1.0 Introduction (max 1 page):

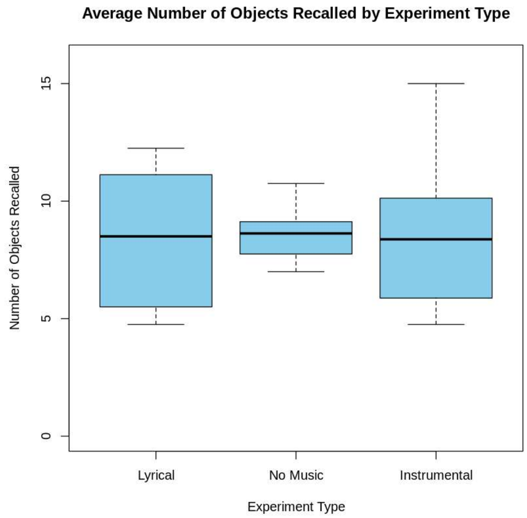
This experiment explores the question “How does the presence and genre of music affect short-term recall memory?” Music is a common external stimuli found coexisting with studying and surveys show around 60% of students listen to music while studying [1]. Visual working memory forms the basis of many cognitive processes and general working memory has been shown to have high correlation with improved academic achievement, particularly in reading and mathematical work [2, 3].

The effects of music on memory retention is an ongoing controversial topic. According to one study on the Mozart Effect and arousal-mood-hypothesis, certain genres, such as classical music, help spatial memory and the brain to absorb and interpret new information [4]. On the other hand, other studies argue that any kind of background music reduces working memory capacity and that it could have detrimental effects, particularly when lyrics are involved [5, 6]. Moreover, other studies suggest that the positive and negative effects of background music averaged to a null impact [7].

With such existing studies that claim contradicting effects of music when studying – beneficial and detrimental, the team has determined to further investigate the topic. Accordingly, this experiment will determine whether music impacts memory retention and if different types of music, specifically instrumental and lyrical, have contrasting effects, and will gather implications to help inform students’ strategies for optimizing cognitive performance in learning environments.

The experiment was participated by a total of 24 students, ages 18-25 without visual or hearing impairments. They were put in 3 groups of 8, with an equal distribution of genders, of the different independent background music variables – instrumental, lyrical, and no music. Participants were each shown images of 16 objects for 30 seconds and asked to recall the objects immediately afterwards for four rounds. With data collected from the experiment, the following null hypothesis was tested: The mean number of images recalled for lyrical, instrumental, and no music are the same. The alternate hypothesis is that the mean of at least one of the three test categories is different.

2.0 Assumption Checks and Validity Discussion (max 1.5 page):
The one-way ANOVA and Tukey tests make the following assumptions about data such that there are no significant outliers, assumption of normality, assumption of equal variance. To assess the validity of no significant outlier assumptions two methods were employed: visualization in a box plot and outlier table from the rstatix package (Figure 1). From the graph, it is visually clear that there are no outliers, and the table shows empty results, confirming no outliers exist.



A data.frame: 0 × 11

Participant	Age	Gender	Experiment.Type	Number.of.Objects.1	Number.of.Objects.2	Number.of.Objects.3	Number.of.Objects.4	Average.Recall	is.outlier	is.extreme
<int>	<int>	<chr>	<fct>	<int>	<int>	<int>	<int>	<dbl>	<lgl>	<lgl>

A data.frame: 0 × 11

Participant	Age	Gender	Experiment.Type	Number.of.Objects.1	Number.of.Objects.2	Number.of.Objects.3	Number.of.Objects.4	Average.Recall	is.outlier	is.extreme
<int>	<int>	<chr>	<fct>	<int>	<int>	<int>	<int>	<dbl>	<lgl>	<lgl>

A data.frame: 0 × 11

Participant	Age	Gender	Experiment.Type	Number.of.Objects.1	Number.of.Objects.2	Number.of.Objects.3	Number.of.Objects.4	Average.Recall	is.outlier	is.extreme
<int>	<int>	<chr>	<fct>	<int>	<int>	<int>	<int>	<dbl>	<lgl>	<lgl>

Figure 1. Box Plot Method (top), identify_outliers table is empty, indicating no outliers (bottom)
Then normality was checked by computing the Shapiro Wilk test for each data point, and the Q-Q plot for demonstration (Figure 2).

