The Impact of Classical and Heavy Metal Music on Reading Comprehension: An Experimental Analysis Across Age Groups

Ashton Chung, Rory Freck, Tiffany Hu, Vienna Truong, Yiting Wang

September 20, 2024

1. Abstract

Studies have shown that listening to classical music benefits key cognitive functions such as short-term memory, which are crucial to reading comprehension. In order to help people optimize their reading comprehension performance, this study aims to explore how listening to music beforehand would affect one's performance in reading comprehension tests among adults of different age groups. In this study, 180 Islanders from the Ironbard island were recruited and randomly assigned into either listening to classical music or heavy metal music. Among the two music type groups, we randomly assigned the participants into three groups of listening times (10/ 20/ 30 minutes). We used a two-factor Random Block Design and blocked the participants by three age groups (young adults, middle-aged and elderly) in order to reduce the variation caused by decline of cognitive functions and differences in cognitive strategies in reading comprehension across ages. After each participant gets their assigned music type and listening time groups, they receive a pre-test on reading comprehension. Then they listen to music according to the group they are assigned to and do a reading comprehension post-test. The differences between the scores of pre-tests and post-tests are used as the response factor to measure the impact of music on reading performance. In our data analysis, we used R to run an ANOVA to examine the mean differences between the factors and presented diagnostic plots to demonstrate the model's validity. The result of our study reveals that listening to classical music significantly improved reading comprehension performance compared to heavy metal, while the length of listening time had no substantial effect. We also conclude that classical music enhances reading comprehension performance more consistently than heavy metal, regardless of listening duration or age group.

2.Introduction

"The Mozart effect" has led the population to believe that listening to classical music can help one become smarter temporarily (Dewar, 2024). While previous studies have shown that there are no significant effects on general intelligence other than some short-term boosting effect on visual-temporal abilities (Rauscher et al 1993; Hetland 2000; Pietschnig et al 2010), very few studies have explored the relationship between music listening before reading and reading comprehension performance. This paper is an exploratory study aiming at answering the following question: does "the Mozart effect" also exist in reading tests among adults?

One possibility is that the characteristics of classical music allows it to stimulate specific brain regions and bring special effects on people's short-term reading performance. On the other hand, the boost in reading comprehension performance might be brought by the arousal effect of listening to any music. In that case, listening to music should have the same effect on reading comprehension performance regardless of music types. This study introduced heavy metal music in order to compare the effects brought by different music types. While classical music is proven to be beneficial for short-term cognitive functions, there are very limited studies on the psychological or neurological effects of heavy metal music with regard to reading comprehension. Comparing outcomes from classical music to heavy metal music allows us to see if the boost of reading performance by classical music is simply just an arousal effect.

We are also interested in the change of effect on reading comprehension by listening to music across different levels of listening times. It is not likely that the effect and listening time is linearly related and can grow infinitely as time increases. It is also not likely that the music would have an effect on reading comprehension performance in a very short period of time, for instance, one minute. In order to examine the amount of time needed for music to have an effect on reading performance, and the sufficient amount of time of listening to music that would optimize the effect on reading performance, we tested both music

types on 3 levels of listening times: 10 minutes, 20 minutes, and 30 minutes. The interactions between music types and listening time are also explored in this study.

However, aging is a factor that can contribute significant variations regarding reading comprehension abilities and approaches among people that this study is not interested in. Aging can bring declines in key cognitive functions such as processing speed, working memory, and attentional resources (De Beni et al., 2007) (Hannon & Daneman, 2009). Due to variations in cognitive functions, human brains apply different reading strategies across the lifespan. While younger adults apply learning-based cognitive functions during reading tasks, older adults might rely more on prior knowledge and context to aid comprehension (De Beni et al., 2007) (Schroeder, 2014). While metacognitive reading comprehension is relatively stable as age increases in a standard lab environment (Norman et al., 1992) (De Beni et al., 2003), Listening to music before doing a reading comprehension test can have different effects on different age groups. Therefore, in order to provide a more accurate explanation of variations caused by music listening, this study blocked participants with three different age groups: young adults (age = 18-29), middle-aged (age = 30-59) and elderly (age >=60).

