Stat 158 Experiment Final Report

Andrew Chen, Jincen Li, Yingsi Huang

Introduction

The purpose of our experiment is to explore to what extent textural features such as color and font affect human abilities of pattern and word recognition and information retrieval when reading text. There are certain combinations of background color and text font that are less compatible and thus more visually straining to read. These factors may influence the time it takes for one to gain necessary information from text. Our experiment has two factors we are interested in (each with three treatment levels): background color (red, white, yellow) and text font (Times New Roman, Pacifico, and *Pinyon Script*). We used a CB[2] design, with our blocking factor being the subject, since we are taking repeated measures from each subject. From the analysis of our results, we found significant effects from the blocks and the font, as well as significant two-way interactions from all three factors.

Methodology

We chose background color and font as our two treatment factors because we encounter them in our daily lives, and they are the two most defining influences of textual readability. We hypothesized that these factors would lead to statistically significant effects on readability. After pilot runs, we noticed that the within-person variability cannot be ignored in our experiment. This chance-like variability would impede us from accurately comparing the effects of conditions. With subject as our blocking factor, we decided to use the CB[2] design.

From our power calculations, we decided to have nine subjects in total. Each subject was given nine lists of words; all lists had the same 30 words, each word appearing 20 times, but presented different orderings of all 600 words. Since our treatment factors each had three levels, we had a total of nine combinations of background and font. We randomly assigned the nine combinations to the nine word lists. On each of the nine runs, the subject was given 20 seconds to identify as many occurrences of the word “face” in the list as possible. The subject circled the word in the list on an iPad to show that he or she identified the word. Therefore, our response of the experiment is the number of correctly identified words in the list. With the words being circle, it is easy for us to check whether the participants found the right word. Also, one of the advantage of this response is the convenience to collect the data. Also, it prevents us from the potential measurement variability of timing (if we use time as our response). We made sure everyone had the same allotted time to find the word by having them facing forward until the experimenter said “go,” signaling the start of the run.

Run #	Subject	Background	Font	Response
19	3	yellow	times	1
20	3	red	pinyon	1
21	3	yellow	pacifico	2
22	3	red	pinyon	2
23	3	white	pinyon	1

TABLE 1. The example data of first five runs of subject #3

Results

Source	#df	Sum of Squares	Mean Square	F value	Pr(>F)
Subject	*8*	*71.556*	*8.9444*	*4.8973*	*0.000521 ****
Font	*2*	*17.556*	*8.7778*	*4.8061*	*0.014958 **
Background	*2*	*4.222*	*2.1111*	*1.1559*	*0.327572*
Font:Background	*4*	*9.778*	*2.4444*	*1.3384*	*0.277135*
Subject:Font	*16*	*33.556*	*2.0972*	*1.1483*	*0.356966*
Subject:Background	*16*	*72.889*	*4.5556*	*2.4943*	*0.013546 **
Error	*32*	*58.444*	*1.8264*		

TABLE 2. The ANOVA table

From our ANOVA analysis of the resulting data, we found that the blocking factor Subject had a very significant effect on the responses we collected. We can see from the boxplots in Figure 1 that the distributions of the responses varied greatly between subjects, with different medians and spreads.

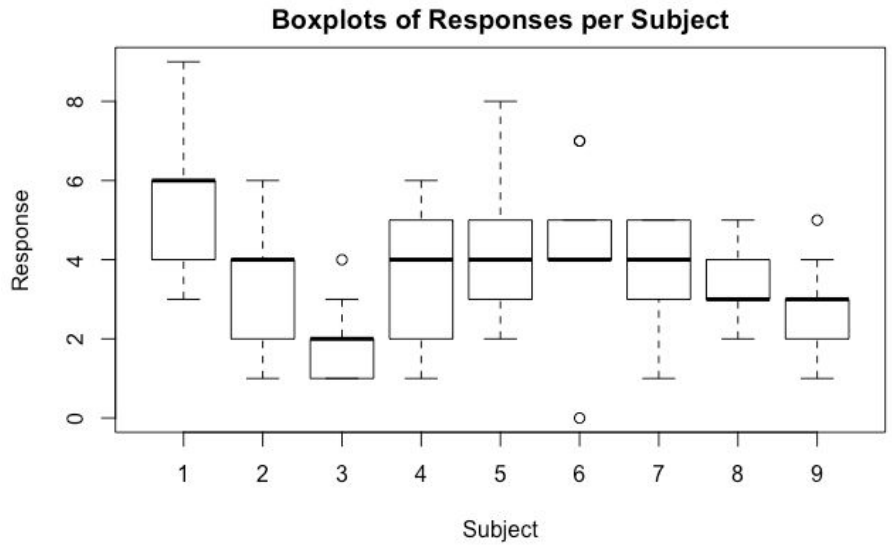

Figure 1

The second most significant effect in our ANOVA analysis came from the interaction between Background and Subject. From the interaction plot between these two factors (Figure 2), we can see that some participants' highest mean responded best with white backgrounds, while others responded best with yellow backgrounds. There was even one subject who responded best with red backgrounds, which we hypothesized would yield the lowest responses. So it is apparent that different subjects vary in their responses to different background colors, affecting the results that we collected.

