Stats 101B Final Research Paper
The Effect of Coffee on Short-Term Memory

Angelique Rubio, Ami Pant, Joyce Yii, Kiana Mohammadinik

Abstract

Coffee is one of the most popular beverages consumed by people from all around the world (Abalo 2021). For many, coffee is fuel - it acts as a source of energy because of its ability to help stave off tiredness and increase alertness throughout the day. However, its effects on memory have not been sufficiently evaluated (Sherman et. al 2016). Thus, in our study, we aim to explore the possible effects that different levels of caffeine might have on short-term memory. Our study is set up to focus on a representative sample of the Island's population using a two-way randomized factorial design by performing a two-way ANOVA in R studio. We sampled a total of 198 participants (ages 20-30) and randomly assigned the participants to the treatment groups so that there were 22 samples in each group. These results could determine if coffee consumption affects short-term memory. Based on our data ANOVA, we failed to reject the null hypothesis that there is not a significant difference in short-term memory due to caffeine intake from different coffees.

---

1. Introduction

The most well-known component of one of the most popular beverages consumed worldwide, coffee, is caffeine (Abalo 2021). It is one of the most commonly consumed stimulants in the world, with upwards of 80% of the population consuming caffeine daily (Sherman et. al 2016). Individuals have identified various reasons for their heavy reliance on caffeine, including an increase in alertness and reaction time. While there is an abundance of research outlining the effects of caffeine on the aforementioned topics, as well as cognition, implicit, and explicit memory, there is limited research on the effects of caffeine on short-term memory (Sherman et. al 2016). There are many techniques researchers utilize to test short-term memory such as immediate object recall, immediate word recall, or pair matching (Syce 2019). Prior research has indicated that immediate word recall is influenced by two related cognitive constructs: attention and short-term episodic memory (Gavett et. al 2012). Therefore, it was decided that short-term vocabulary recall would be an adequate measure for short-term memory.

A myriad of factors were taken into consideration when deciding how participants were chosen for this study, including results from existing studies. While some previous studies have found significantly different results for men and women (Abalo 2021, Johnson-Kozlow et al. 2002), the most significant factors that influenced the effect of caffeine on individuals were age (Syce, May et. al) and the time of day in which the participants were administered their treatment (May et. al. 2005, Syce 2019, Sherman et al. 2016). For that reason, this study selected both male and female participants between the ages of 20 and 30 years old to minimize the variance caused by age and isolate the effect of various levels of caffeine on an individual administered at various times of the day. In order to ensure the caffeine had time to enter the participants' system, we chose to administer the second memory test 20 minutes after applying the treatment.

This study aims to explore the effects of various levels of caffeine on short-term memory using a vocabulary recall test on a simulation of participants. The null hypothesis is that there is no significant difference in short-term memory due to caffeine from different coffees. We propose that the most concentrated level of caffeine, espresso, will have the greatest effect on individuals during the peak performance time for the age group we selected, which would be in the evening (May et al. 2005).

2. Methods and Procedure

2.1 Participants

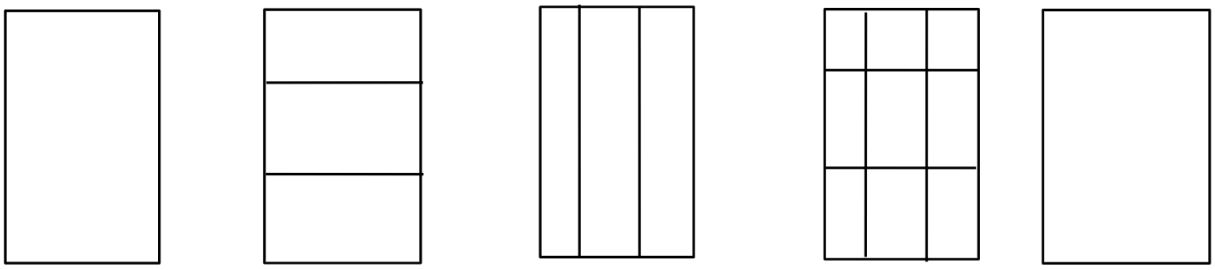
The participants for our study are virtual humans, both male and female, between the ages of 20 and 30 chosen from an online resource known as The Island. Treatments were randomly assigned to the islanders in the contact list of each researcher who collected data for this study. No particular cities were excluded so that our sample can be representative of the entire island population that falls within our designated age range.

2.2 Design

This study follows a two-way randomized factorial design. The parameters for the design are outlined below.

