Is Scrolling Disrupting While Reading?

Katherine Brady, Sun Joo Cho, Gayathri Narasimham, Douglas Fisher, Amanda Goodwin Vanderbilt University

katherine.a.brady@vanderbilt.edu, sj.cho@vanderbilt.edu, gayathri.narasimham@vanderbilt.edu, douglas.h.fisher@vanderbilt.edu, amanda.goodwin@vanderbilt.edu

Abstract: The relationship between scrolling and comprehension has been studied a few times by comparing readers who read in an environment with scrolling to those who read in an environment without scrolling. However, the amount of scrolling done by subjects who had the option to scroll has not been linked to comprehension before. In this paper, we investigate the relationship between how often participants scroll and comprehension. We find that participants who have no preference between paper and digital reading tools do worse when they scroll, but there is no relationship between scrolling and comprehension for other students.

Introduction

As we enter the digital era, it is important to understand how design choices in digital reading environments affect comprehension. In this study, we examine how scrolling, the most popular method of reading long text passages on digital devices, correlates with comprehension. Previous research has found some evidence suggesting that scrolling is worse for comprehension than paging. Research comparing reading on paper to reading on digital devices found that there is no significant difference when the stimuli fits on one screen, but students do worse on digital devices when the stimuli takes up more than one screen (Kim & Huynh, 2008; Singer & Alexander, 2017). Other research suggests comprehension is better for digital readers using a paging interface than those using a scrolling interface (Piolat, Roussey, & Thunin, 1997; Sanchez & Wiley, 2009). These findings are in line with research that shows we use spatial markers to remember what we have read (Baccino, 1994; Rothkopf, 1971; Weger & Inhoff, 2007). When these spatial markers move, as happens while scrolling, readers cannot remember what they read as well, or find previously read passages as quickly.

This study is the first to examine how scrolling frequency is linked to comprehension. If scrolling hurts comprehension, then learners who scroll frequently while reading should perform worse on an open book quiz than those who move the text less often. We recorded the screens of 381 5th to 8th graders in the United States while they read a passage on a laptop and then took an open book test. Scrolling was measured by pulling out a frame from each participant's recording every 0.1 seconds. If the text moved between two frames, that transition was counted as a scroll. We found that students who had no preference between digital and paper reading performed worse when they scrolled frequently, but scrolling had no effect on the learning of other students. We hypothesize that learners who have a medium preference have developed that preference through exposure to both mediums, and are more likely to have strategies for scrolling.

Background

In the most recent synthesis of comparisons between digital and paper reading, Singer and Alexander (2017) conducted a survey of 15 studies occurring between 2001 and 2017 that compared digital and paper reading for texts of different lengths. Seven of these studies used a stimulus which was 500 words or more, and eight used a text which was under 500 words for their stimulus. Singer and Alexander found that the eight studies which used shorter texts either noted there was no difference in comprehension between digital and paper reading, or found that digital reading was better for comprehension. In contrast, six of the seven papers which used longer texts found that students comprehended better in the paper condition. The one exception exclusively studied second language learners and noted that the dictionary lookup functionality in the digital condition was popular. Singer and Alexander hypothesized that the reason digital readers consistently did worse than paper readers on comprehension tests for long passages but not short passages is that the most popular method of navigating long text on the computer screen, scrolling, negatively affects reading comprehension. This hypothesis is supported by a 1997 study that compared scrolling and paging as methods of interacting with digital text. The 1997 study found that the summaries written by undergraduates who paged through a digital document were more coherent than those assigned to navigate the same document with scrolling (Piolat et al., 1997). Part of these results may be explained by the poor user experience of scrolling in many applications in 1997. A study of 10 computer scientists in 1997 found that scrolling mechanisms were very slow and users reported being annoved that to

scroll they had to click on a small scroll bar and that the display did not refresh fast enough to show them their position while they were scrolling (O'Hara & Sellen, 1997).