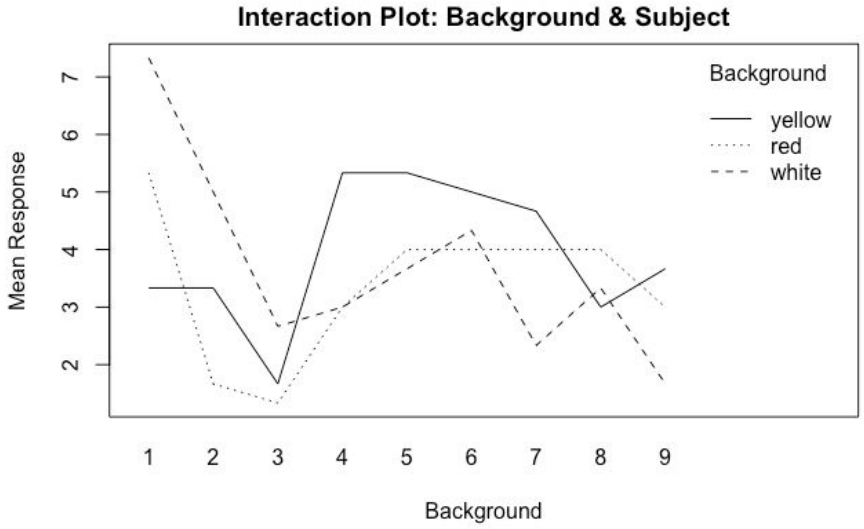

Figure 2

The third most significant effect we found was from the Font factor. Times New Roman yielded the highest mean response (4.3), Pacifico the second highest (3.52), and Pinyon Script the lowest (3.19). The dotchart in Figure 3 shows the the responses for Times are generally higher than those of the other two fonts, except for an outlier within Pinyon of 9. The distributions of the responses generally agreed with our initial hypotheses, that Times would be more readable than Pacifico and Pinyon.

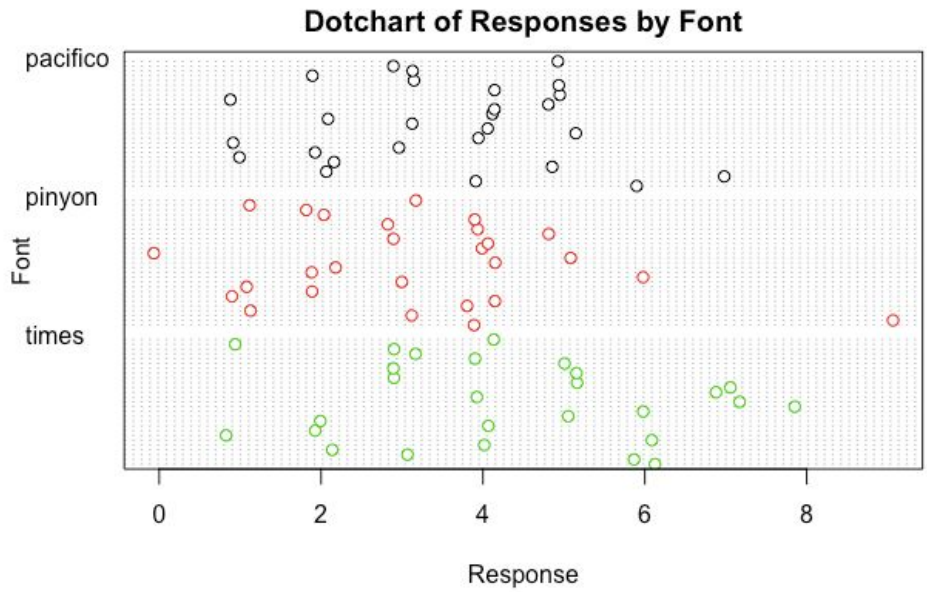

Figure 3

Although declared as not statistically significant in the ANOVA analysis, we also want to point out the interaction between background color and font. Even though Times New Roman generated the highest mean response overall, there were combinations of backgrounds with other fonts that actually generated higher mean responses. In the interaction plot between background and font (Figure 4), we can see that within the red background color, the font Pacifico yielded the highest mean response, which we found surprising. Also, within the yellow background color, Pinyon yielded a higher mean response than Pacifico. These observations suggest that some fonts are more compatible with certain background colors. We can also see from the plot that yellow and white backgrounds with Times New Roman yielded the highest mean responses (4.67), suggesting those combinations to be optimal for readability.