With the block factor of age groups, the study will compare the effects of music type and music length within groups that share similar brain function levels and cognitive reading strategies, and is expected to provide a more accurate prediction of the effect of music on each individual. We are also interested in exploring whether reading comprehension of different age groups would react to music stimulations differently due to different cognitive strategies applied, or would the cognitive process be affected by music regardless of function types.

We hope to contribute to expanding the broader understanding of the interaction between cognitive performance and environmental factors like music. Meanwhile, results from this study can provide some guidelines for optimizing reading comprehension performance in adults preparing for tests, and lead the way for further studies with regard to topics such as possible long-term effects of pre-reading habits on reading comprehension and the impact of music on brain functioning with regard to reading comprehension on a neurological level.

3. Method

3.1 Participants

Our team used an online simulated world called the 'Island' to recruit willing participants for our experiment. We recruited subjects from the region of Ironbard to mitigate any cultural and educational influences that could contribute to unwanted variation. Our research team, composed of five people, found 36 willing participants each to participate in the study, making our total sample size 180 subjects.

The participants were selected based on their age to maintain an equal representation for all age groups. In our experiment, we created three age groups: young adults (18-29 years), middle-aged (30-59 years), and elderly (over 60 years). This stratification based on age was used to explain some variability in age-related cognitive responses. For instance, we speculated that younger adults might have sharper cognitive skills than older adults due to the natural declines in processing speed as we age.

3.2 Design

Our study used the Random Complete Block Design (RCBD) with two factors and a pretest-posttest structure. We used RCBD as our experimental design because it allowed us to control for unwanted variation in the response variable. The nuisance variable results in variation of the response variable that we are not interested in, and to fix this, we employ a block to control for the nuisance. The two design factors we are concerned with are music type (classical or heavy metal) and the duration of listening to music (10 minutes, 20 minutes, and 30 minutes). We will administer a pre-treatment test and a post-treatment test. The difference between the posttest and the pretest is the response variable of this design. Lastly, the blocking factor in this design is age group. We are not interested in the effect that age has on the response variable, so we will block on age.

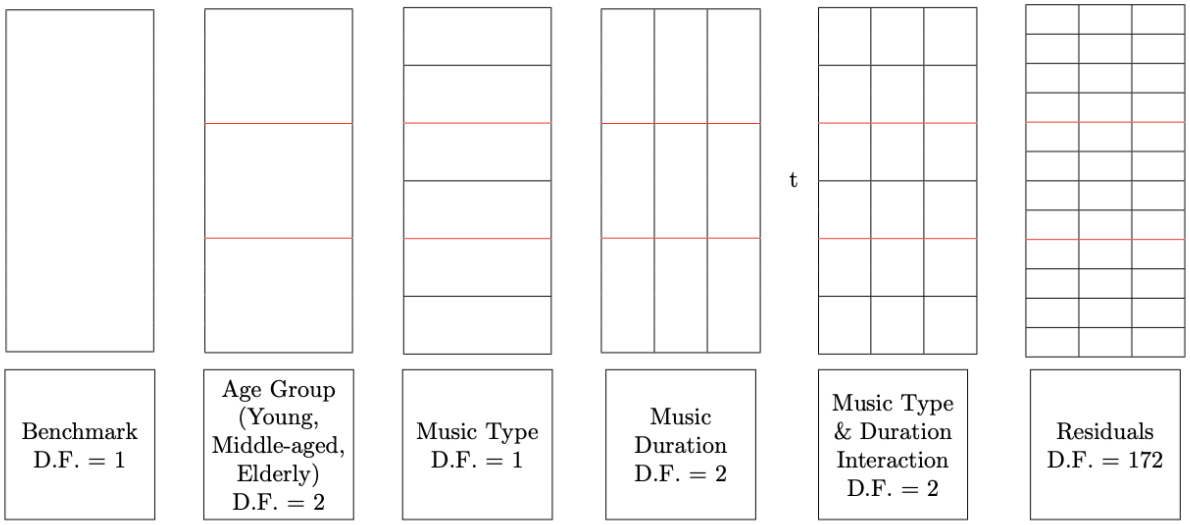

Figure 1: Factor Diagram

3.3 Material and Procedure

Our team wanted to establish a baseline measurement for each participant, so we administered the reading comprehension test before imposing any treatments. The test takes 10 minutes to complete. After the initial reading comprehension test, we blocked participants on age.