More recent studies suggest there are some negative effects from scrolling, but they are not as strong as those Piolat et al. found in 1997. In 2007, a study of Swedish participants found that those who had to use a mouse pointer to scroll while reading reported more stress and mental fatigue than those who had to use the keyboard arrows to page through the document (Wästlund, 2007). However, there was no difference in how those who navigated by scrolling and those who navigated by paging performed on comprehension exams after reading. In 2009, a study of American undergraduates found that among students who were either less familiar with the subject matter of the stimuli or less familiar with reading webpages, comprehension test scores were higher if they clicked through digital pages than if they scrolled while reading (Sanchez & Wiley, 2009). For students familiar with the stimuli or familiar with reading webpages, there was no difference between navigating by scrolling and navigating by paging. A 2011 study of Polish adults found that participants who paged through paragraphs and those who scrolled through the document (Kłyszejko et al., 2011).

Together these studies suggest the impact of scrolling on reading processes and comprehension may be decreasing as the user experience improves and familiarity with scrolling increases. The question we explore in this study is whether scrolling in today's digital age would impact reading or whether perhaps familiarity with this process has reached a point where it no longer influences comprehension and memory. Additionally, each of the studies above compared the occurrence of scrolling to a different behavior (i.e., page turning or clicking through pages). This assumes all scrolling behaviors are similar and have a similar impact on reading. Because readers vary in how they scroll, our study examines the quantity and directionality of scrolling as it relates to comprehension.

One of the reasons scrolling may affect reading, is that semantic memory of text is linked to spatial memory. This was discovered by studies like Rothkopf's 1971 experiment, which found that substantive memory of sub-passages in a document was directly related to incidental information about the spatial location of the sub-passage (i.e. which page it was on; Rothkopf, 1971). More recent studies report similar findings, suggesting the arrival of digital age has not changed readers' reliance on spatial mapping of content. A study of French speakers in 1994 found that after reading a sentence on a computer, subjects could accurately point to where on the screen each word in the sentence had been displayed (Baccino, 1994). In 2007, Wegner and Inhoff used an eye tracker during a reading session (Weger & Inhoff, 2007). They asked readers to look at words they had just read and found that subjects' eyes immediately found the support of spatial mapping, the cognitive load on readers is increased. Since text in a scrolled document doesn't have a fixed position like text on a page, readers who scroll more may perform worse on comprehension tests because they are forced to remember content rather than map content. In contrast, readers who scroll fewer times, like those who only scroll when they must access additional text, may have a lighter cognitive load because they may still be able to map content onto space.

The question remains, though, as to whether readers' increased familiarity and quantity of digital reading experiences have changed how they spatially map content in digital text. For example, a digital reader may have developed compensatory strategies for dealing with the fluid nature of digital texts. Accustomed to scrolling, a reader may map content onto other content, which also moves, like an image or heading providing a constant location in relation to the other content in the text even as scrolling occurs. This may provide a similar decrease in cognitive load as mapping content on space within a physical page. Additionally, it may mitigate the negative impacts of scrolling found in earlier studies of digital reading.

All this assumes, though, increased familiarity with digital reading behaviors like scrolling, which have gained a lot of popularity since the late 90s. A study of 15 German university students in 2010 found that most participants preferred to read a small area of the screen and scroll the text they were reading into that area to reading the whole screen and then scrolling to the text (Buscher, 2010). When questioned about their web reading behavior, experienced users have reported that they use scrolling to move distracting ads off their screen while reading webpages (Hillesund, 2010). Perhaps the biggest proof that user interfaces for digital text are improving and gaining popularity is that more users are opting to read long forms of text on digital devices instead of printing them compared to readers from a decade ago (Freund, Kopak, & O'Brien, 2016).

In this paper, we wish to investigate scrolling in today's digital era. We examine variability of scrolling behaviors and whether scrolling impacts reading comprehension or is an equivalent form of navigation. We conjecture that if scrolling creates more of a cognitive load, then learners who scroll more will perform worse on comprehension and recall tests than learners who scroll the text a whole screen's worth only when they must.

There have been a few studies which counted the number of times a participant scrolled, but to the best of our knowledge, none have been conducted on the scale described here and none have tied the total amount of

scrolling to performance. For example, in 2001, six American academics were recorded searching through 561 documents to find which contained relevant information for a search query (Kelly & Belkin, 2001). Researchers recorded how often users scrolled in each document by recording the number of times they clicked on the scroll bar. They found that there was no difference between the number of times users scrolled when viewing relevant documents and the number of times they scrolled while viewing irrelevant documents. In 2016, Freund et al. recorded the screens of 41 Canadian participants and counted the number of times they scrolled to the bottom to check the length of the article they were reading (Freund et al., 2016). Readers checked the length of the document they were reading more often when the stimulus was a typical webpage with distracting advertisements and pictures than when it was stripped of extraneous content. This suggests scrolling behaviors differ, but does not link such behaviors to comprehension, which is the goal of the current study.

