Quantitative Analysis of Font Type's Effect on Reading Comprehension

Jaret Screws
Clemson University
Clemson, United States
jscrews@g.clemson.edu

December 1, 2016

ABSTRACT

This study investigates the effect that Serifs have upon the readability of text. The readability of a text is determined by words per minute, number of regressions, saccade length, and fixation time. Each participant was given a piece of text to read, the only difference being that one will have a Serif font or a Sans font. Afterwards, the main points of evaluation that was extracted are words per minute (WPM), saccade length, fixation, and regression percentage.

Keywords: reading, serif, sans-serif, comprehension, regression, saccade length, words per minute (WPM)

INTRODUCTION

With a majority of information is passed through the use of text based sources like newspapers, articles, websites, and forums, there is a need to examine the effects that the text size has upon the reader's comprehension along with the amount of time it takes the reader to finish the text. The goal of this study is to view the effects that text font has upon a reader's WPM, number of regressions that the reader takes, average length of the saccades, and the average fixation time.

When it comes to font families, there are two main families of fonts: Sans Serif and Serif. Serif fonts are fonts that have small lines called serifs on the edges of the letters; an example of this is show on figure 1. In this study, the two fonts that was used are Sans-serif and Serif font types. Studies, like those performed by Morrison and Noyes, have shown that in some cases Serif fonts perform better than their ornate counter parts [1]. They attribute this to the possibility that the serifs on the font add a more distinct end to each letter, allowing the reader to quickly discern the end of each letter and allow for increased fluidity while reading. To achieve the most efficient reading pattern, longer saccades and minimized fixations are desired. The increased definition that the serifs add to text will most likely result in a higher readability based upon the WPM, number of regressions, saccade length, and fixation time.

Figure 1: Comparison of Sans Serif and Serif fonts. Displaying the serifs with indicators.

BACKGROUND

Through the decades, many have investigated how font effects reading comprehension. As early as the 1930's, it was shown that two fonts out of a tested ten performed slower [2]. It should be noted that neither of these fonts are still in use today. There are many factors that could cause fonts to perform worse than others. For instance if a particular font type causes a higher percentage of regression than the average 10% to 15% [3], then it will have a higher likelihood of producing a slower overall readability. There are two types of regression that a reader can have: a short with-in word regression, or a longer regression of ten or more letter spaces [3]. The with-in word regressions suggest that the reader is struggling to understand the current word, while the longer regressions suggest that the reader does not understand the text. Another factor that could impact

the WPM of a font could be if it causes more or longer fixations than a more readable font. The fixations can range from under 100 ms to over 500 ms [3]. With this high variance for fixation time, if one font has a significantly lower fixation time, it would suggest that one font may be more readable than another. The final factor that helps to determine if a font type is superior to another is the length of the saccades that occur when reading the text. The length of the saccade can be refereed to as the perceptual span. The purpose of the perceptual span is to bring new information into view while the eye is still fixated upon its current target. The perceptual span has been show to be asymmetric, "extending from about 4 character to the left from the fixation point to about 15 character to the right" [4].

Though the study by Morrison and Noyes showed a significant difference between a Serif and a Sans Serif font, other studies have not been able to show a significant difference between the readability between a Serif and a Sans Serif font [5]. This study used two separate fonts on the same piece of text and also varied the font size of texts that was tested, however there have been conflicting results as to whether or not serifs have an impact on the read ability of a font.

METHODS

Apparatus

Figure 2: An example set up of the used eye tracker.

The eye tracker that shall be used for the experiment is the Gazepoint GP3 Eye Tracker. An example of the eye tracker is show in figure 2. The Gazepoint is a pupil/corneal reflection tracking device. The Gazepoint offers two types of calibration, either a 5-point or a 9-point calibration. The Gazepoint has a 60 Hz refresh rate that results in data being collected roughly every 16ms. The Gazepoint has a degree of accuracy of half a degree. The Gazepoint was mounted to a monitor to conduct the experiment. The monitor that was used for this experiment is a Dell P2213 22" monitor. The P2213 runs at a resolution of 1680 x 1050 and has a refresh rate of 60 Hz.