After blocking the participants on age, we assigned them a unique ID from 1 to 60 within their respective age block. We used a two-step pseudo-randomization process to ensure that the assignment of both treatments was evenly distributed. This process involved first setting a seed of 100 in R for repeatability. Then, we created vectors that represented the treatment types. For example, classical music was assigned '1', and heavy metal was assigned '2.' For duration, 10 minutes was '1', 20 minutes was '2', and 30 minutes was '3'. Each vector contained an equal number of identifiers for each level of the treatment within each block to maintain a balanced design. The treatment vectors were then pseudo-randomized for each participant using the sample() function.

Next, the participants were instructed to listen to the assigned music type for their designated duration. We made sure to leave no gaps between the listening session and the post-treatment reading comprehension test to avoid any dilution in our findings due to the time gaps between treatment and test.

Our procedure is as follows:

1. We identified willing participants on the Island in the region of Ironbard across all cities and recruited the same number of participants per age group to ensure equal representation.
2. We sorted the participants into age categories and used R to assign each participant to both pseudo-randomly assigned treatments.
3. We tested the participants' reading comprehension before imposing treatments to establish a baseline to help us measure the response variable.
4. We gave each participant the assigned music treatment according to the random assignment in each age group, maintaining a balanced design.
5. After music exposure, we again administered the reading comprehension test to measure its effects on the participant's scores.
6. We calculated the difference between the second and first tests to establish the response variable.

3.4 Instruments

Our study utilized the simulated world called the 'Island' to administer the reading comprehension test and the treatments to the participants. The 'Island' was equipped with tools such as listening sessions with various music types and different cognitive tests. We honed our study on classical music and heavy metal and used the standardized reading comprehension test to accurately measure the participants' reading comprehension.

One challenge we faced was the limited options regarding the listening duration. The 'Island' allowed us to play the music for 10 minutes only, so for participants who had been assigned 20 minutes or 30 minutes, we had to manually intervene and start the listening session again once it was over. This ensured that no breaks occurred during the listening duration, allowing for continuous exposure to the music as intended in the study's design.

4. Data Analysis

4.1 Statistical Analysis

Data preparation was necessary before we could begin analyzing it. This involved calculating the difference between the second reading comprehension score and the first and establishing the difference as the response variable. Next, we made a new field for age that classified the participants as young, middle-aged, or elderly.

We ran an ANOVA test with the aov() function. We used F-tests carried out by the summary of the ANOVA to determine if there are significant differences in the change of reading comprehension scores when listening to classical music versus heavy metal or listening to music for 10 minutes, 20 minutes, or 30 minutes. Lastly, we displayed our findings in plots considering the two factors and their interaction.

Next, we ran diagnostic tests to ensure the model abided by the necessary error assumptions of linearity, normality, homoscedasticity, and the absence of influential outliers. Lastly, we displayed our findings in plots considering the two factors and their interaction.

4.2 Sample Size Determination

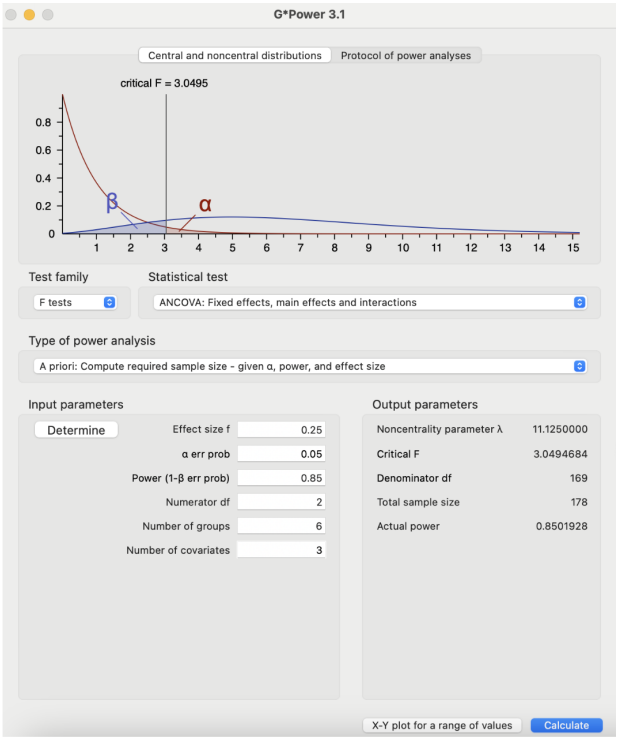

Figure 2: Sample Size using G*Power

We aimed for a power of 0.85 in our study, meaning we have an 85% chance of correctly rejecting a false null hypothesis. We set our significance level (alpha) at 0.05, giving us a 5% chance of mistakenly rejecting the null hypothesis when it's true. We chose a medium effect size of 0.25, representing the expected difference between our groups.