A tibble: 3 × 4

Experiment.Type	variable	statistic	p
<fct>	<chr>	<dbl>	<dbl>
Lyrical	Average.Recall	0.9258169	0.4788114
No Music	Average.Recall	0.8996195	0.2866873
Instrumental	Average.Recall	0.9609343	0.8189524

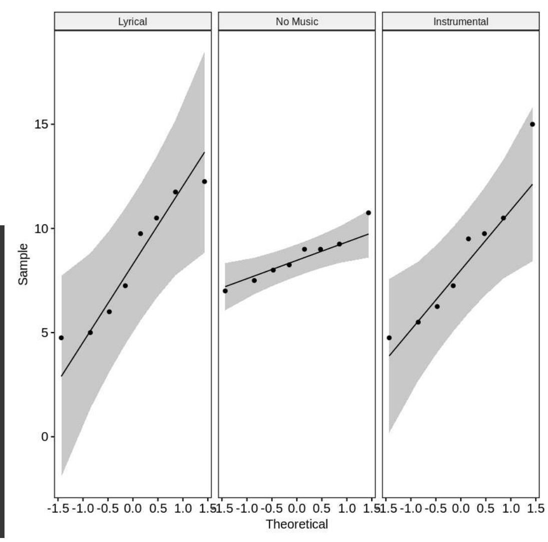


Figure 2. Shapiro Wilk test (left) and QQ plot (right).
The p-value is greater than 0.05, confirming normality. Additionally, the points in the plot lie approximately along the reference line, so normality is assumed. Figure 3 checks the equal variance assumption via Levene's test. Since the p-value is < 0.05, we reject the hypothesis and conclude that the variances are not equal. This assumption is not valid for this test, and ANOVA in this form is not acceptable. However, alternative methods are out of scope for this course, thus, we moved forward to conducting the ANOVA test.

A anova: 2 × 3

	Df	F value	Pr(>F)
	<int>	<dbl>	<dbl>
group	2	4.877205	0.01820853
	21	NA	NA

Figure 3. Levene's test for checking equal variances

3.0 Results (max 1.5 pages):

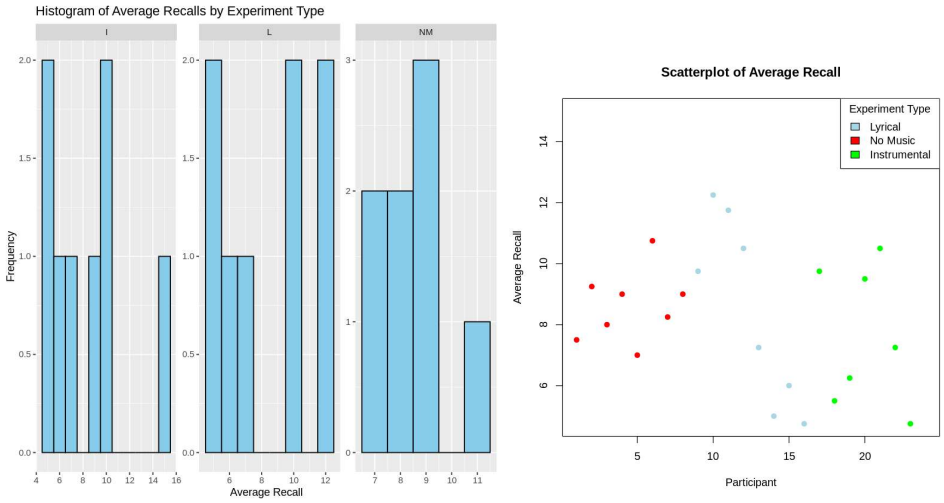


Figure 4. Histogram of average recall count frequency by music type (left) and scatterplot of average recall values (right).

The histogram and scatter plot in Figure 4. provide a visual representation of the distribution of average recall for each experimental condition (I -- Instrumental, NM -- No music, L -- Lyrical). This allows for observing the shape, central tendency and spread of the data.

	Experiment.Type	mean	median	std_dev
	<chr>	<dbl>	<dbl>	<dbl>
1	I	8.56	8.38	3.35
2	L	8.41	8.5	3.03
3	NM	8.59	8.62	1.17

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Experiment.Type	2	0.16	0.081	0.011	0.989
Residuals	21	152.33	7.254		

Figure 5. Summary statistics of data (left) and ANOVA table (right).

Figure 5 shows the summary statistics organized by experiment type. Generally, all experiment types from the sample have similar recall performance means and medians, while there is greater variability in Instrumental and Lyrical data compared to no music. The F-value for the ANOVA test is < than critical value f*(v1 = 2, v2 = 21, a = 0.05) = 3.47, meaning we fail to reject (FTR) H_o that the suggesting all music types are equal at the 95% significance level.