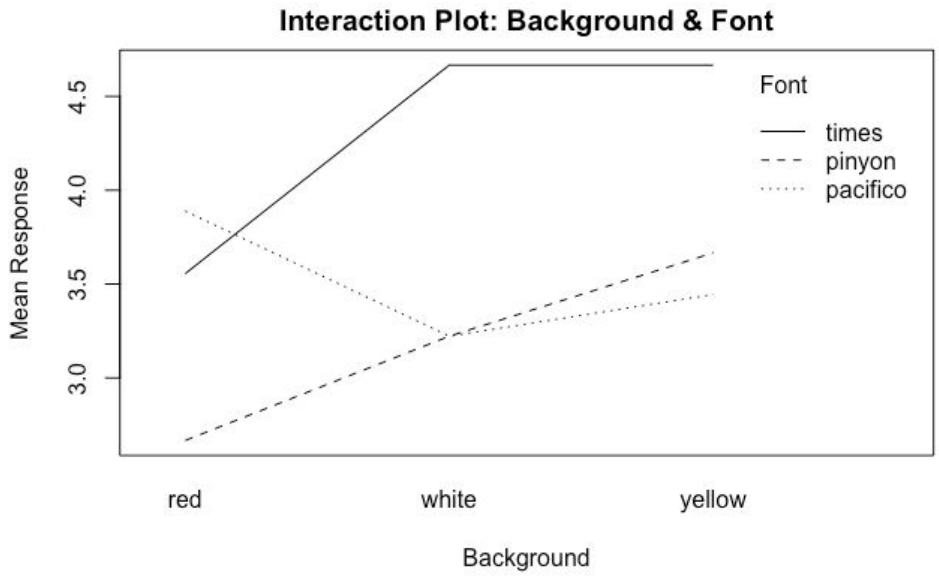

Figure 4

To check our assumptions for the CB[2] design, we also plotted the residuals of the model against the fitted values. Looking at Figure 5, we noticed that the data points were roughly centered around 0 and randomly scattered, without any significant pattern based on treatment combination, thus satisfying the constant variance assumption.

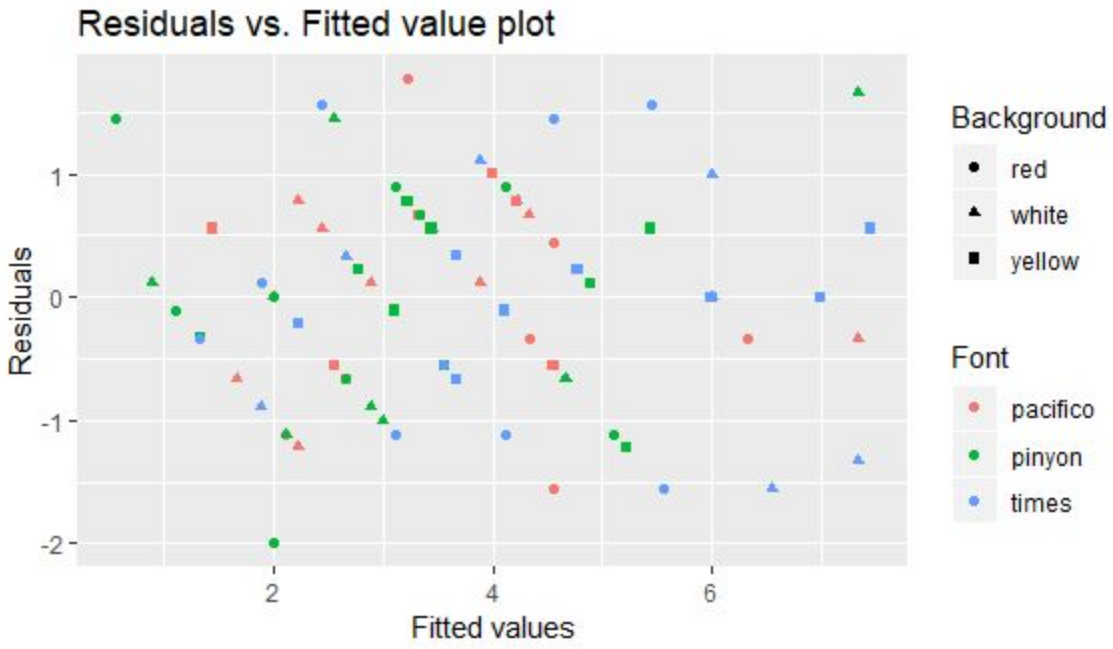

Figure 5

Conclusion

From analyzing the results of the experiments through ANOVA and interaction plots, we found that different combinations of color and font significantly affected individuals' visual perception and readability of text. The distribution of the responses varied between the subjects, which indicated that individual differences are significant. When we ran the analysis of the experiment, we first looked at the effects of our treatment factors separately. Between the three levels of background color, white background yielded the highest mean response, and among the three levels of text font, Times New Roman gave the best performance. We also found that certain text fonts seemed more compatible with specific background colors (i.e. red background with Pacifico and yellow with Pinyon Script), suggesting a significant two-way interaction between font and background. The other two-way interaction we found significant was between background and subject, since different subjects seemed to perform best under varying background colors. Overall, white and yellow backgrounds with Times New Roman resulted in the best mean response; in which case, optimal readability is obtained and individuals can retrieve the most information from the text.

