The Influence of Text Annotation Tools on Print and Digital Reading Comprehension

Gal Ben-Yehudah
The Open University of Israel
galby@openu.ac.il

Yoram Eshet-Alkalai
The Open University of Israel
yorames@openu.ac.il

Abstract

Recent studies that compared reading comprehension between print and digital displays found a disadvantage for digital reading. A separate line of research has shown that effective use of learning strategies and active learning tools (e.g. annotation tools: highlighting, underlining and note-taking) can improve reading comprehension. Here, the influence of annotation tools on comprehension and their utility in diminishing the gap between print and digital reading was investigated. In a 2 X 2 experimental design, ninety three undergraduates (mean age 29.6 years, 72 women) read an academic essay in print or in digital format, with or without the use of annotation tools. In general, the results reinforce previous research reports on the inferiority of digital compared to print reading comprehension. Our findings indicate that for factual-level questions using annotation tools had no effect on comprehension in both formats. For the inference-level questions, however, use of annotation tools improved comprehension of printed text but not of digital text. In other words, text-annotation in the digital format had no effect on reading comprehension. In the digital era, digital reading is the norm rather than the exception; thus, it is essential to improve our understanding of the factors that influence the comprehension of digital text.

Keywords: digital reading, print reading, annotation tools, learning strategies, self regulated learning, monitoring.

Introduction

In recent years, information consumers are faced with a rapid growth in the availability of digital in lieu of printed text, as evidenced from the proliferation of online textbooks, newspapers, encyclopedias and online academic journals (Cargill, 2011; Hamblen, 2011; Heider, Laverick, & Bennett, 2009; The Economist, 2011). This shift, from print towards digital reading, is especially critical in the higher education and the school systems, where learning from digital texts becomes increasingly more common (Cargill, 2011; Heider et al., 2009; Thayer et al., 2011).

Reading from digital displays poses a wide range of challenges for readers (Altonen, Mannonen, Nieminen & Nieminen, 2011; Bus & Neuman, 2009; Quinn & Stark-Adam, 2007; Van Den Broek, Kendeou & White, 2009), mainly because of the large reading distance from a computer screen (as opposed to the short reading distance from a printed book), the long lines of text on the modern wide computer screens and the problem of shifting the eye gaze from line to line while reading (Evans, Charland & Saint-Aubin, 2009). In addition, text fragmentation, associated with the nonlinear nature of hypertext, results in a decrease in text coherence (Albrecht & O'Brien, 1993; Ozuru, Dempsey & McNamara, 2009). Reduced text coherence in digital displays creates disorientation (Armitage, Wilson & Sharp, 2004), presents readers with a high cognitive load (Ackerman & Goldsmith, 2011; Rouet, 2000) and harms text

comprehension (Chang & Ley, 2006; Rouet, 2000; Van den Broek et al, 2009). Consequently, many studies report that text comprehension from digital displays is inferior to comprehension from print (e.g. Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012; Eshet-Alkalai & Geri, 2007; 2010). Ackerman and Goldsmith (2011) claim that the differences between print and digital comprehension result mainly from differences in Self Regulated Learning (SRL) that are dictated by these two media. Some studies (e.g. Lauterman & Ackerman, 2013; Thiede, Anderson & Therriault, 2003) report that under specific conditions, learning strategies, such as writing keywords, can improve monitoring of the learning process; SRL, in turn, improves the quality of learning.