Method

We recorded the screens of middle schoolers while they read a website. We then used a computer program to calculate when they scrolled, and checked the results of the program against our own intuition by looking at visualizations of students at the far ends of the spectrum and summary statistics such as the average length of a student's scroll. Once scrolling was calculated for all participants, we linked it to comprehension using a statistical model which took other factors such as prior knowledge into account.

Experimental Setup

Our participants were 381 fifth to eighth graders (N=87 fifth graders, 78 sixth graders, 83 seventh graders, and 132 eighth graders) who were learning in the classrooms of 11 teachers across 3 schools in an urban district in the Southeastern United States. Eighth graders were oversampled because the stimulus was a passage designed for 8^{th} graders by the National Assessment of Educational Progress (NAEP). We received demographic data on 369 of the participants from the school board. Two hundred and nine (56%) of the students were female. 165 were White, 157 were Black, 34 were Hispanic, 12 were Asian and 1 was an American Indian. According to the 2010 census, our sample had about 12% more Blacks and less Whites than the town's distribution, but otherwise matched city-wide demographics. There were 87 students noted as economically disadvantaged and 66 spoke a language other than English at home.

All students did a mix of digital and paper reading as part of their regular classwork according to their teachers, though no classroom used one-to-one digital devices in class. All digital reading was recorded by iMotions (IMotions A/S, 2016). The recording contained an estimate of where the participant was looking aligned with the digital content being viewed, but we only used the screen recording for this study.

For the content, we divided an NAEP passage¹ into two parts with the dividing point chosen due to the natural end of a section. Therefore, the first part had 434 words and the second part had 674 words. Each participant read one part on paper and one part on a laptop with the order of the reading randomly assigned. Students were interviewed individually in a quiet classroom in their school. We measured student familiarity with the topic and preferences regarding digital and paper reading tools using a pre-test. After reading both sections, they took an open book post-test in which most of the questions referred to specific sentences or paragraphs in the stimuli.

When students started the digital reading, the researcher showed them how to use the highlighting, dictionary lookup and touch screen tools which were available in the digital environment. We did this because these tools change enough between devices and platforms that it was unlikely all students were already familiar with the tools we used, though their mechanisms were the same as similar tools found in PDF viewers.

Calculating Scrolling

We developed a computer program to analyze the screen recording video. Note that both passages were too long to fit on a single screen, and therefore participants were required to scroll down to see the full content. To determine number of scrolls, the program pulled out frames from the video every 0.1 seconds and used image processing to figure out where each line of text was relative to the top and bottom of the screen as well as which lines were present in a frame. This information was used to calculate how far down the page a participant had scrolled and whether they had zoomed into the screen. To calculate scrolling, we listed the frames in the order they appeared in the video. If the text moved more than half a line's height between one frame and the next, that counted as a scroll. We used a cutoff above zero because the image processing was occasionally off by a few pixels. This method worked for 378 of the 381 students we gathered data on. The screen capture program crashed for two of the students, and for one student the image processing algorithm was unable to identify the

¹ https://www.nationsreportcard.gov/reading_2011/testyourself_g8_passage_ann.aspx

text in the frames due to the student highlighting all the text on the screen. The code for this program can be found at <u>https://github.com/kbrady/eye-tracker-pipeline</u>.

In addition to the total number of scrolls, we also calculated how many scrolls went up versus down on the page. We theorized that scrolling up to possibly review the document would have a different relationship with comprehension than scrolling down. One hundred and forty-one students (37%) never scrolled up after the experimenter showed them how to scroll and 241 (64%) had less than five frame transitions in which they scrolled up. For perspective, the medium number of frame transitions in which a student scrolled down after the instructor gave them control of the computer was 14. This meant that for most participants, the number of timeframes they scrolled down and the total number in which they scrolled ended up either being the same or very close.