Stimulus

The text for the stimuli was taken from the first few sentences on the Wikipedia page for eye tracking. The text was split and applied each to a white background that measures 1680x1050 pixels. Two images were created by altering the text to Sans font while the other two were created by altering the text to Serif font. To attempt to keep the font relatively the same size, it was necessary for the two stimuli to be different font sizes. The Sans stimuli is set to 108pt font while the Serif was set to 128pt font, which can be viewed in figure 3 and figure 6 respectively.

Subjects

The subjects for this study were 10 Clemson University students between the ages of 18 to 24 were tested. Of the groups tested 5 were male and 5 were female. The subjects were randomly split into two groups, group A and group B. Group A was shown a Sans stimuli first and then a Serif stimuli. While, group B was shown a Serif stimuli first and then a Sans sitmuli.

Experimental design

For the experiment, 2 factor (font type) at 2 levels (text porition) was utilized, or 2x2. This will result in four pieces of stimuli. On each of the stimuli, an independent area of interest (AOI) was place over each individual word. These AOI's was used to determined the distance of regressions and how many words the saccades covers. The AOI's will also allow it to be determined which words cause for longer fixations and will analyze the length of the fixations between the font types. So that subjects will not gain any benefit from seeing one piece of text and then another, each subject will only be shown one piece of text. This means that the experiment and the analysis is being done within subjects.

Procedures

Participants were seated and greeted based upon the script. The participants were given a brief survey to determine their age, sex, and if they had any vision problems. Each participant went through the Gazepoint calibration to ensure the validity of the test. The participants were given instructions to read the text for however long it took them to fully understand it. This portion of the experiment was timed, however there was no time limit placed upon the individual participant.

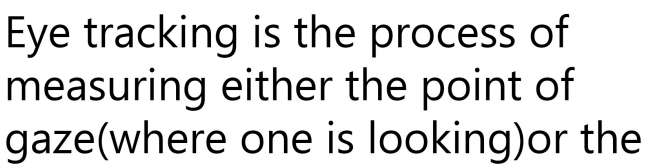

Figure 3: The first Sans stimuli.

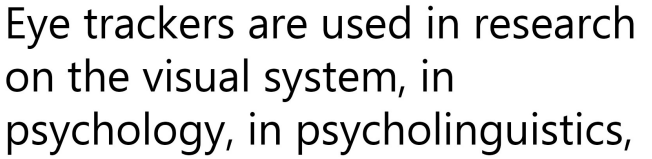

Figure 4: The second Sans stimuli.

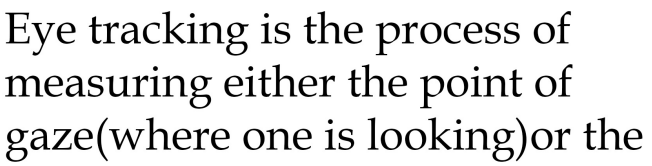

Figure 5: The first Serif stimuli.

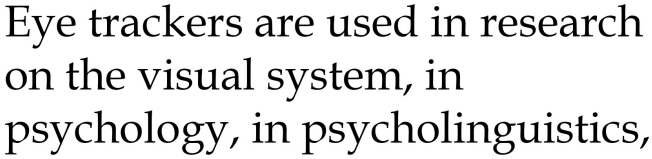

Figure 6: The second Serif stimuli.

This was to insure that there was no pressure placed upon the participants so they would not attempt to rush to finish the text without fully reading it. Once the subject completed the reading they were given a one question survey to determine that they had actually read the text. After they had completed the test, the participants were thanked according to script and sent along their way.

RESULTS

The experiment resulted in 10 trials being done with each person reading 2 pieces of stimuli, each giving a .csv file produce by Gazepoint. Each .csv file was analyzed using a combination of python and R scripts to find the scan path, the fixations, the saccade amplitude, the number of words between fixations, the number of re-fixations, and the WPM. The python scripts were used to generate the quantitative data from the .csv files that would then be given to the R scripts for statistical analysis using ezANOVA.