The recent surge in learning from digital displays and the accumulating evidence on the inferiority of comprehension from digital texts (Ackerman & Goldsmith, 2011; Eshet-Alkalai & Geri, 2007), raises the need for a better understanding of the role that learning strategies and active learning tools play in digital reading. The wealth of studies on printed text suggest that a reader can improve monitoring of the learning process by utilizing learning tools such as annotation, highlighting, writing keywords, summaries and reflection (e.g., Pressley, 2000). Recently, studies comparing learning strategies, such as the use of active learning tools for print and digital reading, have offered the following observations:

- **Keywords:** Thiede et al. (2003) reported that writing keywords after reading an essay, improved the readers' monitoring of the learning process and their SRL. Lauterman and Ackerman (2013) elaborated on Thiede et al.'s work and showed that utilizing keywords as a learning strategy can help reduce differences in text comprehension and SRL between print and digital text-reading**.**
- **Note taking**: Oppenheimer and Mueller (2012) explored the effect of note-taking during a lecture as a learning strategy. They showed that students who took lecture notes by pen-writing typically showed a deeper processing of the content in their notes and performed better than those who took notes on a laptop computer, mostly by transcribing the lecture "as is".
- **Text highlighting and underlining:** Fowler and Barker (1974) found that the act of highlighting text, while reading from paper, improved retention and understanding. By distinguishing between **active highlighting** (i.e., readers actively select and highlight important parts of the text) and **passive highlighting** (i.e., readers read an essay, in which important sections were already highlighted for them) they were able to show that active highlighting was a more effective learning strategy than a passive one. Underlining was also found to be an effective strategy for improving retention, but it was inferior to highlighting. Leutner et al. (2007) found that a computer-based training program designed to improve readers' use of text-highlighting and self regulation led to the greatest improvements in text comprehension, relative to other training programs. In their comparison between readers from electronic devices and readers from paper, Schugar, Schugar and Penny (2011) discovered that print readers employed active learning strategies (highlighting, annotating and bookmarking), whereas, e-readers rarely utilized these strategies.

The current study focused on the effect of text-highlighting and annotation (i.e. taking notes directly on the text page), as learning strategies, on text-comprehension from print and digital displays. Our study differs from Leutner et al. (2007), because we did not train the participants on these strategies prior to the test phase. However, they were encouraged to use highlighting and annotation tools to improve their comprehension of the text. According to our literature review, this is the first study which directly compares the effect of active highlighting and annotation on comprehension of print versus digital text. Present day students tend to take lecture-notes on laptops, which offer computer programs with a wide variety of active learning tools (e.g. highlighting, commenting and tagging). Therefore, it is vital that we improve our understanding of these learning tools and their utility for monitoring learning in the digital age.

Methodology

Participants: Ninety three students from the Open University of Israel participated in the study. The participants were 72 women and 28 men; all of them were studying towards a bachelorette degree. The participants age ranged between 20 and 54 years (M = 29.6). All of the participants were native Hebrew speakers, had normal or corrected-to-normal vision and received course credit for their time. They were randomly assigned to one of four experimental groups (N was approximately 23 per group).

The experimental groups did not differ on their mean age, level of higher education or reading comprehension ability (see below "Technologies in Medicine" test). In all of the groups, a large majority of the participants indicated that they had little or no prior knowledge in the topic of the test material (geology).

Materials: Participants completed three tasks in the following order:

- A demographic questionnaire
- A test of text comprehension ("Technologies in Medicine" task), which served as a baseline measure for participants' level of reading comprehension
- An experimental text-comprehension task ("Fossils" task)

The demographic questionnaire included questions on language history, daily use of technology (including annotation habits) and prior knowledge of geology (the topic of the experimental text). The baseline comprehension measure was the "Technologies in Medicine" test - an essay about new technologies in medicine. This text is used by the Israeli National Institute for Testing and Evaluation as an assessment of college-level comprehension skills. The text consists of 458 words (5 paragraphs, 31 lines), and it is followed by ten multiple choice questions that assess reading comprehension. The "Technologies in Medicine" test is administered in print and without time constraints. The reliability of this test was acceptable (Cronbach's alpha of 0.614).

