2.63$, $B=0.53$, $SEB = 0.20$, $p<.01$. While for location questions there was a strong positive relationship between number of dwells on relevant paragraphs and hypertext comprehension ($p< .01$), for integrate questions this relation was not significant ($p<.20$). No other interactions were significant.

In a third hierarchical regression analysis, we used as predictors the variable dwell times on paragraphs (relevant, distractor, irrelevant), distinguishing between first and second runs. Hypertext comprehension didn't vary as a function of dwell times for the first visits to the three types of paragraphs (all $p> .18$). Nevertheless, hypertext comprehension was negatively related to dwell times for the second and subsequent visits to irrelevant paragraphs ($t(20) = -2.34$, $B=-3.76$, $SEB=1.61$, $p<.05$), positively related to dwell times for the second and subsequent visits to relevant paragraphs ($t(20) = 2.40$, $B=0.47$, $SEB=0.19$, $p<.05$), and not related to dwell times on distractor paragraphs ($p= .14$). In this model, hypertext comprehension was not related to either reading comprehension skills ($p=.41$) or question type ($p=.10$). In addition, we looked at the interaction terms of the hierarchical linear model. Results revealed a significant interaction between dwells times during the second and subsequent visits to irrelevant paragraphs and students' reading comprehension skills: $t(20) = -2.27$, $B=-1.74$, $SEB = 0.76$, $p<.05$. For poor comprehenders (1 *SD* below the mean), dwell times during the second and subsequent visits to irrelevant paragraphs were not apparently related to their hypertext comprehension ($p=.09$), but for skilled comprehenders (1 *SD* above the mean), it showed a significant negative relation (p<.01). No other interactions resulted in significant effects.

In sum, the data from verbal protocols and eye tracking only partially supported Hypothesis 3, and they revealed unexpected results. The expected three-way interaction between question type, reading comprehension skills, and deep processing was not significant in any of the measures taken. Specifically, results revealed that verbal utterances reflecting the use of textual context to infer sections' relevance, visits to relevant paragraphs, and reading times during revisits to relevant paragraphs led to better hypertext comprehension, regardless of students' reading comprehension skills. Data from the number of dwells revealed that this positive effect was evident for locate questions, but not for integrate questions. In addition, the analyses showed an unexpected result related to scanning and reading comprehension skills. Verbal utterances reflecting text scanning, as well as revisits to and longer rereading times of irrelevant paragraphs, led to lower hypertext comprehension. Interestingly, this effect was qualified by students' reading comprehension skills: while the effect of scanning was null for poor comprehenders, it was detrimental for good comprehenders.

Discussion

The results from our study provide new evidence on how reading comprehension skills and task type affect the way students scan and process hypertext, as well as how this behavior relates to hypertext comprehension. In the coming section we discuss the relationships between reading comprehension skills, task type, and navigation, emphasizing how our results may add to the current knowledge of what means to be a good comprehender in Internet problem solving tasks. Then, we address potential instructional applications of the results. Finally, we discuss some limitations of the study and propose future lines of research.

Factors Influencing Scanning and Deep Processing of Hypertext

How do reading comprehension skills and question type influence the way students scan and process hypertext while answering a question? Efficient hypertext readers quickly scan sections of irrelevant information for the question they are trying to answer, while carefully processing those parts of the hypertext containing information that is more relevant. Results from our study provided only partial support for this view. On the one hand, highly skilled readers quickly scan and revisit less often sections of the hypertext that don't contain relevant information, especially on integrate questions. Discarding irrelevant information may be a less demanding activity when the students' goal is to simply locate a specific piece of information (Rouet, 2003). Distinguishing between less and more relevant sections when the question demands the integration of separate pieces of information distributed across different hypertext sections may require the student to process the information at a deeper level (Cerdán & Vidal-Abarca, 2008), which may be supported by students' reading comprehension skills.

On the other hand, our results do not show any evidence of a positive relation between reading comprehension skills and the deep processing of relevant sections, contrary to what has been reported in prior hypertext studies that analyzed highly proficient readers (Coiro & Dobler, 2007; Zhang & Duke, 2008). A possible explanation for this lack of effect is the fact that we studied a sample of students with a normal distribution of reading comprehension skills. In this case, students with medium to high reading comprehension skills may require some navigation training in order to take advantage of their reading skills to better identify relevant sections within hypertext (Authors, 2008b). Although the participating students on average have been working with computers for more than five years, they indicated that they read pages from Wikipedia® only once or twice a month.