Scan Paths

The python scripts that generated the scans paths for the experiment used a velocity based formula to determine fixations. The python scripts produced a visualized form of the scan path highlighting each AOI that was fixated upon in red and all that aren't fixated upon in green. The circles on the image are the fixations themselves and the larger the circle is the longer that the fixation lasted. As expected the type of font does not appear to have any affect of the scan path that is taken. The scan path almost always starts in the top left moving right and dropping one row at a time as it approaches the bottom of the text as show in figure 7, in figure 8, in figure 9, and in figure 10. It should be noted that each of these scan paths are from a different trial but each follow the same general path mainly focusing context words and have fairly short jumps between words normally only one to two words. It does not matter what the text is or what the font is it still follows the expected path. It can also be seen that in some cases the participant would go back to re-read words or they would fixate on the same word multiple times(re-fixations).

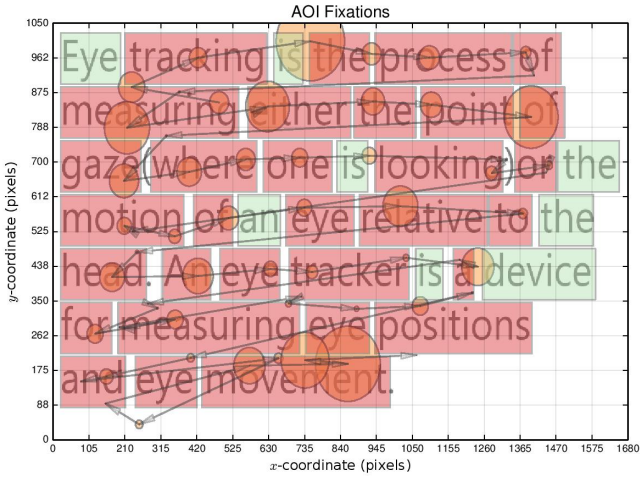

Figure 7: An example scan path on the first serif stimuli.

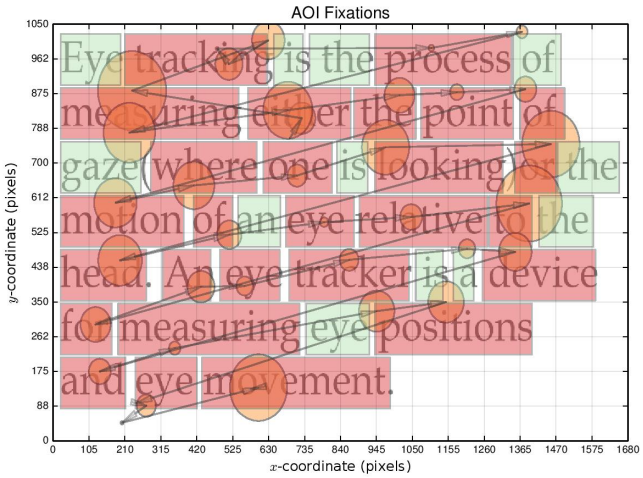

Figure 8: An example scan path on the first sans stimuli.

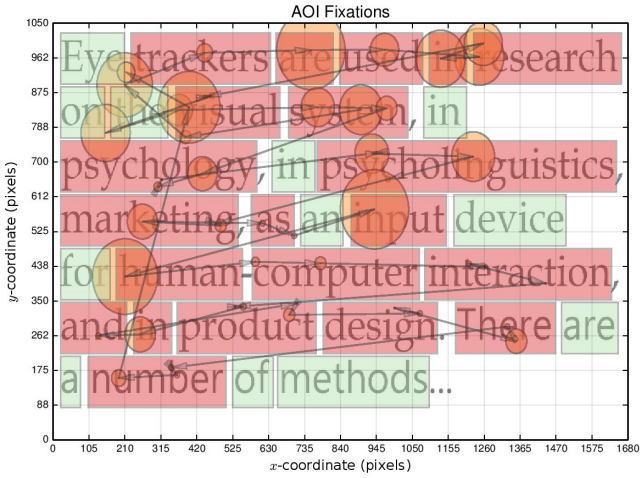

Figure 9: An example scan path on the second serif stimuli.