The experimental task - the "Fossils" task – was a reading comprehension test on an essay about the formation and scientific importance of fossils. The four-page text, adapted from a college-level text book, had 858 words (19 paragraphs, 71 lines) and 2 text-figures. Participants read the text without time constraints, then the text was removed and they answered ten multiple choice comprehension questions (each with four possible answers). Participants were informed in advance that they will be tested on the text and that it would not be available to them while answering the comprehension questions; therefore, they were encouraged to read the text until they felt ready to answer the comprehension questions. Half of the comprehension questions related to facts that appeared in the text (factual questions), while the other half required a deeper understanding of the text (inference questions). The reliability of the Fossils test, which was obtained during a pilot study, was acceptable (Cronbach's alpha of 0.652).

Procedure: The Fossils task had a 2 × 2 experimental design. The independent variables were the display format (print/digital) and the use of annotation tools (with/without text-highlighting and annotation). The dependent variables were the response accuracy for the comprehension questions, the text reading time (*reading time*) and the time spent answering the questions (*answer time*). Each participant was randomly assigned to one of the four experimental reading conditions: print reading with annotation tools (P+), digital reading with annotation tools (D+), print reading without annotation tools (P-), and digital reading without annotation tools (D-). At the beginning of each task, participants received detailed instructions on the use of annotation tools. A highlighter and a pen were used in the print reading condition and the Adobe Reader annotation tools were used for text-highlighting and commenting in the digital reading

condition. The display format of the comprehension questions was identical to that of the text reading condition. The reading instructions were identical for the print and the digital reading conditions. In the condition with annotation, participants were instructed to use annotation tools (highlighting and comments) during reading to improve comprehension and memory. In the condition without annotation, participants were explicitly told that the use of annotation tools is not allowed. All of the participants were tested in the same environment (a university computer room). Participants in the digital reading condition performed the task on desktop computers, with a 19" screen.

Results

The majority of the participants reported of an infrequent use of annotation tools; 63.4% indicated that they use them sometimes or very rarely.

Table 1 shows the mean performance on each experimental condition (i.e., P+, P-, D+, D-) for the three dependent variables (i.e., accuracy, reading time and answer time). To explore the effect of our experimental manipulation on performance, analyses of variance were conducted for each dependent measure. Results show a significant effect of the experimental condition on all of the dependent measures: accuracy ($F(3,89) = 2.73$, $p < .05$), reading time ($F(3,88) = 3.1$, $p < .05$), and answer time ($F(3,87) = 2.86$, $p < .05$). Planned independent samples t-tests (1-tailed) were conducted to investigate the source of this effect. Consistent with previous studies, text comprehension was found to be better for printed relative to digital displays. This difference was significant only for the annotation condition (P+ group vs. D+ group), on both accuracy ($t(47) = 2.8$, $p < .005$) and answer time ($t(45) = -2.8$, $p < .005$). In other words, we found that using annotation tools while reading a printed text leads to more accurate and faster responses than using the same tools when reading a digital text.

Table 1. Mean (SD) performance per condition, for each dependent measure.

Condition	Accuracy (%)	Read (min)	Answer (min)
P+	86.4 (12)	10.2 (4.5)	4.1 (2.1)
D+	73.9 (18)	10.5 (2.9)	5.8 (2.0)
P-	79.6 (16)	7.8 (3.7)	4.3 (1.9)
D-	73.8 (24)	8.3 (3.2)	5.7 (3.8)

As expected, the use of annotation tools significantly increased text reading time in both formats, by approximately 29% relative to reading the text without annotation tools (P+ vs. P-, $t(45) = 1.9$, $p = .06$; D+ vs. D-, $t(43) = 2.4$, $p = .02$).

An interesting pattern emerged from the analysis of performance according to question type (factual/inference) (Figure 1). For the factual questions, there was no significant effect of the experimental condition on accuracy. For the inference questions, however, there was a significant effect of condition on accuracy ($F(3,56) = 3.1$, $p < .05$). To explore the source of this effect, planned independent samples t-tests (1-tailed) were conducted on the inference data. These analyses showed that display format (print/digital) significantly influenced accuracy in the “annotation” condition (P+ vs. D+, $t(47) = 2.3$, $p < .05$), but only marginally in the "without annotation" condition (P- vs. D-, $t(42) = 1.4$, $p = .08$). Regarding the overall effect of annotation on performance, surprisingly, it was significant only for the printed text condition (P+ vs. P-, $t(46) = 1.7$, $p < .05$).