Since the laptop used to display the text was a touch screen, a lot of participants opted to zoom in to see the text or picture better (109 of the 378 students we got data for zoomed in at least once). If the zoom level on two consecutive frames was not the same, we did not count that pair as a scroll even if the text moved vertically. Due to the sensitivity in multiple directions of touch screen interface, it was very easy for students to accidentally scroll while zooming. For this reason, we felt it was best not to count text movement during a zoom as a scroll.



Figure 1. The median student measured by scrolling while reading the second section

Visualizing Scrolling

We checked that the scrolling measured by the computer program matched our own intuition of how much users scrolled while reading by visualizing each participant's scrolling. This allowed us to see the scroll positions detected by the computer, compare them to the screen recording, and quickly check whether the ranking of participants by number of scrolls made sense. An example visualization is shown in Figure 1. This image shows the median student measured by number of scrolls while reading the second part. The image on the left shows the vertical position of the screen plotted against the time since the session began. The three outtakes on the right-side show what the student's screen looked like during three sequences when it was stable. The top outtake shows what the student saw from 48 seconds after the session started to 2 minutes and 52 seconds. This outtake corresponds to the shaded area above the 2:47 tick in the graph on the left. The final outtake shows what was on the student's screen from 2 minutes and 53 seconds after the session started to 3 minutes and 48 seconds after the session started.

The x-axis in the graph is used for time. In general, students did not zoom in enough to cut off the horizontal sides of the screen, so they could see the full horizontal width of the text (as shown in the outtakes to the right in Figure 1). Thus, the shaded area in the graph only represents the vertical position of the screen. We overlaid the plot on a stitched together picture of the whole document so we could see which parts of the document were visible to the student at each time point during their reading session.

Figure 2 shows the same visualization applied to the students who scrolled the most and the least during the second part. The frequent movement of the blue band in 2.a shows that the student scrolled quite a

bit, while the stability of the blue band in 2.b shows that the student only scrolled in one short burst a little over 7 minutes into the session. Visualizing scrolling in this way also allowed us to see certain macro behaviors. For example, the thin blue downward spikes in the first minute of Figure 1 and Figure 2.a show that the screen was moved down and up very quickly shortly after the page loaded. This was the period in which the researcher showed the participant how to use the tools in the digital environment, including scrolling and zooming on the touch screen. Another macro behavior is reviewing the document, which is shown after 3:30 in Figure 2.a. We believe training the computer program to recognize and classify these behaviors would allow us to gain further insights into how students scroll while learning.



We also wanted to check whether students who scrolled more, scrolled shorter lengths than their peers. As shown in Figure 2, students on the far ends of the distribution seemed to support this theory, but we did not have time to look through the visualizations for all students. Therefore, we plotted the average length of a downward scroll for each condition against the number of times a student scrolled. As shown in Figure 3, there was a relationship between the number of times a reader scrolled and the average length of a scroll, with those who scrolled less often making longer scrolls.



<u>Figure 3</u>. Average length of a downward scroll in pixels plotted against the number of times a student scrolled while reading the first section

Statistical Models

We linked scrolling to post-test results (i.e., comprehension) using generalized linear mixed models. Each model was fit using the Laplace approximation for maximum likelihood. The model took pre-test scores, grade level, section read digitally, reading modality preferences and ethnicity into account as covariates. In addition to directly testing whether the number of times a subject scrolled had a relationship with comprehension, we considered the effects of holding these variables controlled. There were eight students who did not have a complete post-test and therefore had to be left out of the model. With the three students whose video recordings could not be processed, this limited the final model to 370 participants.

Results

Among students who had no preference between reading digitally versus on paper, we found a significant negative relationship between scrolling down and comprehension, controlling for the other covariates in the model. This group represented 39% of the subjects and was larger than either the set of students who said they preferred reading on paper (24%) or the set who said they preferred to read on digital devices (37%). The students without a preference between digital and paper were disproportionately likely to be non-white, but otherwise seemed demographically equivalent to the rest of the class. We did not find any other relationships between scrolling and comprehension.

There was no relationship between scrolling and pre-test scores, reading aptitude or ethnicity. The frequency with which students scrolled roughly corresponded to how much they liked digital environments. Students who preferred reading on digital devices scrolled more than their peers, followed by students who liked both digital and paper reading.