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Average.Recall ~ Experiment.Type, data = stats)

$Experiment.Type
                          diff       lwr      upr     p adj
No Music-Lyrical       0.18750 -3.206792 3.581792 0.9893744
Instrumental-Lyrical   0.15625 -3.238042 3.550542 0.9926078
Instrumental-No Music -0.03125 -3.425542 3.363042 0.9997031
```

Figure 6.: Tukey's test conducted on data by pair.

Since we FTR H_o, pairwise comparisons with Tukey's test are **not needed**. However, it has been provided in Figure 6 for demonstration purposes. Since p adj > alpha (0.05), we can once again conclude: FTR that for each pair of music types, the mean is equal.

4.0 Conclusion and Summary (1 page):

Using the various different statistical methods outlined in the report, the question of how the presence of different music affects studying, particularly short-term memory recall was explored. The setup of the experiment involved 24 individuals, all within the age range of 18-25. Each participant was assigned a music type {no music, instrumental, lyrical and instrumental} and participants underwent 4 rounds of the same activity, where images changed each round. All participants were screened to ensure any hearing impairments they may have would not affect the results of this experiment.

To recap, the following hypothesis were developed for this particular experiment:

- Null hypothesis: The mean number of images recalled for lyrical, instrumental, and no music are the same.
- Alternate hypothesis: The alternate hypothesis is that the mean of at least one of the three test categories is different.

The statistical methods employed allowed for data visualization and inspection (checking for assumptions), summary statistics, and finally, whether we fail to reject the null hypothesis or should reject. From the main findings (particularly the analysis of variance), it is evident that we fail to reject the null hypothesis because of the fact that the F-value given in the ANOVA table is less than the critical value (0.011 < 3.47).

To summarize, this means there is insufficient evidence to reject the null hypothesis, which states that the mean number of images recalled for the different musical types is the same.

Applying this to the original research question of how different music can impact studying or cognitive load, the above findings demonstrate that there is no inherent difference in participant performance.

The finding that different types of music do not significantly impact participant performance also suggests areas for further research. Future studies could explore additional factors that may

influence the relationship between music and cognitive load, such as volume or familiarity with the music genres. Additionally, different experiments can be introduced where several different factors are tested, rather than just short-term visual memory load.

Some limitations in this experiment include small sample size (≥24), the music selection (where the genre was limited to just R&B/funk), the type of test offered (memorization), and even the length of the experiment (4 rounds of 30 second exposure). In terms of the statistical analysis, the equality of variance test shows not equal which limits the ability to complete an Analysis of Variance (ANOVA). However, for the sake of this study, we checked the equality of variance and and while it led to unequal variances, ANOVA, as well as the other tests were conducted.

Additionally, one major limitation in this study involves the lack of repetition in trials, as the experiment designed for only one test per participant. This is considered a limitation because it plays a part in statistical power, in factors such as an increase in variability in the data. Given this limitation, it decreases the consistency and accuracy of the obtainted results.

5.0 References (1 page):

Slide 6:
[1] Study International (2023) *Music for homework: 4 best genres to help you study better*, *Study International*. Available at: https://studyinternational.com/news/best-music-for-homework-students/#:~:text=Around%2060%25%20of%20students%20tend,juggled%20studying%20with%20another%20task (Accessed: 11 April 2024).

Slide 7:
[2] M.-L. Amundsen, P. E. Garmannslund, and H. S. Stokke, Differences Between Visual Working Memory Among Students, https://files.eric.ed.gov/fulltext/EJ1236791.pdf.

[3] T. L. Blankenship, M. O'Neill, A. Ross, and M. A. Bell, "Working memory and recollection contribute to academic achievement," Learning and individual differences, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4669898/ (accessed Apr. 12, 2024).

Slide 8:
[4] Sridharan, D. *et al.* (2007) 'Neural dynamics of event segmentation in music: Converging evidence for dissociable ventral and dorsal networks', *Neuron*, 55(3), pp. 521–532. doi:10.1016/j.neuron.2007.07.003. Available at: https://www.sciencedirect.com/science/article/pii/S0896627307005004 (Accessed: 11 April 2024).

[5] Lehmann, J.A. and Seufert, T. (2017) 'The influence of background music on learning in the light of different theoretical perspectives and the role of working memory capacity', *Frontiers in Psychology*, 8. doi:10.3389/fpsyg.2017.01902. Available at: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2017.01902/full (Accessed: 11 April 2024).

[6] M. R. Vasilev, J. A. Kirkby, and B. Angele, “Auditory distraction during reading: A bayesian meta-analysis of a continuing controversy,” Perspectives on Psychological Science, vol. 13, no. 5, pp. 567–597, Jun. 2018. doi:10.1177/1745691617747398

Slide 9:

[7] T. L. Blankenship, M. O’Neill, A. Ross, and M. A. Bell, “Working memory and recollection contribute to academic achievement,” Learning and individual differences, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4669898/ (accessed Apr. 12, 2024).