Factors Influencing Hypertext Comprehension

How do reading comprehension skills, question type, and scanning and deep processing affect hypertext comprehension? Results show that high hypertext comprehension is related to the processing of relevant hypertext sections, as indicated by students' use of textual context to infer sections' relevance, inspections, and reading times during revisits to relevant paragraphs. This effect is in line with a meta-analysis of hypertext studies using log files as an indicator of navigation (Authors, 2008a), which report a correlation of .29 between visits to pages relevant to the task to be solved and performance. Interestingly, our results qualify this effect: data on number of inspections to relevant paragraphs revealed that the positive relation with hypertext comprehension was evident for locate questions but not for integrate questions. These questions require the student to not only locate and identify relevant information from separate sections but also to integrate them to provide a coherent response (Cerdán & Vidal-Abarca, 2008). Nevertheless, contrary to what was expected, students' reading comprehension skills do not help students who process relevant sections to answer integrate questions specifically. Actually, the overall positive effect of reading comprehension skills on hypertext comprehension in two of the three regression models analyzed supported the claim that printed reading comprehension skills may be a prerequisite for efficient navigation in Internet problem solving (Brand-Gruwel et al., 2009).

Unexpectedly, the results showed that the scanning of irrelevant hypertext sections reflected by verbal utterances reflecting text scanning, as well as by revisits and longer rereading times of irrelevant paragraphs, lead to lower hypertext comprehension. This effect is particularly evident in students with high reading comprehension skills. This effect suggests that students' skills do not automatically help in solving Internet problem solving tasks, as previously suggested (e.g., Authors, 2011b; Coiro, 2011). Why did some of the good comprehenders in our study fail to employ efficient navigation on some questions? A potential explanation may be related to the cognitive mechanisms used by good comprehenders. These students are highly fluent in several aspects of text processing, such as idea identification, inference generation, or macroidea elaboration (Perfetti, 2007). When it comes to text searches, they might also show better task performance than less proficient comprehenders if they are allowed to read the text first (Authors, 2011a). This way, they can profit from their efficient text processing and may use the generated text representation to guide their search (Payne & Reader, 2006). However, this does not necessarily mean that they would be good searchers in the absence of such representation. Indeed, when middle and high school students are given the opportunity to either read a short text first or to search the questions first, good comprehenders tend to read the text first and then look at the questions to search (Authors, 2014), suggesting that they may not feel confident searching without the background of a mental text representation. In sum, while proficiency with printed texts still allows students to partially succeed in Internet problem solving tasks (Brand-Gruwel et al., 2009), our results suggest that there are other components related to search behavior that make a student an expert reader and navigator in the digital medium, such as being able to use contextual cues to predict text relevance.

Instructional Applications

A major skill to be a proficient reader of hypertext is to regulate the scanning and deep processing of text to correctly identify text sections relevant to the learning task while scanning a hypertext by means of analyzing contextual cues, such as the introduction to a section or paragraph headings, and to immediately process those particular sections.

One way to instruct students to navigate efficiently is by means of eye movement modeling examples, which display experts or successful students’ gaze and verbalizations while performing a particular task on the computer. These videos have recently been successfully applied to different instructional settings to foster procedural learning such as medical diagnosis (Jarodzka et al., 2012), fish locomotion (Jarodzka, Van Gog, Dorr, Scheiter, & Gerjets, 2013), or strategic processes such as text–picture integration (Pluchino, Tornatora, & Mason, 2013). The results from our study could be used to generate models of efficient strategies displaying students’ gaze and verbalizations.

Limitations and Future Research

It should be noted that in our study most participants had a low level of prior knowledge on the topic studied. Recently, Coiro (2011) found a positive relationship between high school students’ prior knowledge and hypertext comprehension, which was independent of their reading comprehension skills. In a similar vein, Rouet, Ros, Goumi, Macedo-Rouet, and Dinet (2011) have shown that activating high school students’ prior knowledge on a topic by providing them with a summary of the relevant topic information before solving topic-related tasks, improves their assessment of hyperlink relevance in a search engine results pages. Future research should extend those results to explore how prior knowledge may influence the way adolescents scan and deep process hypertext during Internet problem solving tasks.