Response Variable	Difference in Percentage Correct on Vocabulary Test		
Treatment 1 (Caffeine)	Control	Coffee	Expresso
Treatment 2 (Time)	Morning	Afternoon	Evening

The factor diagram is illustrated below.



We chose to focus on the time of day because the literature revealed that is a significant factor in the way caffeine influences an individual. For this study, the morning was defined between the hours of 7 AM-10 AM, afternoon 12 PM-3 PM, and evening 6 PM-9 PM. The three levels of caffeine consumption were chosen based on the available options provided on the island, with the control being 250 mL of water (0 milligrams of caffeine), coffee 250mL (approximately 95 milligrams of caffeine), and espresso 60mL (approximately 64 milligrams of caffeine).

2.3 Material and Procedure

Step 1: Find subjects on the Island who consent to participate in our experiment and generate our contact list of participants.
Step 2: Randomly assign these individuals into different treatment groups

Group 1: Water, Coffee, or Espresso in the morning
Group 2: Water, Coffee, or Espresso in the afternoon
Group 3: Water, Coffee, or Espresso in the evening

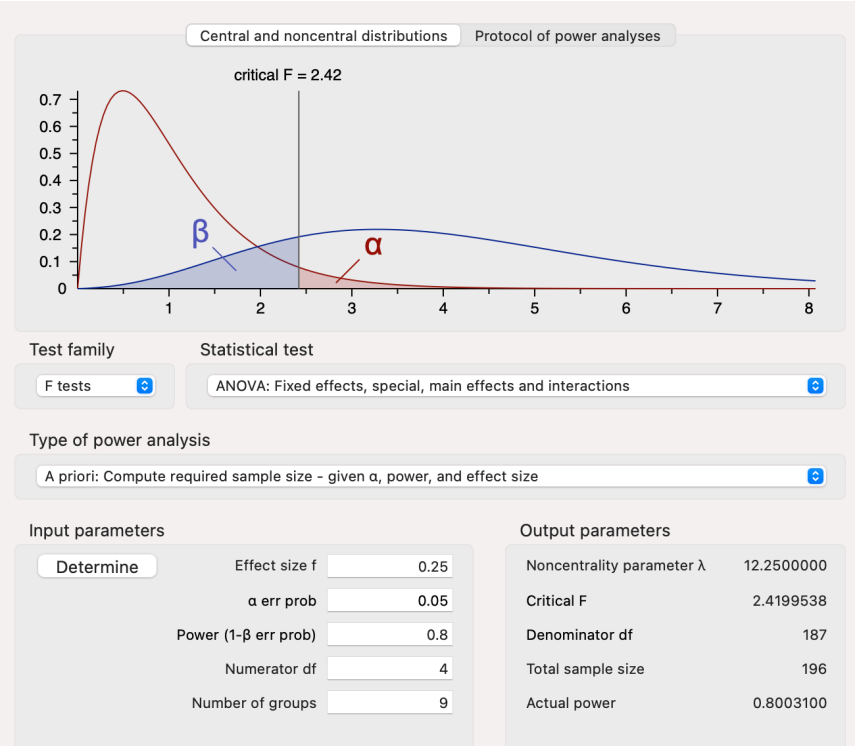
Step 3: For each unit (one islander) measure their performance on the memory vocabulary test (and convert the raw scores given by the island into percentages to make data manipulation easier).
Step 4: For each unit, administer the assigned treatment (coffee, water, or espresso in the morning, afternoon, or evening).
Step 4: For each unit, measure their performance on the memory vocabulary test 20 minutes after administering treatment.
Step 5: For each unit, compute the difference in percentages before and after treatment which will be used as our response variable.

2.4 Instruments

We conducted our experiment on individuals from virtual populations living on The Islands (https://islands.smp.uq.edu.au/). The G*Power statistics applet was utilized to determine the sample size based on the experiment design. We selected participants (male and female) between the ages of 20 and 30 from varying cities between the different islands. On The Islands, we utilized the memory test vocabulary task to measure short-term memory before and after administering a water, coffee (250mL), or coffee espresso (60 mL) treatment at the three-time points: morning, afternoon, and evening. All data was recorded on a Google Sheets document and then analyzed using base R as well as the “ggplot2”. “ggfortify” and “readr” R Studio packages. The “ggplot2” and “ggfortify” packages were employed to create the plots for data visualizations.

