Text as Visual Distraction

Sam Seabrook
Computer Science Department
Clemson University
Clemson, South Carolina 29634
smseabr@clemson.edu

Chris Georgiades
Computer Science Department
Clemson University
Clemson, South Carolina 29634
chrisg@clemson.edu

ABSTRACT

In this paper we describe the effects of textual distractions on a person's memory. The implications of this work are varying from questions like should an employee be allowed to listen to music while working; to should it be illegal to do something like text and drive. Similar things to our experiment are the dual-task paradigm and subtitling which are talked about in the background section. We went about proving our hypothesis by setting up an experiment in which we are showing an image with or without a text distraction and asking questions based on one's memory of said image. To further show the level of distraction we tracked the eye movement of the user to show exactly where and how long the user was looking. We hope to show that when given a second task there will be a significant difference in correct answers about the image do to the textual distraction.

Author Keywords

Distraction, text, visual search, eye-tracking.

INTRODUCTION

Goals

This experiment has a statistical goals and an analytical goal which are both of equal importance in our eyes. The statistical goal is to gather data about whether or not the subject answered the questions following the experiment correctly. Another statistical goal is to collect the eye tracking data including how long the subject was looking at the distraction and how often. These are both very important as they provide some real, measured data on whether or not a simple distraction in these controlled cases actually affected the subjects in their goal.

Our analytical goal is to take this data we collect and be able to surmise a firm and definite conclusion on whether in this case text specifically beyond a reasonable doubt does cause us to perform poorly at remedial to complex tasks.

Motivation

Have you ever heard someone say something along the lines of "I can listen to you and read this at the same time", or "Yes I'm listening, just because I'm watching television doesn't mean I'm not". We have also heard these things and wanted to design an experiment to get some hard data on this subject. Although initially it just sounds like this experiment is to prove a comical question there are some very serious implications of statements like these. There are some people that have lost their jobs because they do other things at work and some people have even lost their lives while texting and driving. These things are what led us to want to do this experiment and get some real data on whether or not, and how much, distractions affect our ability to do simple or complex tasks.

Hypothesis

It is hypothesized that the presence of text along the bottom of the picture grid will cause a significant drop in performance during testing.

BACKGROUND

The implications of subtitling on attention have been well researched in recent years. In the paper by Mcnamara, et. al. [1] they used a technique to improve searching in a distracted scene. The scene was distracted by common objects. What they did was have search bubble in the scene and they were trying to get the user to go straight from bubble to bubble without looking at other objects. This is concurrent with our hypothesis because it shows that searching when guided was better than a subject just searching an image in general. There are a lot of limiting factors to how viable this research is with ours as it wasn't measuring data on the distraction itself, but showing that this technique could indeed improve the search of the subject. It is still much related in many ways and a good read if one wanted another reference to distractions and effect. Our experiment also parallels the Kallenbach, et. al. paper [2]. The main thing to get from this paper in relation to ours was it showed that when there was short text and a big image the users hardly looked at the text. This is in line with our hypothesis in that the user unconsciously picked one thing to do and was not able to equally multi-task. Another couple of papers very similar to the one previously mentioned are *Ogata, et. al.* [3] and Qvarfordt, et. al. [6] Although it is not directly related to our paper *Burke, et. al. [4]* is relevant because it shows that distractions effected what a user looked at. In this experiment the flashing banners received the most attention. So it shows that a distraction can affect how a person performs at a task. This is very important to our research in that our goal is to

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.