Appendix

Raw Data (no notes were taken)

Participant	Background	Font	Response	Notes
1	red	times	6	
1	red	pinyon	4	
1	red	pacifico	6	
1	white	pinyon	9	
1	white	times	6	
1	yellow	pinyon	3	
1	yellow	times	3	
1	yellow	pacifico	4	
1	white	pacifico	7	
2	red	pinyon	1	
2	red	times	2	
2	yellow	pinyon	4	
2	yellow	times	4	
2	white	pinyon	4	
2	yellow	pacifico	2	
2	white	pacifico	5	
2	red	pacifico	2	
2	white	times	6	
3	yellow	pinyon	1	
3	red	pacifico	1	
3	yellow	pacifico	2	
3	red	pinyon	2	
3	white	pinyon	1	
3	white	pacifico	3	
3	red	times	1	
3	yellow	times	2	
3	white	times	4	
4	red	times	2	
4	white	pacifico	1	
4	yellow	pacifico	4	
4	white	times	5	
4	white	pinyon	3	
4	yellow	pinyon	6	
4	red	pacifico	5	
4	red	pinyon	2	
4	yellow	times	6	
5	white	pinyon	2	
5	yellow	times	8	
5	white	times	7	
5	red	pacifico	4	
5	red	times	4	
5	red	pinyon	4	
5	yellow	pinyon	5	
5	yellow	pacifico	3	
5	white	pacifico	2	
6	red	pinyon	0	
6	yellow	pinyon	4	
6	yellow	pacifico	4	
6	white	pinyon	4	
6	yellow	times	7	
6	white	pacifico	4	
6	red	pacifico	5	
6	red	times	7	
6	white	times	5	
7	white	pacifico	1	
7	white	pinyon	3	
7	red	pinyon	5	
7	yellow	pacifico	5	
7	white	times	3	
7	yellow	times	5	
7	yellow	pinyon	4	
7	red	times	3	
7	red	pacifico	4	
8	white	times	5	
8	yellow	pinyon	3	
8	red	pacifico	5	
8	yellow	times	4	
8	red	pinyon	4	
8	white	pacifico	3	
8	yellow	pacifico	2	
8	red	times	3	
8	white	pinyon	2	
9	red	pinyon	2	
9	yellow	times	3	
9	white	pinyon	1	
9	white	pacifico	3	
9	white	times	1	
9	red	times	4	
9	red	pacifico	3	
9	yellow	pacifico	5	
9	yellow	pinyon	3	

R Code

```
#######################
#Reading in data and creating ANOVA table (Table 2 in text)
library(ggplot2)
experiment$Participant <- as.factor(experiment$Participant)
mod <- lm(Response ~ Participant + Font * Background + Participant : Font + Participant :
Background,data = experiment)
anova(mod)
#######################
```

```
######################
#Boxplot of responses grouped by subject (Figure 1 in text)
boxplot(Response ~ Participant, experiment, main = "Boxplots of Responses per Subject", xlab =
"Subject", ylab = "Response")
######################

######################
#Interaction plot between subject and background (Figure 2 in text)
interaction.plot(experiment$Participant, experiment$Background, experiment$Response, xlab =
'Background', ylab = 'Mean Response', trace.label = 'Background', main = "Interaction Plot:
Background & Subject")
######################

######################
#Dotchart of responses grouped by font (Figure 3 in text)
dotchart(jitter(experiment$Response), groups = experiment$Font, color = experiment$Font,
main = "Dotchart of Responses by Font", ylab = "Font", xlab = "Response")
######################

######################
#Interaction plot between font and background (Figure 4 in text)
interaction.plot(experiment$Background, experiment$Font, experiment$Response, xlab =
'Background', ylab = 'Mean Response', trace.label = 'Font', main = "Interaction Plot: Background
& Font")
######################

######################
#Residuals vs. Fitted Values Plot (Figure 5 in text)
ggplot(data = experiment) +
  geom_point(aes(x = mod$fitted.values,y = mod$residuals,color = Font,shape = Background)) +
  ggtitle("Residuals vs. Fitted value plot") +
  xlab("Fitted values") +
  ylab("Residuals")
######################
```