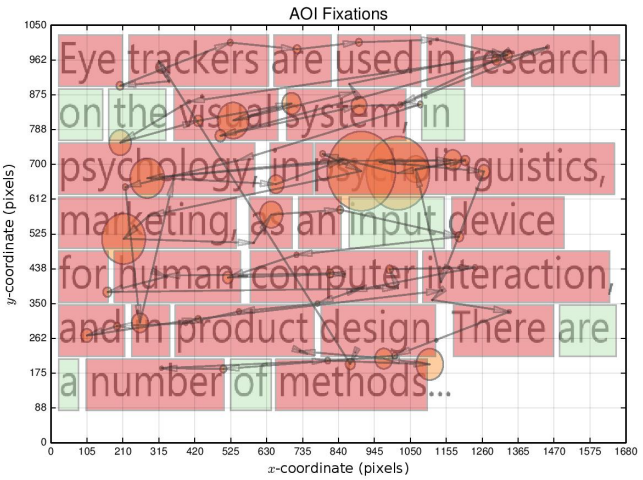

Figure 10: An example scan path on the second sans stimuli.

Fixations

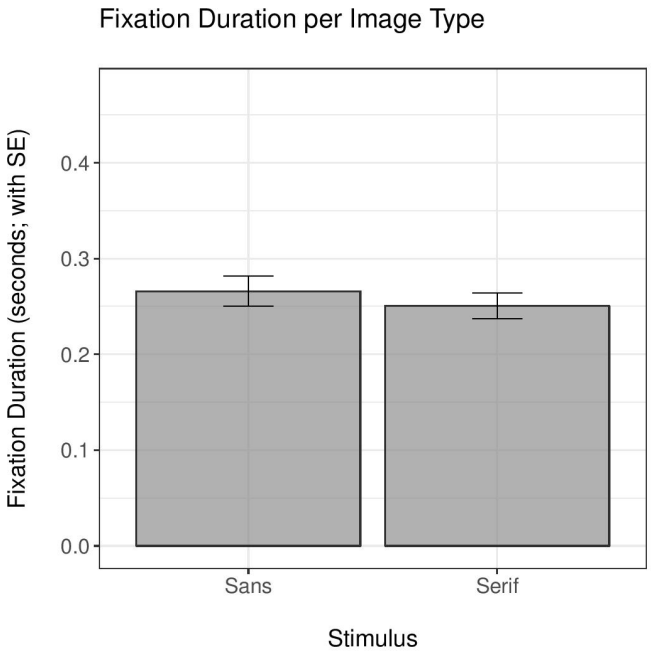

Figure 11: The average time spent on fixations in seconds per font type.

When it comes to reading, speed fixations have a one of the largest impacts, and the time spent per fixation will affect the reading speed greatly. The average time spent in each fixation is show in figure 11. The average time spent in fixations for the Sans font was 0.27 seconds per fixation, while the average time spent in fixations for the Serif font was 0.25 seconds per fixation. This is a difference of 0.02 seconds per fixation. Using one-way ANOVA it results in F = 0.523, p <= 0.05, which is not significant.