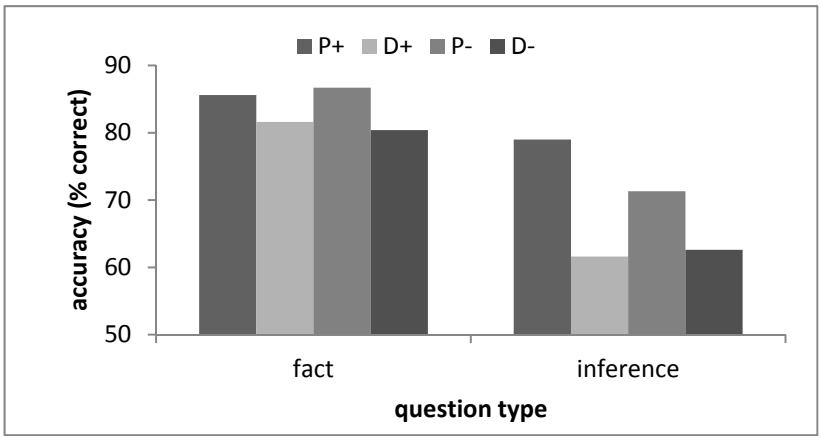

Figure 1. Mean accuracy for each reading condition by question type (factual, inference). The four experimental conditions are: reading with the use of annotation tools for print (P+) and digital (D+) text, and reading without use of annotation tools for print (P-) and digital (D-) text.

Discussion and Conclusions

This study compared the effect of using annotation tools (i.e., text-highlighting, underlining and note-taking) on the comprehension of a college-level text displayed either in print or in digital format. Active text-annotation is known to be an effective learning strategy for improving comprehension of printed text (Fowler & Barker, 1974), but readers of digital texts rarely employ this strategy (Schugar et al., 2011). Therefore, in this study, we decided to compel readers to employ annotation tools so that we could examine whether using this strategy would indeed improve the comprehension of digital texts.

In general, this study replicated previous findings, showing that comprehension of a printed text is better than comprehension of a digital one (Ackerman & Goldsmith, 2011; Eshet-Alkalai & Geri, 2007; 2010). However, contrary to our expectation, using annotation tools to improve comprehension had a limited effect on performance. For the print reading condition, we found that highlighting improved comprehension only for the inference questions and had no effect on the performance for factual questions. This finding suggests that employing an active learning strategy during reading from print is effective mainly for deeper processing and understanding of the text, and not for learning that is based on memorization of facts. For the digital reading condition, highlighting did not improve comprehension at all.

Various studies (e.g. Leutner et al., 2007) have shown that by training readers to utilize effectively learning strategies, one can improve SRL and reading comprehension. Unfortunately, Leuner's study tested SRL and comprehension for printed texts and did not include digital ones. In our study, we did not train our participants; however, we did encourage them to use annotation tools while reading. Despite this emphasis, the expected positive influence of text-highlighting and note-taking on performance was limited to print reading.

Our findings can be discussed from three different perspectives: the **usability perspective**, the **cognitive perspective** and the **information economics perspective**. According to **the usability perspective**, our findings could be interpreted in terms of usage proficiency, following claims that the experience and comfort that users feel in a learning environment play a pivotal role in their performances (Eshet-Alkalai & Chajut, 2010; Hargittai, 2002; Shneiderman, 1998). The infrequent use of annotation tools during digital reading, which was reported by most of the participants in our study, could explain the limited effect of these tools on comprehension.

Correspondingly, the familiarity of our participants with the use of annotation tools for printed text could explain the positive impact of these learning tools on the comprehension of printed text. Since sample size was limited in this study, we could not analyze the effect of age on participants' performance; however, we plan to do this in a follow-up study.