2.5 Sample Size Determination

We determined the sample size of our study using the statistical software G*Power. We used a two-way factorial design with an effect size of 0.25, power of 0.8, and significance level of $\alpha = 0.05$. With all this and based on our numerator df of 4, the sample size required is 196. However, we used a sample size of 198 to achieve a balanced design, so that there are 22 samples for each group.



3. Results

3.1 Evaluating Model Validity with Residual Diagnostics

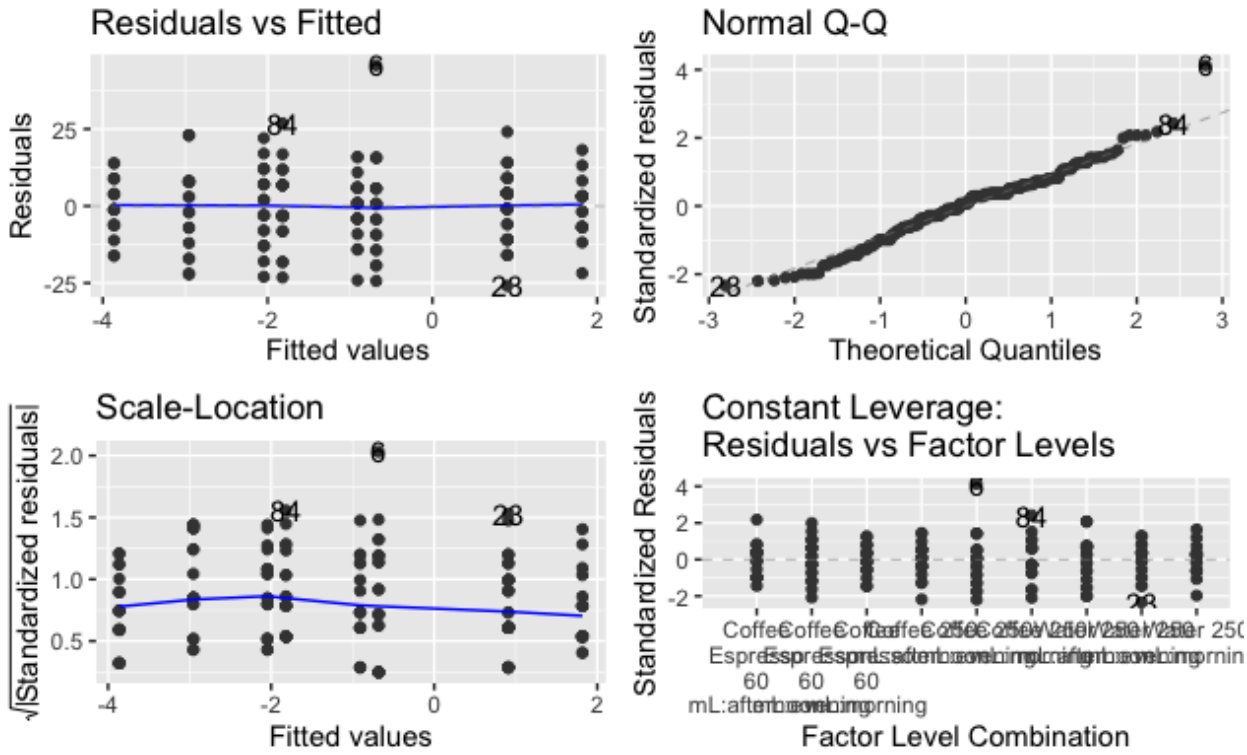


Figure 1: Residual Diagnostics of the Model. To check whether our model was correct we analyzed the residual, normal QQ, scale-location, and constant leverage plots. The residual and scale location plots do not portray any significant pattern or fan-shaped. The points in the normal QQ plot closely follow the y = x line. The constant leverage plot does visualize some high leverage points: 5, 198, and 84. None of these points are outside Cook's distance.

