An Eye Tracking Study of How Pictures Influence Online Reading

David Beymer[1], Peter Z Orton[2], and Daniel M Russell[3]

[1] IBM Almaden Research Center, 650 Harry Road, San Jose, California 95120 USA
[2] IBM Center for Advanced Learning, 20 Old Post Road, Armonk, New York 10504 USA
[3] Google, 1600 Amphitheatre Parkway, Mountain View, California 94043 USA
{beymer, porton}@ us.ibm.com, drussell@google.com

Abstract. We present an eye tracking study to measure if and how including pictures – relevant or irrelevant to the text – affects online reading. In a between-subjects design, 82 subjects read a story on a computer screen. The text was accompanied by either: (a) pictures related to the text, (b) pictures unrelated to the text (advertisements), or (c) no pictures. Reading statistics such as reading speed and regressions were computed, as well as measures of picture gazes. When pictures related to the text were replaced with advertisements, we observed a number of significant differences, including speed, regressions, and re-reading.

Keywords: Eye tracking, viewing pictures, online reading.

1 Introduction

To understand the detailed structure of how people read text, psychologists and HCI researchers have turned to eye gaze tracking as a valuable analysis tool. In eye gaze tracking, a camera tracks and records where a subject's eye is looking; these gaze points are mapped to the text to follow the subjects' reading behavior. Eye tracking analysis has revealed how the eye moves during the reading process – see Rayner and Pollatsek [1] and Rayner [2] for excellent summaries. The eye reads a line of text in discrete chunks by making a series of fixations and saccades. A *fixation* is a brief moment, around 250 ms, where the eye is paused on a word or word group, and the brain processes the visual information. A *saccade* is a fast eye movement, usually forward in the text around 8-12 characters, to take in the next section of text. A *regression* is a backwards motion in the text, and it indicates confusion. The trace of these eye tracking parameters reveals much about the reader's cognitive state as well as the nature of the reading material. For instance, more difficult passages of text will yield longer fixations, shorter saccades, and a higher regression rate.

While eye tracking researchers have also studied how we process pictures (see Yarbus [3], Loftus and Mackworth [4]), surprisingly little work has been done on how we process the *combination* of text and pictures. Given our everyday exposure to rich combinations of text and pictures on the web, this gap is a little surprising. Carroll *et al* [5] studied how subjects view cartoons, looking at the processing of cartoon captions and graphics in *The Far Side* cartoons. They found that the text was read first and occupied most of the subjects' time. Similar results were reported by Rayner *et al* [6] for print advertisements and by Hegarty *et al* [7] for diagrams. The interplay between text and advertisements on web pages is being explored; Burke, *et al* [8] studied the negative effect of banner ads, showing that they slow down subjects in a web search task. In the Eyetrack III study [9], media researchers studied how subjects read online news sites. They found that (a) ads mixed in with the main text are viewed more than ads in the periphery, and (b) size matters for ads, with larger ones viewed more than smaller. The Norman Nielsen Group [10] recommends that pictures relate to content and don't look like ads.

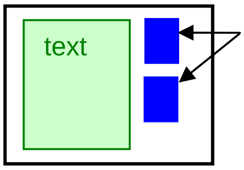
An important unaddressed problem is how pictures influence the *reading for comprehension* task. That is, given an online article with accompanying pictures, how do the pictures affect the reader when asked to read the article for comprehension? This issue is important for online instructional material as seen in e-learning systems.

In this paper, we present an eye tracking study of how different types of pictures affect reading a fixed passage of text. For a fixed, single-page story, we present three conditions to subjects: (a) *on task* – pictures relate to the story, (b) *advert* – pictures are advertisements, and (c) *none* – pictures are replaced by blank space. Will the differences in pictures alone cause detectable eye tracking differences in reading? For example, will advertisements slow the reader down? Understanding the influence of pictures on reading could help extend cognitive modeling in psychology or put HCI/usability rules-of-thumb about the use of pictures on more solid footing.

2 Experiment

In our experiment, we collected eye tracking data from 82 subjects in a between-subjects design; Table 1 shows the breakdown between conditions (A) – (C). The subjects were employees of a major computer company. We recruited subjects at two company cafeterias, offering them a cafeteria voucher in return for their participation.

Table 1. Page layout for the story and pictures. N is the number of subjects in each condition.