significantly suggest that text serves as a distraction from the additional details on screen. The *Maglio, et. al.* [5] paper discusses designing peripheral text that least affects a user's ability to do his or her main task. Just the mere fact that different text in the periphery affects the user differently is pertinent to our research. If you are interested in our kind of research this is a must read paper. One big thing that is tied to our research is the psychology dual task paradigm. This is the closest related thing to our research. The dual task paradigm is a psychology problem where one gives someone two tasks to complete to see the results verse a single task. In the Gordon et. al. paper [7] they take a look the effect of different tasks on each other and how people performed them together vs. by themselves. This is very similar to our research except we focus specifically on text as a distraction and not just as another task. Although they are not the same we can learn much from this paper about how to get better data and run the experiment better. We also looked into subtitling. Specifically at the Bryant et al paper[8]. Their experiment was to see whether subtitles were distracting or helpful. From this paper we concluded it would be viable for us to do further research alongside subtitling to try and find a way for the text to distract a user the least. So in conclusion although we cannot find exactly what our experiment is doing there are a lot of papers that are directly correlated with ours and we recommend taking a look especially at all the aforementioned papers if you are interested. We will also look into future research alongside subtitling after the results and data are calculated from this experiment.

METHOD

Apparatus

The primary equipment used for the diagnostic eye-tracking research was the Tobii model ET-1750 display. The Tobii ET-1750 has a resolution of 1280x1024 on a 17 inch TFT display with two embedded infrared light emitting diodes that allow the software to track eye movements of the user at the monitor. The sampling rate is 50Hz (binocular) with a latency of 20ms and an accuracy of 0.5° (bias error). The Tobii eye tracking server runs on the Windows platform with Tobii's Studio software (v2.2.8) and is connected via an IOGEAR (GCS102U) 2-Port USB KVM switch.

Design and Stimuli

The design consisted of a single factor within-subjects study with one independent variable and one dependent variable. The two levels of the independent variable manipulated were the presence or omission of a textual prompt along the bottom of the stimulus while the participant surveyed the image. The dependent variable is the number of correct answers to a series of questions following each view of a stimulus.

The presentation order of the two stimuli was counterbalanced throughout the trials for each participant. Figure 1 shows the stimulus used for the text distraction portion of the experiment. Despite the absence of text in the second stimulus, both stimuli have the same areas of interest (AOI's) so the gaze times measured cover the same area on the image as shown in Figure 2.

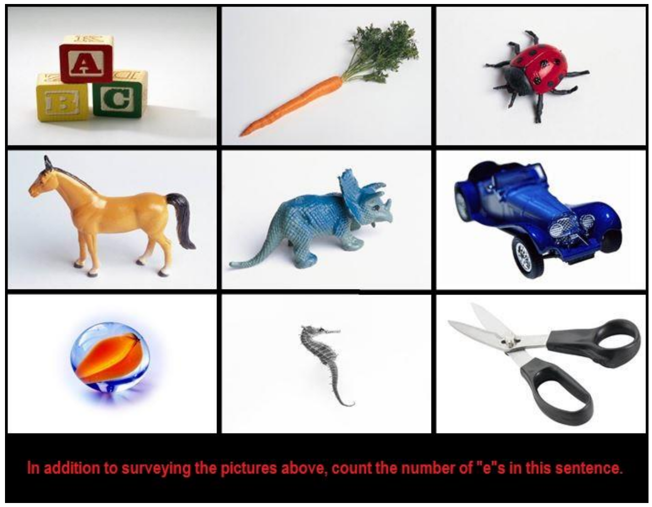


Figure 1: Stimulus Image with Text Distractor

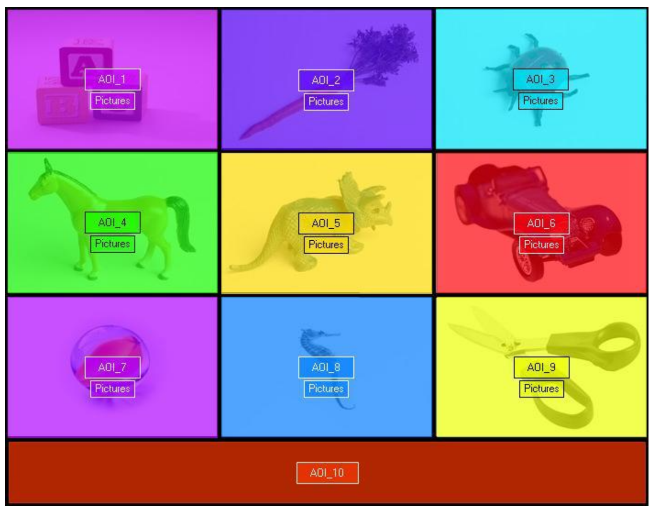


Figure 2: Stimulus with AOI overlay

Subjects

The experiment consisted of 25 male subjects with normal color vision between the ages of 18 and 24. The data from two of the subjects were thrown out because of faulty results due to head and body movement during the test. The experiment took each subject approximately ten minutes.