3.2 Mean Differences in Memory by Coffee and Time of Day Per Group

Coffee	Time of Day	Difference in Memory
Coffee 250 mL	Afternoon	-0.9090909
Coffee Espresso 60 mL	Afternoon	0.9090909
Water 250 mL	Afternoon	-2.9545455
Coffee 250 mL	Evening	-0.6818182
Coffee Espresso 60 mL	Evening	-2.0454545
Water 250 mL	Evening	0.9090909
Coffee 250 mL	Morning	-1.8181818
Coffee Espresso 60 mL	Morning	-3.8636364
Water 250 mL	Morning	1.8181818

Table 1: Mean Differences in Memory by Each Design Factor

3.3 Tukey HSD Adjusted P-Values and Plots

Comparisons	Difference	Lower	Upper	P-Value Adjusted
Coffee Espresso-Coffee	-0.530303	-5.187855	4.127249	0.9609252
Water-Coffee	1.060606	-3.596946	5.718158	0.8527720
Water-Coffee Espresso	1.590909	-3.066643	6.248462	0.6992817
Evening-Afternoon	0.3787879	-4.278765	5.036340	0.9798658
Morning-Afternoon	-0.3030303	-4.960583	4.354522	0.9870661
Morning-Evening	-0.6818182	-5.339371	3.975734	0.9362493

Table 2: Post-Hoc Analysis of Differences in Change in Memory Test Results Between Different Levels of the Coffee Treatment and Different Times of Day. Tukey's Honesty Significance indicates that there is not sufficient evidence to suggest any significant differences of means of outcome between the levels of each factor as all of the outputted adjusted p-values are larger than the significance level of 0.05.