On average, we found that students who scrolled through the shorter section had 27.5 tenth of second timeframes in which they scrolled and students scrolling through the longer passage had 42.7. It seems that about 11 timeframes (1.1 seconds worth) per student are due to the researcher showing them how to scroll at the beginning of the session. If we exclude scrolls in the first minute, the average drops to 16.9 for the shorter section and 31.6 for the longer section. There was a wide range of scrolling behaviors in both conditions. The standard deviation was 19.6 timeframes for the short passage and 29.1 timeframes for the longer passage, with very long tails to the right in both cases.

As you might expect, the amount of time a student spent reading the digital text was positively correlated to how much they scrolled. This correlation was not so strong that we believe scrolling to be an approximation of time, nor did dividing the number of times a person scrolled by the time they spent reading produce different results in our analysis. On average students scrolled once every 14 seconds, but most scrolled more often than that (the median was a scroll every 10.5 seconds).

We also found a correlation between the number of times a student scrolled up and the number of times they scrolled down. This relationship is shown in Figure 4. Initially we thought this might be occurring because too many frames when the screen was still were being classified as scrolls due to bugs in the image processing algorithm. To check this hypothesis, we tried only counting scrolls in which the text moved at least a full line instead of half a line. Using this larger cutoff resulted in some students being labeled as scrolling much less. When we visualized those students, we found that there were many cases where we did not agree that the student had scrolled less than their peers. However, even with the higher cutoff there was a strong relationship between the number of times a student scrolled up and the number of times they scrolled down.



Figure 4. Number of upward scrolls plotted against number of downward scrolls

Discussion

Preferring text to be available in both paper and digital editions is not unusual among adults. A survey of New Zealand college students found that 49.7% preferred that their textbooks be available in both print and online (Traut & Toland, 2014). However, many of the reasons cited by adults for preferring materials in two formats such as reading the digital one on the train and the paper one at home and saving money by reading mostly online and printing critical passages do not apply to middle schoolers. We therefore conjecture that the two fifths of our subjects who did not have a preference between digital and print reading may have not had a chance to develop a preference due to not getting as much exposure to each environment as their peers.

If this conjecture is correct, then it would mean our results line up with Sanchez and Wiley's (2009) and support the hypothesis that negative effects from scrolling go away as students get more experience scrolling. It is notable that among this population, only down scrolling had a negative relationship with comprehension, not scrolling up. This supports the spatial hypothesis as an explanation of why scrolling is bad for some students. Frequently scrolling down allows the reader to only read the top line of the screen and thus erase spatial ques such as position on the screen which might have helped their memory. Scrolling up indicates that a learner took the opportunity to review or checked the length of the text, neither of which would impact their spatial memory.

One way of adding spatial cues while reading from a screen is to move text by a page at a time so that the spatial cues of the screen are re-introduced. Another method would be to train readers to use other spatial markers like text headings and pictures to orient the text they are currently reading in a space which is not dependent on the current placement of the screen. It is notable that learners who prefer to read from paper scroll less than their peers, suggesting they are employing the first strategy and learners who prefer to read digitally scroll more than their peers, suggesting they are using the latter strategy. Learners without a preference may still be using the screen position to remember what they have read but not adapting their scrolling behavior accordingly. Thus, in this group, those who scroll more are not comprehending the text as well. Training these learners to either scroll less often or pick up on spatial cues like headers and pictures might improve their comprehension.

Historians of printed books have noted that our modern method of reading printed books is only a few centuries old. Sixteenth century English bibles were littered with indices and concordance lists which encouraged discontinuous reading much like the links on Wikipedia sites today (Hillesund, 2010). Annotations and diaries from the era confirm that many readers did take a discontinuous approach to reading their Bibles. In the intervening centuries, long form printed text meant to be read from beginning to end has become popular and methods of reading it efficiently have been developed. As more educational tools move online, it becomes more necessary to train students in practices which will help their digital literacy, which may not be the same as the best practices for reading on paper. This may include introducing them to alternative navigation techniques or different ways of viewing space.