Picture Conditions		N
A) on task	pictures relate to text	27
B) advert	pictures are ads, same size as A	27
C) none	pictures replaced by blank space	28

The story presented to our subjects, taken from a science news web site written at an 8th-grade reading level, is on changes to the Earth caused by the 2004 Asian tsunami. The content was selected to go beyond common knowledge to allow for testing of retention. The story is 7 paragraphs long and contains 444 words total. In the *on task* condition, the two pictures include (1) an aerial shot of damaged coastline, and (2) a color-coded map showing depth changes in the ocean floor. In the *advert* condition, we selected ads for National Geographic and the New York Times as they were reasonably consistent with the science and news theme. Picture size is the same between conditions (A) and (B).

Our eye tracking setup includes the Tobii 1750 eye tracker, a camcorder taking a head-and-shoulders shot of the subjects, and three IBM T40 laptops. After sitting the subjects at a distance of around 60-70 cm from the Tobii and running a 5-point Tobii calibration, the experiment is presented in an instrumented browser. It includes: (1) instructions to read the story for comprehension, (2) a questionnaire asking for: name, first language, and a self-estimate of web usage, (3) the tsunami story itself, and (4) a 3-question, multiple-choice post-test of retention.

Data are recorded and analyzed by WebGazeAnalyzer (WGA) [11]. During reading analysis, WGA finds reading fixations by looking for a linear, horizontal grouping of fixations, calling the result a *gaze line*. Next, the analysis system uses a robust linematching algorithm to match gaze lines against lines of text from the story itself. From these matches, we can measure where and what the subject read, the reading speed, regressions, and additional statistics that we now report.

3 Results

Table 2 summarizes a number of eye tracking statistics, grouping them into measures of speed, distraction, and retention. For all but two rows, there are significant differences between the *on task* and *advert* conditions. First, consider reading speed. Using a speed metric of the 1^{st}-pass speed (equal to the 1^{st}-pass gaze duration / characters read) [2], [12], *on task* readers are 19% slower than *advert* readers, a significant difference ($F(1,52) = 10.23$, $p < 0.005$). Furthermore, this speed difference is consistent with similar significant differences in fixation duration and saccade length. For *on task* subjects, fixation durations are 7.6% longer ($F(1,52) = 5.35$, $p < 0.03$) and saccade length is 15% shorter ($F(1,52) = 7.51$, $p < 0.01$) compared to *advert* subjects.

Table 2. Reading statistics reveal that on task pictures slow the reader's speed, and advertisements increase the regression rate (shown as reg. rate). All but two rows have significant differences between the *on task* and *advert* conditions; the p-value is given in the right column. Standard deviations are shown in parentheses.

Reading Statistic	Pictures			Significance level, *p*
	On task	Advert	None	
Speed: 1^{st} pass speed (char/sec)	40.2 (7.8)	49.9 (13)	45.9 (13)	$p < 0.005$
Fixation duration (ms)	269 (35)	250 (26)	260 (42)	$p < 0.03$
Saccade length (char)	10.1 (1.8)	11.9 (2.9)	11.1 (2.7)	$p < 0.01$
Distractions: Reg. rate (reg/sec)	0.43 (.17)	0.54 (.19)	0.53 (.22)	$p < 0.03$
Picture gaze duration (sec)	1.26 (1.1)	0.86 (1.1)	n/a	*not significant*
Number of picture gazes	1.69 (1.3)	1.05 (.83)	n/a	$p < 0.05$
Re-read on picture return (char)	6.25 (14)	23.8 (26)	n/a	$p < 0.03$
Retention (% correct)	80.2 (19)	80.2 (24)	80.9 (21)	*not significant*

The second group of rows in Table 2 compares conditions based on measures of distraction. While the *on task* subjects spend more time viewing pictures than *advert* subjects, evidence from regressions and re-reading suggest that pictures are a distraction for *advert* subjects. Increased picture viewing for *on task* subjects comes as no surprise, because pictures do relate to story content. To measure picture viewing, we look at (a) picture gaze duration, and (b) the number of distinct picture gazes. First, for picture gaze duration, *on task* subjects spend 44% more time fixating on the pictures compared to *advert* subjects. Due to high variance in the time data, however, this difference is not significant (picture viewing is quite unstructured and variable among subjects). Second, subjects in the *on task* group have 60% more distinct picture gazes compared to the *advert* group, a significant difference ($F(1,52) = 4.471$, $p < 0.05$).