Saccade Amplitude

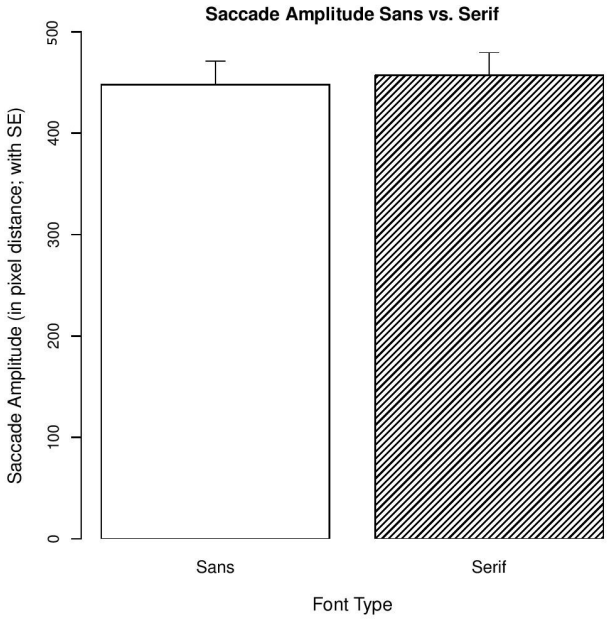

Figure 12: The average saccades size in pixels per font type.

A major part of reading is the size of the saccades that are taken in between fixations. The average size of the saccades are shown in figure 12. The average size of the saccades for the Sans font was 448 pixels, while the average size of the saccades for the Serif font was 457 pixels. This is a difference of 9 pixels. Using one-way ANOVA it results in F = 0.783, p <= 0.05, which is not significant.

Words Between Fixations

The average number of words between each fixation was taken by using a positive number for the how many words forward and a negative number for regressions. This number was totaled and then the average was taken from this (note that re-fixations were not included in this count). The average number of words between fixations is shown in figure 13. The average number of words between fixations for the Sans font was 0.354 words, while the average number of words between fixations for the Serif font was 0.480 words. This is a difference of 0.126 words. Using one-way ANOVA it results in F = 0.663, p <= 0.05, which is not significant.

Re-fixations

The average number of re-fixations was taken by counting each time the subject fixated on the same word that they had fixated on in the previous fixation. The average number of re-fixations are shown in figure 14.

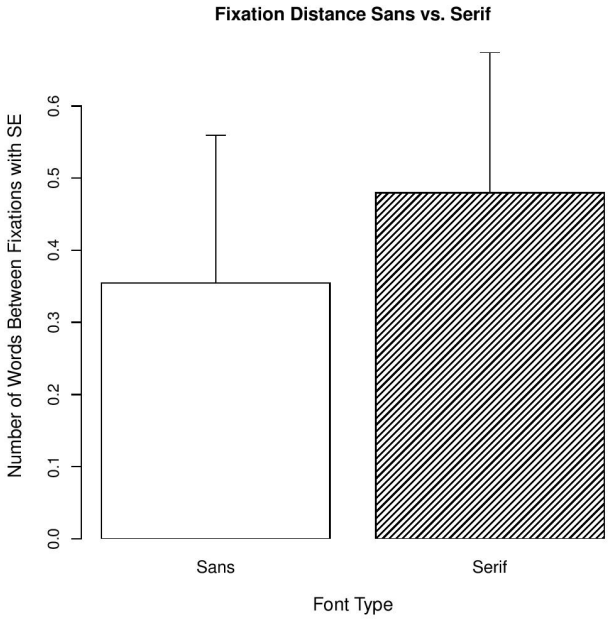

Figure 13: The average number of words between fixations per font type.

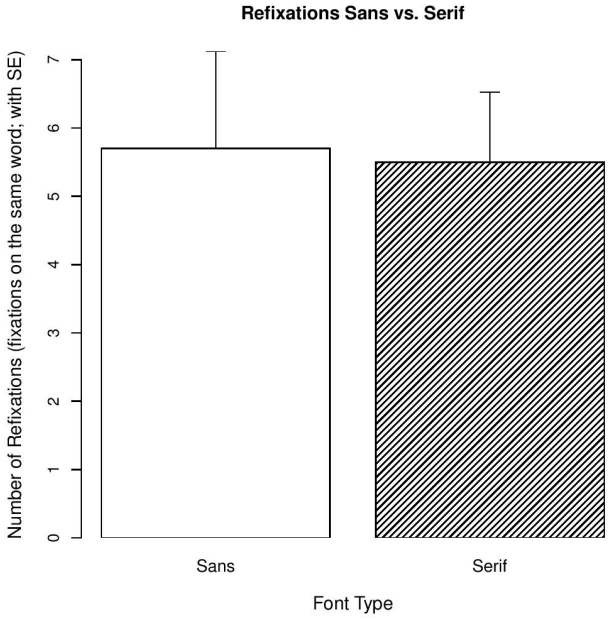

Figure 14: The average number of re-fixations per font type.