There is strong evidence that reading from a digital display involves a higher cognitive load than print reading (Ackerman & Goldsmith, 2011; Chang & Ley, 2006; Niederhauser et al., 2000; Rouet, 2000). According to the **cognitive perspective**, text comprehension depends on the interaction between the structure of the text and the cognitive abilities of the reader. The structural differences between print and digital formats have a dramatic effect on readers' ability to self regulate learning and effectively employ monitoring strategies, which influence their level of comprehension (e.g. Ackerman & Goldsmith, 2011; Rouet, 2000). According to this perspective, our finding, that annotation did not influence comprehension in the digital condition could be explained, at least partially, by the high cognitive load that results from the structure of a digital text.

The results of this study confirm the basic premise of **information economics**, which is that information does not have an absolute universal value (Ahituv, 1989). Since learners differ in their preferences for information formats, their performance is affected by the media in which the information is presented (Speier et al, 2003). In this respect, our findings support claims by Davitt Maughan (2001) and Saranto and Hovenga (2004) that the familiarity of information consumers with the information display format increases the real value they assign to the information, and consequently influences their comprehension and ability to retain information.

Acknowledgments

This research was supported by a grant to the authors from the Open University of Israel's research fund. The authors thank Iris Fait for her dedicated assistance in data collection and analysis.

References

Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning: On screen versus on paper. *Journal of Experimental Psychology: Applied, 17*, 18-32.

Ackerman, R., & Lauterman, T. (2012). Taking reading comprehension exams on screen or on paper? A metacognitive analysis of learning texts under time pressure. *Computers in Human Behavior, 28*, 1816-1828.

Ahituv, N. (1989). Assessing the value of information: Problems and approaches, *Proceedings of the 10th Annual International Conference on Information Systems*, Boston, MA, (December) 1989, 315-325.

Albrecht, J., & O'brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Learning, Memory, and Cognition, 19*(5), 1061-1070.

Altonen, M, Mannonen, P., Nieminen, S., & Nieminen, M. (2011). Usability and compatibility of e-book readers in an academic environment: A collaborative study. *International Federation of Library Associations, 37*(1), 16-27.

Armitage U., Wilson, S., & Sharp, H. (2004). Navigation and ownership for learning in electronic texts: An experimental study. *Electronic Journal of E-Learning, 2*(2): http://www.ejel.org/volume-2/vol2-issue1/issue1-art17.htm.

Bus, A. G., & Neuman, S. B. (editors) (2009). *Multimedia literacy development*. Routledge, New York.

Cargill, S. (2011). Infographic: Textbooks of tomorrow. *Getting Smart, November, 11, 2011.* http://gettingsmart.com/news/2011/11/infographic-textbooks-of-tomorrow.

Chang, S. L., & Ley, K. (2006). Learning strategy to compensate for cognitive overload in online learner use of printed online materials. *Journal of Interactive Online Learning, 5*(1)*,* 104-117.

Davitt Maughan, P. (2001). Assessing information literacy among undergraduates: A discussion of the literature and the University of California-Berkeley assessment experience, *College & Research Libraries 62*(1) (2001), 71-85.

Eshet-Alkalai, Y., & Geri, N. (2007). Does the medium affect the message? The influence of text representation format on critical thinking, *Human Systems Management, 26*, 269-279.

Eshet-Alkalai, Y., & Geri, N. (2010). Does the medium affect the message? The effect of congruent versus incongruent display on critical reading. *Human Systems Management, 29*(4), 243-251. DOI:10.3233/HSM-2010-0730.

Eshet-Alkalai, Y. & Chajut, E. (2010). You Can Teach Old Dogs New Tricks: The Factors That Affect Changes over Time in Digital Literacy. *Journal of Information Technology Education, 9*: 173-181.