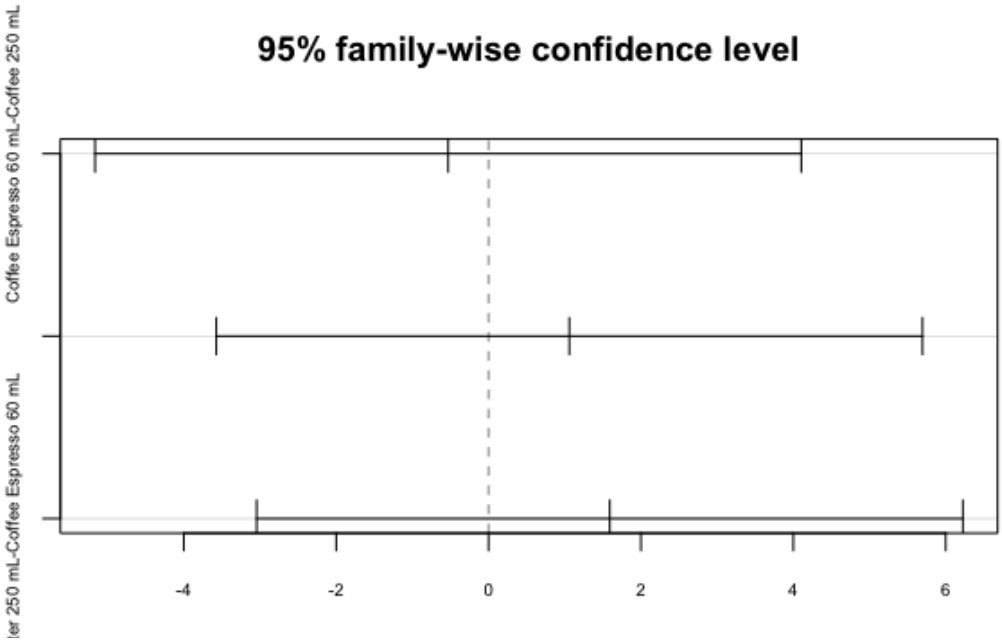


Figure 2: Tukey's HSD Plot For Memory Differences Among Coffee Types

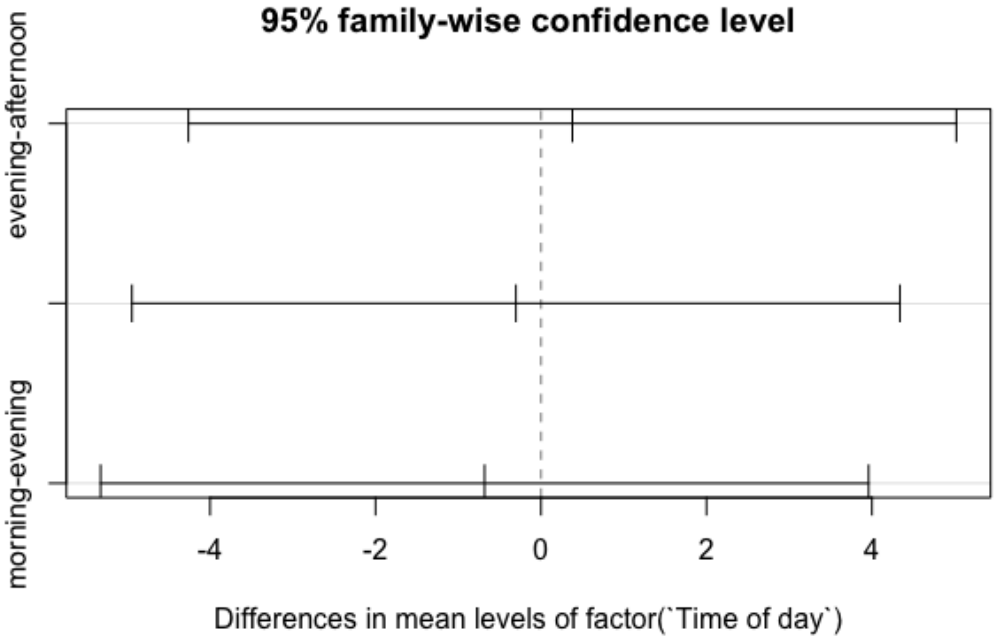


Figure 3: Tukey's HSD Plot For Memory Differences By Different Times of Day

3.4 Box plots

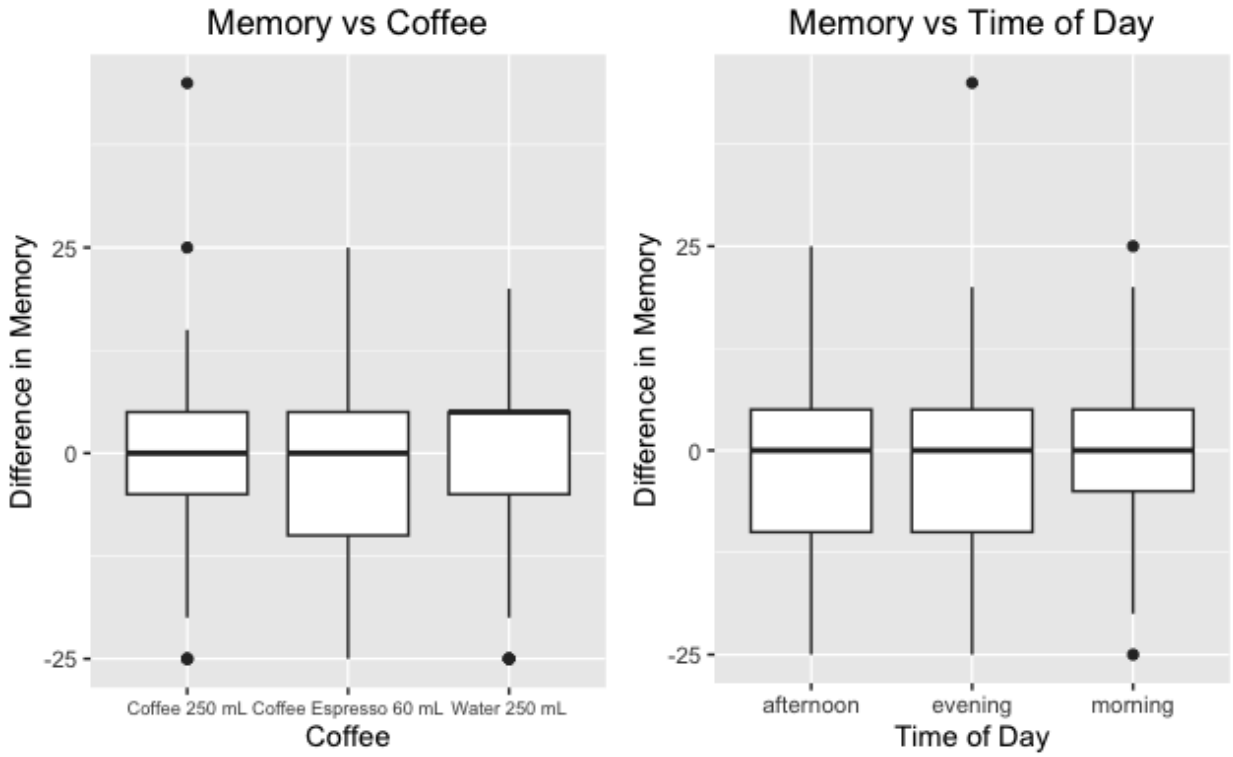


Figure 4: Side by Side Boxplots Separately Examining the Effects of Time of Day and Coffee Type on the Difference of Memory Test Results. The median of each group is depicted by the black lines cutting across each "box with the box itself depicting the first and third quantiles and potential outliers depicted by the black dots outside of the boxes.