Procedure

Each subject will be seated directly in front of the Tobii monitor (Figure 3) and asked to read an Informed Consent Form describing their rights as a participant, as well as the experiment objectives, potential benefits, and risks. The investigator then asked the subject if he/she has any questions.

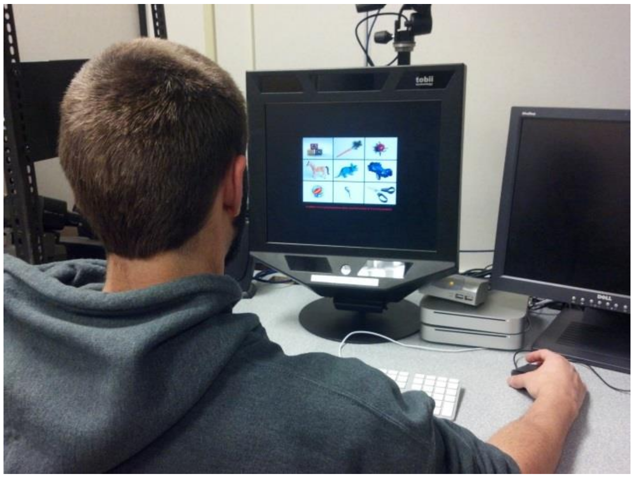


Figure 3: Subject participating in the experiment

Before the calibration could begin, each participant was required to pull the chair up to the desk, sit straight without slouching or leaning, and center themselves on the monitor. The Tobii software gave feedback in the form of two circles representing the subject's eyes. Once the subject's eyes were centered in the feedback window, the subject was instructed to follow a red ball around the next screen using only his/her eyes, without attempting to predict its path or adjusting themselves in the seat. Once calibration was complete, the investigator ensured that each point of interest on various locations of the screen was adequately measured for the test. If any points had too large of an error, the specific point(s) could be recalibrated.

Next, the test was executed and a short instruction screen prompted the subject to carefully survey the following image and remember as much detail as possible in order to answer questions later regarding what they recalled. Once the participant read and understood the instructions, he/she was prompted to left-click anywhere on the screen to show the first stimulus. Depending on how each experiment was counterbalanced, the subject would have 25 seconds to view the stimulus either displaying the group of items with or without a textual prompt along the bottom of the screen. If the stimulus shown to the subject contained text, the following screen prompted each participant to answer the five questions from Table 1. However, if there was no text below the image in the stimulus (blank) then each participant would answer the questions from Table 2 below.

1. How many "e"s did the bottom sentence contain?
2. In which corner were the scissors located?
3. Was the insect pictured a grasshopper?
4. What color was the car on the right side of the screen?
5. Which direction did the horse face?

Table 1: Post-Text Questions

After the subject answered the first five questions, an intermediate screen of instruction was displayed for ten seconds that stated a different stimulus would be displayed next. The subject should again focus on details and later recall as much about the image as possible. Depending on the counterbalanced sets, the subject then answered the remaining five questions that had not been previously shown in the questionnaire.

1. How many pictures were on the picture grid?
2. Was there a dinosaur picture?
3. What was the round object?
4. What vegetable was displayed?
5. How many letter blocks were there?

Table 2: Post-Blank Questions

Upon the completion of the last five questions, the subjects were then informed of the correct answers to each question and thanked for their participation in the experiment.