Another limitation of the current study is that we could experimentally manipulate the type of question but not the students’ reading comprehension skills. Therefore, any causal claim made regarding the role of students’ skills is just tentative and is open to further exploration. In addition, we studied a limited sample of participants, as is usually the case in studies that involve time consuming analyses such as think-aloud protocols or eye tracking. Future studies should aim to replicate our findings with different materials and different grade levels. Our future efforts will proceed in this direction.

References

Aiken, S. & West, G. (1991).*Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Authors, 2008a.

Authors, 2008b.

Authors, 2009.

Authors, 2011a.

Authors, 2011b.

Authors, 2014.

Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education, 53*, 1207-1217. doi:10.1016/j.compedu.2009.06.004

Cerdán, R., & Vidal-Abarca, E. (2008). The effects of tasks on integrating information from multiple documents. *Journal of Educational Psychology, 100(1)*, 209-222. doi: 10.1037/0022-0663.100.1.209

Cerdán, R., Vidal-Abarca, E., Martínez, T., Gilabert, R., & Gil, L. (2009). Impact of question-answering tasks on search processes and reading comprehension. *Learning and Instruction, 19*, 13–27. doi: 10.1016/j.learninstruc.2007.12.003

Coiro, J. (2011). Predicting reading comprehension on the internet contributions of offline reading skills, online reading skills, and prior knowledge. *Journal of Literacy Research, 43(4)*, 352-392. doi:10.1177/1086296X11421979

Coiro, J. & Dobler, E. (2007). Exploring the online comprehension strategies used by sixth-grade skilled readers to search for and locate information on the Internet. *Reading Research Quarterly, 42*, 214-257. doi:10.1598/RRQ.42.2.2

Cromley, J. G., & Azevedo, R. (2009). Locating information within extended hypermedia. *Educational Technology Research and Development, 57*(3), 287-313. doi: 10.1007/s11423-008-9106-5

Duggan, G. B., & Payne, S. J. (2009). Text skimming: The process and effectiveness of foraging through text under time pressure. *Journal of Experimental Psychology: Applied, 15*(3), 228-242. doi: 10.1037/a0016995

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis*. Cambridge, MA: MIT Press.

Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, A. C., & Brodowinska, K. (2012). Comprehending and learning from internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly, 47*(4), 356-381. doi: 10.1002/RRQ.027

Jarodzka, H., Balslev, T., Holmqvist, K., Nyström, M., Scheiter, K., Gerjets, P., & Eika, B. (2012). Conveying clinical reasoning based on visual observation via eye-movement modelling examples. *Instructional Science, 40(5),* 813-827. doi: 10.1007/s11251-012-9218-5

Jarodzka, H., Van Gog, T., Dorr, M., Scheiter, K., & Gerjets, P. (2013). Learning to see: Guiding students' attention via a model's eye movements fosters learning. *Learning and Instruction, 25*, 62-70. doi:10.1016/j.learninstruc.2012.11.004

Kaakinen, J.K., & Hyönä, J. (2011). Online processing of and memory for perspective relevant and irrelevant text information. In M.T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 223-242). Charlotte, NC: Information Age Publishing.

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

Organisation for Economic Co-operation and Development (OECD) (2011). *PISA 2009 results: Students on line digital technologies and performance (Volume VI).* PISA, OECD Publishing. doi: 10.1787/9789264112995-en

Payne, S. J. & Reader, W. R. (2006). Constructing structure maps of multiple on-line texts. *International Journal of Human Computer Studies, 64*, 461-474. doi: 10.1016/j.ijhcs.2005.09.003

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383. doi: 10.1080/10888430701530730

Pluchino, P., Tornatora, M. C., & Mason, L. (2013). *Improving text and picture integration during reading through eye-movement modeling.* Talk presentated at the 2013 meeting of the European Associtíon for Learning and Instruction, Munich, Germany.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology 62*, 1457–1506. doi: 10.1080/17470210902816461