Future Work

We do not know why there is such a strong relationship between the frequency of downward scrolling and the frequency of upward scrolling. This suggests the metrics for scrolling in this paper are overly broad and a system which could recognize and categorize macro behaviors such as reviewing and checking length would get a better picture of how readers use navigation. Both the statistics we gathered and the visualizations of student scrolling showed in Figures 1 and 2 demonstrate that there is a lot of variance in how middle schoolers navigate scrollable text. In particular, we would like to know whether students who prefer reading digitally employ different navigation strategies than those who prefer paper or do not have a preference. The differences in scrolling amounts we observed between these groups suggest that they are employing different strategies.

References

Baccino, T. (1994). Spatial Coding and Discourse Models During Text Reading. *Language and Cognitive Process*, *9*(2), 143–155. http://doi.org/10.1080/01690969408402114

Buscher, G. (2010). Eye Tracking Analysis of Preferred Reading Regions on the Screen. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems* (pp. 3307–3312). Atlanta: ACM.

Freund, L., Kopak, R., & O'Brien, H. (2016). The effects of textual environment on reading comprehension: Implications for searching as learning. *Journal of Information Science*, 42(1), 79–93. http://doi.org/10.1177/0165551515614472

Hillesund, T. (2010). Digital reading spaces: How expert readers handle books, the Web and electronic paper.

First Monday, 15, 1-19. http://doi.org/10.5210/fm.v15i4.2762

IMotions A/S. (2016). iMotions Biometric Research Platform 6.3. Copenhagen, Denmark.

- Kelly, D., & Belkin, N. J. (2001). Reading time, scrolling and interaction: exploring implicit sources of user preferences for relevance feedback. *Proceedings of the 24th Annual International ACM SIGIR Conference* on Research and Development in Information Retrieval, (558), 408–409. http://doi.org/10.1145/383952.384045
- Kim, D.-H., & Huynh, H. (2008). Computer-based and paper-and-pencil administration mode effects on a statewide end-of-course English test. *Educational and Psychological Measurement*, 68(4), 554–570. http://doi.org/10.1177/0013164407310132
- Kłyszejko, Z., Soluch, T., Wieczorek, A., Chmiel, K., Sarzyńska, J., & Szóstek, A. (2011). The influence of text visualization on the screen on eye movements and information processing. In *Proceedings of the Conference: Interfejs użytkownika - Kansei w praktyce, Warszawa* (pp. 113–120).
- O'Hara, K., & Sellen, A. (1997). A Comparison of Reading Paper and On-Line Documents. In *Proceedings of CHI '97, Human Factors in Computing Systems* (pp. 335–342). Atlanta, Georgia, U.S.A. http://doi.org/10.1145/258549.258787
- Piolat, A., Roussey, J.-Y., & Thunin, O. (1997). Effects of screen presentation on text reading and revising. International Journal of Human-Computer Studies, 47(4), 565–589. http://doi.org/10.1006/ijhc.1997.0145
- Rothkopf, E. Z. (1971). Incidental memory for location of information in text. *Journal of Verbal Learning and Verbal Behavior*, *10*(6), 608–613. http://doi.org/10.1016/S0022-5371(71)80066-X
- Sanchez, C. a., & Wiley, J. (2009). To Scroll or Not to Scroll: Scrolling, Working Memory Capacity, and Comprehending Complex Texts. *Human Factors*, 51(5), 730–738. http://doi.org/10.1177/0018720809352788
- Singer, L. M., & Alexander, P. A. (2017). Reading on Paper and Digitally: What the Past Decades of Empirical Research Reveal. *Review of Educational Research*, 1–35. http://doi.org/10.3102/0034654317722961
- Traut, L., & Toland, J. (2014). Online vs. Printed Course Materials: Student Preferences and Influential factors. In *ICIS2014 SIGGreen Workshop*. Retrieved from https://siggreen.wikispaces.com/Accepted+Papers
- Wästlund, E. (2007). Experimental Studies of Human-Computer Interaction: Working memory and mental workload in complex cognition. http://doi.org/10.1016/j.chb.2004.02.007
- Weger, U. W., & Inhoff, A. W. (2007). Long-range regressions to previously read words are guided by spatial and verbal memory. *Memory & Cognition*, 35(6), 1293–1306.

Acknowledgments

This work was completed with the support of an interdisciplinary grant from Vanderbilt University (Amanda Goodwin is the primary investigator).