RESULTS

Eye movement data collected during the recorded experiments was exported from Tobii Studio and examined for statistical significance. Various metrics were used to gather additional information from the study including time to first fixation, mean and sum of fixation durations, and number of fixations per AOI. The metric used to measure each subject's performance was the number of questions he answered correctly out of five for each stimulus. This information was analyzed using a paired T test as shown in Table 3.

	Mean	Std. Dev.	T Value
Blank Stimulus	4.8261	0.3876	-
Text Stimulus	3.5652	0.7878	-
Paired	1.2609	0.1831	6.8877

Table 3: Paired T Test for Performance

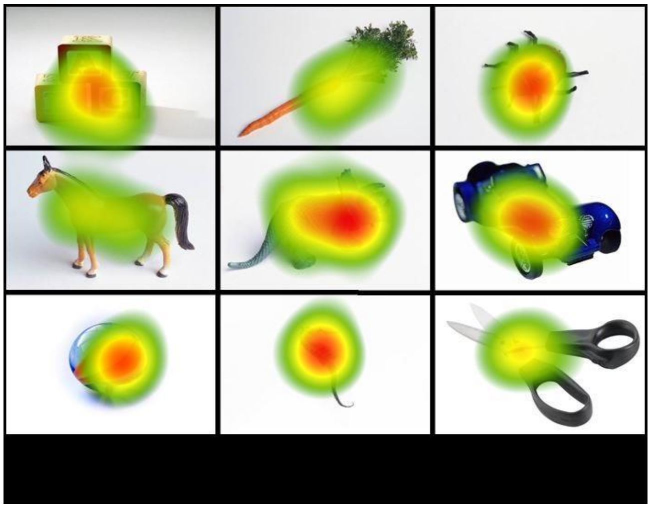


Figure 4: Media without text

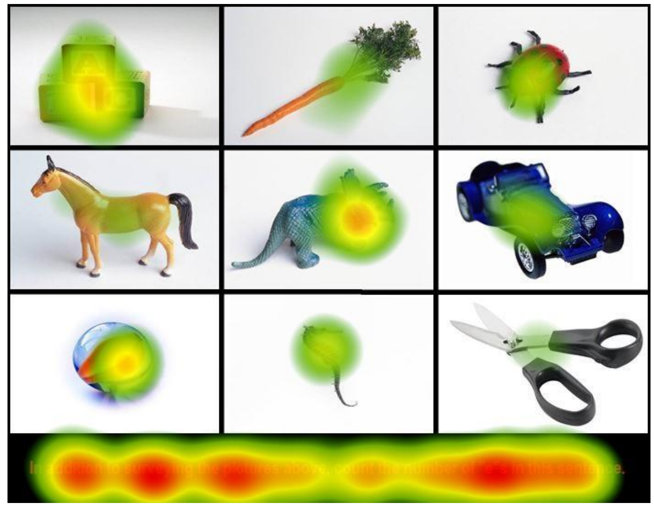
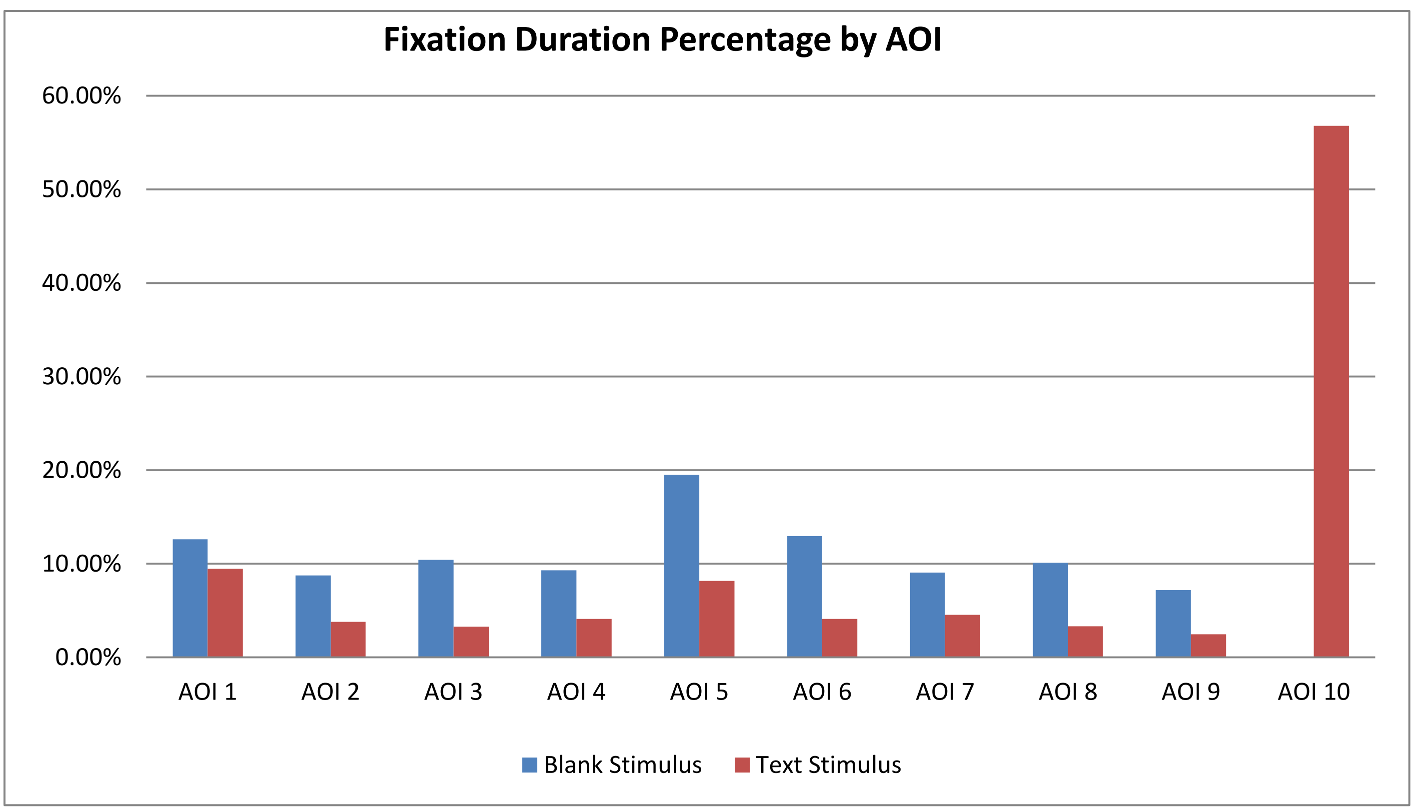


Figure 5: Media containing text

The results of this analysis provided quantitative information to determine the probability of the team's hypothesis. The figures calculated, T=6.8877 and P<.0001, suggest performance on the text stimulus is significantly worse than performance on the blank stimulus. The heatmaps in Figures 4 and 5 show the areas that were most fixated upon in the blank and text stimuli respectively.

There was no significant difference between the stimuli for time to first fixation (p=.0719). Despite different AOI's being fixated upon depending on the presence of text, the test shows there was no single AOI that significantly grabbed the subjects' attention.

Using the visit count metric, which measures saccadic crossovers, the study suggested the presence of text had a significant impact on the average visit count to each AOI ($p<0.05$). AOI's in the non-text stimulus consistently had more visits recorded than the stimulus with text, meaning subjects could survey each image multiple times when there was nothing to distract them. Similarly, Figure 6 below shows the percentage of fixations recorded in each AOI when comparing the two pictures. The area containing text accounted for over 55% of the average fixation duration for the time spent observing the text stimulus.