Evans, M. A., Charland, A. R., & Saint-Aubin, J. (2009). A new look at an old format: Eye-tracking studies of shared book reading and implications for eBook and eBook research. In Bus, G. and Neuman, S. (eds.) *Multimedia and Literacy Development*. Routledge: New York. pp. 89-111

Fowler R. L., & Barker, A. S. (1974). Effectiveness of highlighting for retention of text material. *Journal of Applied Psychology, 59*(3)*,* 358-364

Hamblen, M. (2011). Amazon: E-books now outsell print books. *Computerworld*, May, 19, 2011. http://www.computerworld.com/s/article/9216869/Amazon_E_books_now_outsell_print_books.

Heider, K., Laverick, D., & Bennett, B. (2009). *Digital textbooks: The next paradigm shift in higher education?*. *AACE Journal. 17*(2), 103-112. http://www.editlib.org/p/27054.

Lauterman, T., & Ackerman, R. (2013). Overcoming the screen inferiority in text learning. In Y. Eshet-Alkalai, A. Caspi, S. Eden, N. Geri, Y. Kalman, & Y. Yair (Eds.), Proceedings of *the Chais conference on instructional technologies research 2013: Learning in the technological era* (Vol. 8, pp. 100-106). Raanana, Israel: Open University of Israel.

Leutner, D., Leopold, C., & Elzen-Rump, V. (2007). Self-Regulated Learning with a Text-Highlighting Strategy: A Training Experiment . *Journal of Psychology, 215*(3): 174-182.

Niederhauser, D. D., Reynolds, R. E., Salmen, D., & Skolmoski, P. (2000). The influence of cognitive load on learning from hypertext, *Journal of Educational Computing Research 23*(3) (2000), 237-255.

Oppenheimer, D. M., & Mueller, P. A. (2012). The Pen Is Mightier than the Keyboard: Longhand and Laptop Note-Taking. Talk given at the Psychonomic Society Annual Meeting, Minneapolis, MN. (see p. 34)

Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction, 19,* 228-242.

Pressley, M. (2000). What should comprehension instruction be the instruction of? In M. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.). *Handbook of Reading Research* (Vol. III). Mahwah, NJ: Lawrence Erlbaum.

Quinn, S., & Stark-Adam, P. (2007). What are the differences in reading news in print and online?. *The Poynteronline,* April, 2007.

Rouet, J. F. (2000). Hypermedia and learning – cognitive perspectives. *Journal of Computer Assisted Learnin*g, *16*, 97-101.

Saranto, K., & Hovenga, E. J. (2004). Information literacy-what it is about? Literature review of the concept and the context. *International Journal of Medical Information 73*(6), 503-513.

Schugar, J. T., Schugar, H., & Penny, C. (2011). A nook or a book: Comparing college students' reading comprehension level, critical reading, and study skills. *International Journal of Technology in Teaching and Learning, 7(2),* 174-192.

Shneiderman, B. (1998). Designing the User Interface: Strategies for Effective Human-Computer Interaction, 3rd ed., Addison Wesley, Reading, Massachusetts, 1998, pp. 67-78.

Speier, C., Vessey, I., & Valacich, J. S. (2003). The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance, *Decision Sciences* 34(4) (2003), 771-797.

Thayer, A., Lee, C.P, Hwang, L. H., Sales, H., Sen, P., & Dalal, N. (2011). The imposition and superimposition of digital reading technology: the academic potential of e-readers. Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems, ACM, New York, NY, USA. doi>10.1145/1978942.1979375.

The Economist (2011). The books business: Great digital expectations. http://www.economist.com/node/21528611 [accessed, October, 2011]

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95(1),* 66-73. doi: 10.1037/0022-0663.95.1.66

Van Den Broek, P., Kendeou, P. & White, M. J. (2009). Cognitive processes during reading: Implications for the use of multimedia to foster reading comprehension. In G. Bus, &. S. Neuman, *Multimedia and Literacy Development. Routledge*: New York. pp 57-74.