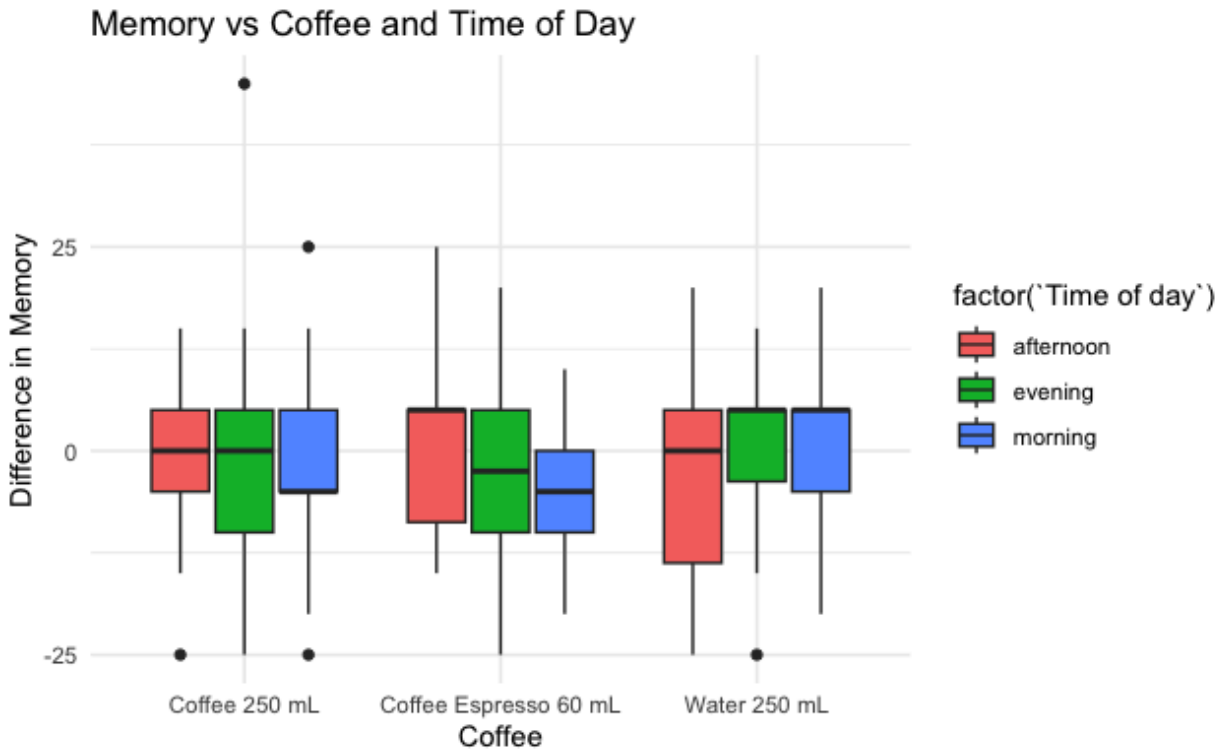


Figure 5: Side by Side Box Plot of Memory Test Difference vs. Times of Day Grouped by Coffee Type. The median of each group is depicted by the blacklines cutting across each "box with the box itself depicting the first and third quantiles and potential outliers depicted by the black dots outside of the boxes.