DISCUSSION / CONCLUSION

Motivated by current hot topics such as texting while driving and the ever growing use of technology in the world today, we explored through a simple yet extensive experiment, whether text is enough of a visual distraction to impede one's ability to complete a task. In contrast to previous works we focused on specifically text as a distraction to try and conclude whether text alone was enough of a distraction to inhibit the user's ability to complete a task. The first and easiest thing to conclude from the results was that there was a significant drop-off in correct answers to the post experiment questions when there was text and when there wasn't. When text wasn't present the users averaged 96.5 percent correct answers. When the text was present users answered 71.3 percent questions correctly. There are many things that factor into why it would drop off of that so we turned to the eye tracking data for more conclusive results that it was indeed the text that caused this drop-off. There are a few things that were used to determine our final conclusion. Did the user look at the text longer than the image that was bigger and more in his field of vision? What percentage of visual fixation time does one have to look at the image to successfully answer the questions? Is it conclusive that the text as a distraction inhibited the user's ability to answer the questions? From the data collect from the Tobi eye tracking system one can conclude that people fixate on the text longer than the image which is shown in the results graphs. Also, unexpectedly we found that when text was present, the fixations on the image were centered on one or two areas, and while text was not present the whole image was fixated upon. This leads us to conclude that even while users were looking at the image, when text was present, they weren't as focused on it as they were when text was not present. There was one question when text was present that was answered wrong almost every time. We looked at the ones that got it right and found that they had actually just picked one thing to focus on and didn't really even look at the image. So this supports our hypothesis that there is a significant difference in ability to complete a task when there is no distraction. So in conclusion we find that our original hypothesis was correct. This is because we found that users answered 25 percent more questions correct when there was no text present. We can conclude that it was due to the text as a distraction because the data collected shows that people fixated longer on the text than the image even if they were a subject that had the image first and knew we were asking questions about it. Also we found that people didn't focus as well on the image even when they were looking at it when text was present. This was found using the heat map images, showing that only certain areas were fixated upon with text where as all of the image was fixated upon about equally when text was not present. There is much future work that can be done from this experiment to test other variables and try the experiment from different angles such as have the image as the distraction to see if people still fixate on the text. From our data and experiment we conclude that the text draws a person's eyes and attention away from their task significantly enough to say that our hypothesis was correct.

ACKNOWLEDGMENTS

We would like to thank Dr. Duchowski for his guidance in creating our experiment and the Clemson University School of Computing for allowing us to utilize its eye-tracking lab to conduct this research.

REFERENCES

1. Ann McNamara, Reynold Bailey, and Cindy Grimm. 2009. Search task performance using subtle gaze direction with the presence of distractions. *ACM Transactions on Applied Perception*. 6, 3, Article 17 (September 2009), 19 pages.
2. Jan Kallenbach, Silja Narhi, and Pirkko Oittinen. 2007. Effects of extra information on TV viewers' visual attention, message processing ability, and cognitive workload. *Computers in Entertainment.* 5, 2.
3. Keiji Ogata, Yasuhiro Seya, Katsumi Watanabe, and Tohru Ifukube. 2012. Effects of visual cues on the complicated search task. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design* (NordiCHI '12). ACM, New York, NY, USA, 478-485.
4. Moira Burke, Anthony Hornof, Erik Nilsen, and Nicholas Gorman. 2005. High-cost banner blindness: Ads increase perceived workload, hinder visual search, and are forgotten. *ACM Transactions on Computer-Human Interaction.* 12, 4 (December 2005), 423-445.
5. Paul P. Maglio and Christopher S. Campbell. 2000. Tradeoffs in displaying peripheral information. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (CHI '00). ACM, New York, NY, USA, 241-248.
6. Pernilla Qvarfordt, Jacob T. Biehl, Gene Golovchinsky, and Tony Dunningan. 2010. Understanding the benefits of gaze enhanced visual search. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (ETRA '10). ACM, New York, NY, USA, 283-290.
7. Logan, Gordon D.; Burkell, Jacqueline Dependence and independence in responding to double stimulation: A comparison of stop, change, and dual-task paradigms. Journal of Experimental Psychology: Human Perception and Performance, Vol 12(4), Nov 1986, 549-563.
8. Ross Bryant, Jjill Constante, Mike Manna, Tony Serechia, Tawanna Starks.2004. Visual Reaction to Subtitling in Television and Media(7 pages)