The average number of re-fixations for the Sans font was 5.7, while the average number of re-fixations for the Serif font was 5.5. This is a difference of 0.2 re-fixations. Using one-way ANOVA it results in F = 0.913, p <= 0.05, which is not significant.

WPM

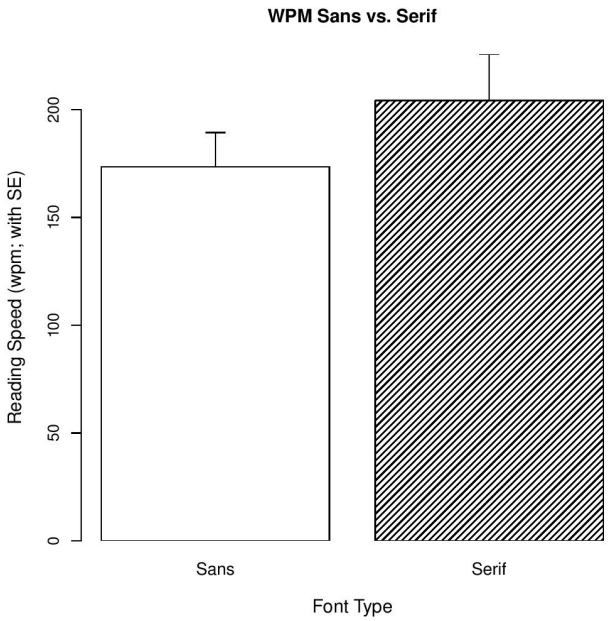

Figure 15: The average number of re-fixations per font type.

The average WPM was taken by taken the list time stamp for each .csv and dividing it by the number of words in the stimulus and multiplying that by 60. The average WPM is shown in figure 15. The average WPM for the Sans font was 173.5, while the WPM for the Serif font was 204.1. This is a difference of 31.4 WPM. Using one-way ANOVA it results in F = 0.266, p <= 0.05, which is not significant.

DISCUSSION

The hypothesis that a Serif font would result in a more efficient medium was categorically proven to not be significant. In every factor that was test (fixation time, saccade length, WPM, re-fixations, and words between fixations) were all shown to not be significant. This could be due to the fact that there was not enough of a negligible difference between the fonts at this text size to prove beneficial.

Improvements

Ways to improve the study would have been to use larger pieces of text at the same font size but to provide a more consistent WPM. Another improvement would have been to get a larger sample size to lower the margin of error and to get a more diverse range of readers. Another possible improvement would have been to include a harder piece of text so that participants would have had to spend more time re-reading for context.

Future Work

Due to the wide number of variables that come with reading there is a plethora of future research that is available. Some future work that could be done with font type based analysis would to be to test different fonts' affects on reading efficiency, the effect that line spacing has on reading comprehension, or the possible implications that age has upon reading efficiency with different types of font.

References

[1] S. Morrison and J. Noyes, *A comparison of two computer fonts: Serif versus ornate sans serif*, `http://usabilitynews.org/a-comparison-of-two-computer-fonts-serif-versus-ornate-sans-serif/`, Accessed: 2016-09-28, 2003.

[2] D. Paterson and M. Tinker, *Journal of Applied Psychology*, vol. 16, pp. 605–613, 3 1932.

[3] K. Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychological Bulletin*, vol. 124, pp. 372–422, 3 1998.

[4] J. M. Henderson and F. Ferreira, "Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control," *Journal of Experimental Psychology*, vol. 16, pp. 417–429, 3 1990.

[5] D. Beymer, D. Russell, and P. Orton, "An eye tracking study of how font size and type influence online reading," Unpublished paper, 2008.