3.5 Interaction plot

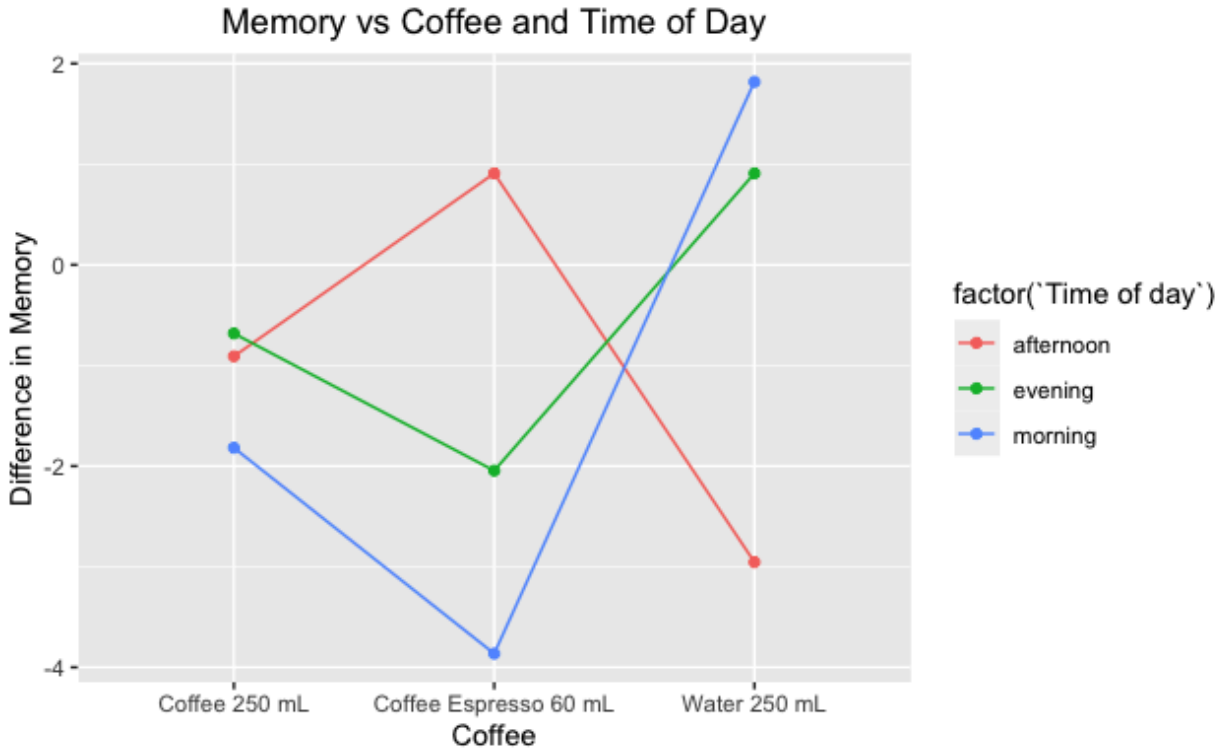


Figure 6: Interaction plot of Times of Day and Coffee Types with Difference in Memory Test Results as Response. Due to a lack of parallel patterns in the interaction plot, we can make the conclusion that interaction between the predictors exists.

3.6 Data Analysis

The model, model validity check, data visualization, two-way ANOVA, and post-hoc analysis were all conducted on R and R studio. We created a model with the aov() function, which included coffee and time of day design factors as well as the interaction between the two. The inclusion of the interaction in the model was further confirmed with the interaction plot (Figure 6) since there is a lack of parallel patterns among the lines. This interaction plot was created with the ggplot() function along with stat_summary() functions.

Model validity was checked with the autoplot() function from the "ggfortify" package. The model validity plots, residuals, normal qqplot, scale location, and constant leverage, convey model validity. The lack of any significant pattern or fan shape portrays that there is constant variance. In addition to this, the lack of any pattern in the residual plot indicates a normal distribution spread. The points following the y = x line in the normal QQ plot also convey normality. There are some high leverage points in the constant leverage plot: 5, 198, and 84. Since none of these points are outside Cook's distance, none of the high leverage points are bad, high leverage points.

Once the validity of the model was confirmed, we calculated all the average differences in memory test scores for each unique group, Coffee 250mL afternoon, Coffee Espresso 60mL afternoon, Water 250 mL afternoon, Coffee 250mL morning, Coffee Espresso 60mL morning, Water 250 mL

morning, Coffee 250mL evening, Coffee Espresso 60mL evening, and Water 250 mL evening, via the aggregate() function. An analysis of the multiple comparisons of these means was conducted using the TukeyHSD() function. This would allow us to visually determine the source of variation in the treatment. All the different boxplots, which were plotted with the base ggplot() function along with geom_boxplot() included also serve a similar purpose. They help visualize any outliers and the overall spread of the effect of time of day and coffee type on the difference in short-term memory, both separately and together.

After confirmation of model validity, we also conducted the two-way ANOVA on the collected data using R and RStudio software with the anova() function. We intended to test if caffeine will influence an individual's performance on a vocabulary test using F-testing within treatments to determine if there is a significant difference between groups. Our control group (water 250 mL) will serve as a baseline statistic which will then be compared to our treatment groups.