Despite *on task* subjects' increased attendance to the pictures, pictures seem to hurt *advert* subjects through increased regressions and re-reading. Subjects in the *advert* condition have a significantly higher rate of regressions than *on task* subjects ($F(1,52) = 5.014$, $p < 0.03$), which we attribute to the distracting nature of ads. Finally, to measure re-reading caused by a picture gaze, we note the text exit and re-entry points. If the re-entry point is located before the exit point, then the picture gaze is causing re-reading. As shown in the Table 2 row "re-read on picture return," the ads caused significantly more re-reading than the on task pictures ($F(1,29) = 5.34$, $p < 0.03$).

Finally, there are no significant differences between conditions for retention of the material as measured in the 3-question multiple-choice post-test. While the ads may have caused regressions and re-reading of the material, they did not impair comprehension.

4 Discussion

This paper is the first eye tracking study of the effect of different types of pictures (on task, advertisements, and no pictures) on the task of reading for comprehension. One finding is that on task pictures slow readers down, decreasing 1st-pass reading speed, lengthening fixation duration, and making saccades shorter. We attribute this to the extra effort the reader is making to relate the pictures to the text – the cognitive effort to relate pictures and text is slowing down the reader. On the other hand, advertisements appear to be distracting the reader by causing more regressions and re-reading of the material. No impact was found, however, from pictures on the retention of the material.

This study is important for design issues in e-learning. The negative impact of ads on the regressions and re-reading would argue against e-learning sites that are funded by ads placed on the same page as the instructional material. It would be better to charge users (or their organizations) up front and keep the site ad-free.

Returning to the issue of distraction from advertisements, it is interesting to note the rise of contextual advertising and its potential to increase the attraction of ads. In contextual ads, the web page content is scanned to determine those ads that may interest the reader, and only those ads are presented (Google AdSense, Chitika [13]). Thus, the ad will target the web page's intended audience, potentially creating a distraction that is hard to resist. The effect of contextual ads on a subject's performance and task completion would make an interesting future eye tracking study.

References

1. Rayner, K. and Pollatsek, A. The Psychology of Reading. Lawrence Erlbaum Associates, (1989).
2. Rayner, K. Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, 124:3 (1998), 372-422.
3. Yarbus, A. Eye Movements and Vision. New York: Plenum Press (1967).
4. Loftus, G. and Mackworth, N. Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 4 (1978), 565-572.
5. Carroll, P., Young, J. and Guertin, M. Visual Analysis of Cartoons: A View from the Far Side. In *Eye Movements and Visual Cognition*, K. Rayner, ed. (1992), 444-461.
6. Rayner, K., Rotello, C., Stewart, A., Keir, J., and Duffy, S. Integrating Text and Pictorial Information: Eye Movements When Looking at Print Advertisements. *Journal of Experimental Psychology: Applied*. 7:3 (2001), 219-226.
7. Hegarty, M. The Mechanics of Comprehension and the Comprehension of Mechanics. In *Eye Movements and Visual Cognition*, K. Rayner, ed. (1992), 428-443.
8. Burke, M., Hornof, A., Nilsen, E., Gorman, N. High-Cost Banner Blindness: Ads Increase Perceived Workload, Hinder Visual Search, and Are Forgotten. *ACM Trans. On Computer-Human Interaction*, 12:4 (2005), 423-445.
9. Eyetrack III: Online News Consumer Behavior in the Age of Multimedia. The Poynter Institute, http://www.poynterextra.org/eyetrack2005/main.htm
10. Norman Nielsen Group, http://www.nngroup.com
11. Beymer, D. and Russell, D. WebGazeAnalyzer: A System for Capturing and Analyzing Web Reading Behavior Using Eye Gaze. In *CHI 2005 Extended Abstracts*, ACM Press (2005), 1913-1916.
12. Liversedge, S., Paterson, K. and Pickering, M. Eye Movements and Measures of Reading Time. In *Eye Guidance in Reading and Scene Perception*. G. Underwood, ed. (1998) 55-75.
13. Tedeschi, B. Internet Banner Ads Look to Get More Interesting (and Thus Less Easy to Ignore). *The New York Times*, July 11, 2005.