Rouet, J. F. (2003). What was I looking for? The influence of task specificity and prior knowledge on students' search strategies in hypertext. *Interacting with Computers, 15(3)*, 409-428. doi: 10.1016/S0953-5438(02)00064-4

Rouet, J.-F., Ros, C., Goumi, A., Macedo-Rouet, M. & Dinet, J. (2011). The influence of surface and deep cues on primary and secondary school students' assessment of relevance in Web menus. *Learning and Instruction, 21*, 205-219. doi:10.1016/j.learninstruc.2010.02.007

Rouet, J.-F., Vidal-Abarca, E., Bert-Erboul, A., & Millogo, V. (2001). Effects of information search tasks on the comprehension of instructional text. *Discourse Processes, 31, 163-186*. doi: 10.1207/S15326950DP3102_03

Tapia, J. (2004). *La Revolución Francesa [The French Revolution]*. Retrieved from http://ntic.educacion.es/w3//eos/MaterialesEducativos/bachillerato/historia/revfran/index.htm

Van Gog, T., Paas, F., Van Merriënboer, J. G., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied, 11*(4), 237-244. doi: 10.1037/1076-898X.11.4.237

Vidal-Abarca, E., Gilabert, R. Martínez, T., Sellés. P., Abad. N., & Ferrer, C. (2007). *TPC. Test de Procesos de Comprensión* (Test of Comprehension Strategies). Madrid, ICCE.

Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B.Zimmerman & D.Schunk (Eds.),*Self-regulated learning and academic achievement: Theoretical perspectives*, (2nd ed., pp.153-189). Mahwah, NJ: Erlbaum.

Zhang, S., & Duke, N. K. (2008). Strategies for internet reading with different reading purposes: A descriptive study of twelve good internet readers. *Journal of Literacy Research, 40*(1), 128-162. doi: 10.1080/10862960802070491

Figure 1. Screen capture of the main Wikipedia document used in the study.

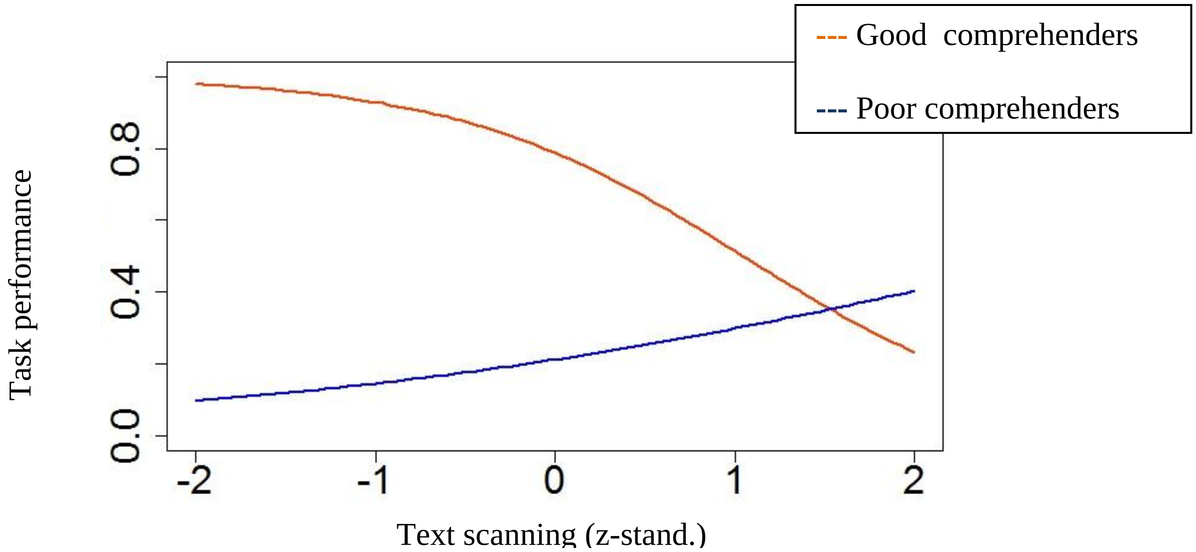

Figure 2. Effect of reading comprehension skills and use of text scanning, as revealed by think aloud protocols, on task performance.