ANOVA Results

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Coffee	2	86.6	43.308	0.3379	0.7137
Time of Day	2	15.4	7.702	0.0601	0.9417
Coffee:Time of Day	4	538.4	134.596	1.0500	0.3827
Residuals	189	24227.3	128.187		

Table 3: Two way ANOVA Table with Interactions. The p-value of 0.7137 which is much larger than the significance level of 0.05 indicates that coffee by itself does not have a statistically significant effect on the memory test results as the difference of memory test results before and after the consumption of coffee is shown to be not significant. The p-value of 0.9417 which is also much larger than the significance level of 0.05 indicates that the time of day that coffee is consumed also does not have a statistically significant effect on memory test results by itself. The p-value of 0.3827 for the interaction term also indicates that the interaction between the two factors of coffee and time of day does not make a statistically significant difference in the memory test results either. Based on these results, we fail to reject the null hypothesis and make the conclusion that there is no statistically significant association between coffee, time of day for drinking coffee, the interaction between the two factors, and short-term memory.

4. Discussion

Our study aimed to explore the effects of caffeine on individuals' performance on a vocabulary test at various times of the day. The results of our study ultimately indicate that neither the caffeine, time of day, nor their interaction yielded significant changes in an individual's performance on a vocabulary test. This can be concluded by the p-values for the F-test of each factor and their interaction being greater than our previously specified threshold of 0.05.

A deeper analysis of our data reveals that there was a difference in the control condition for each time of day, despite these differences not being significant. In the morning, the average vocabulary test score was approximately 3.62% less than the control after the coffee treatment and 5.67% less after the espresso treatment. For the afternoon, participants performed approximately 2.05% better after the coffee treatment and 3.86% better for the espresso treatment. In the evening, the test scores were approximately 1.59% less than the control after the coffee treatment and 2.95% less after the espresso treatment. These averages are not consistent with our proposed hypothesis, which is that test scores would improve after both the coffee and the espresso treatments in the evening given the age range of the participants in our study. Furthermore, it was unexpected that participants' scores would worsen in the control condition in the afternoon since that is closer to peak performance time, while the average difference in memory was actually the highest during the morning which is the farthest from peak performance time. Overall the largest average difference occurred after the espresso treatment in the morning, indicating a potential intervention that warrants further observation and experimentation.

The boxplots made from our data provides us insight into the factors within our study. The results shown in our boxplots are consistent with the null hypothesis that there is no significant difference in short-term memory due to caffeine from different coffees. Moreover, our boxplots contradicts our initial conclusion that the most concentrated level of caffeine, espresso, will have the greatest effect on individuals in the evening. Overall, the interaction between coffee types and times of day does affect the memory results but not enough to conclude that it creates a statistically significant difference in short-term memory.

Our results were shocking given their inconsistency with previous literature, suggesting there may be ways to improve this study and replicate it in order to yield more significant results. One potential reason our results did not reveal a substantial influence of caffeine on an individual’s vocabulary test performance is that we did not allow adequate time for the caffeine to be absorbed into the bloodstream of the individual. Coffee can take anywhere from 15 minutes to 90 minutes to absorb before one can begin to see or feel any effects; moreover, this time can also vary based on various factors of the individual, including their BMI, metabolic rate, and history of caffeine consumption. If this study were to be replicated in the future, we would advise taking this into account and perhaps adopting a repeated measures design or adding in blocks to account for these sources of variation. Overall, this study offers an excellent baseline for future research to continue to explore the effects of caffeine on short-term memory.

References

Abalo R. (2021) Coffee and Caffeine Consumption for Human Health. *Nutrients*. 13(9):2918.

Alharbi, W. D., Azmat, A., & Ahmed, M. (2018). Comparative effect of coffee robusta and coffee arabica (Qahwa) on memory and attention. *Metabolic brain disease*, *33*, 1203-1210.

Gavett, Brandon. Horwitz, Julie. (2012). Immediate List Recall as a Measure of Short-Term Episodic Memory: Insights from the Serial Position Effect and Item Response Theory, *Archives of Clinical Neuropsychology*, 27(2), 125–135.

Johnson-Kozlow, M., Kritz-Silverstein, D., Barrett-Connor, E., & Morton, D. (2002). Coffee consumption and cognitive function among older adults. *American journal of epidemiology*, *156*(9), 842-850.

May, C. P., Hasher, L., & Foong, N. (2005). Implicit memory, age, and time of day: paradoxical priming effects. *Psychological Science*, *16*(2), 96-100.

Sherman, S. M., Buckley, T. P., Baena, E., & Ryan, L. (2016). Caffeine enhances memory performance in young adults during their non-optimal time of day. *Frontiers in psychology*, *7*, 1764.

Syce, D. V. (2019). Can a cup of black coffee enhance cognitive function and short-term memory?. *Al Ameen Journal of Medical Sciences*, *12*(2), 90-94.