Our experiment follows a two-way randomized complete block design. Using G*Power, we determined that we would need 178 participants. To keep the study design balanced and ensure equal representation across our six groups (2 music types $\times$ 3 durations), we increased the sample size to 180 participants. This adjustment ensures a balanced design.

5. Results

5.1 ANOVA Analysis

Table 1: ANOVA Table: Effects of Music Length, Type, and Age Group

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Music Length	2	133	67	0.567	0.568
Music Type	1	23598	23598	200.506	<2e-16
Age Group	2	245	122	1.039	0.356
Music Length:Music Type	2	520	260	2.21	0.113
Residuals	172	20243	118		

Figure 3: ANOVA Table

The ANOVA results for the experiment investigating the effect of Music Length and Music Type on the Test Score Differences (with Age Group as a blocking factor) provide a detailed overview of the influence of these factors on reading comprehension performance.

The ANOVA table shows that the Music Type has a significant effect on the response variable (Test Score Difference), with a p-value of less than 2e-16. This suggests that there is a statistically significant difference in reading comprehension performance depending on whether participants listened to Classical or Heavy Metal music before their tests. Specifically, the low p-value indicates a very strong association between the type of music and reading performance, providing evidence that the choice of music influences cognitive performance during the test.

The Music Length variable itself, however, does not have a statistically significant effect on the response variable, with a p-value of 0.568. This suggests that the duration of music listening—whether it was for 10, 20, or 30 minutes—did not make a substantial difference in reading performance across the sample population. While the duration of listening was not significant, this could also imply that music's effect might be more related to the type rather than the length of exposure.

The interaction term between Music Length and Music Type has a p-value of 0.113, which is above the traditional significance threshold ($p < 0.05$), suggesting that there is no significant interaction between the

two factors. This means that the impact of Music Length on reading comprehension did not vary significantly based on the Music Type. For example, increasing the duration of listening to music had a similar impact on reading performance for both Classical and Heavy Metal music.

Interestingly, the Age Group block did not produce a significant result (p-value = 0.356), which indicates that controlling for age group did not explain much of the variation in test score differences. This result suggests that the effect of listening to music on reading comprehension is relatively consistent across different age groups. In other words, younger, middle-aged, and older adults did not exhibit notably different responses to the music conditions in terms of their reading comprehension.

5.2 Post-Hoc Analysis

To further explore the differences between the groups, a post-hoc Tukey's test was conducted, allowing us to assess pairwise comparisons between the groups for both Music Type and Music Length.

Table 2: Post-Hoc Analysis of Differences

Comparison	Diff	Lwr	Upr	p adj
2-1	-0.8833	-5.5661	3.7998	0.8962
3-1	-2.1000	-6.7827	2.5827	0.5401
3-2	-1.2167	-5.8994	3.4661	0.8125
HeavyMetal-Classical	-22.9000	-26.0922	-19.7078	0.0000
middle-aged-elderly	2.7044	-1.9946	7.4035	0.3639
young adults-elderly	0.6566	-4.1314	5.4445	0.9437
young adults-middle-aged	-2.0479	-6.6396	2.5439	0.5437
2:Classical-1:Classical	1.8785	-6.1946	9.9517	0.9849
3:Classical-1:Classical	1.9785	-6.0946	10.0517	0.9810
1:HeavyMetal-1:Classical	-18.3397	-26.4129	-10.2666	0.0000
2:HeavyMetal-1:Classical	-21.9849	-30.0581	-13.9118	0.0000
3:HeavyMetal-1:Classical	-24.5183	-32.5914	-16.4451	0.0000
3:Classical-2:Classical	0.1000	-7.9732	8.1732	1.0000
1:HeavyMetal-2:Classical	-20.2183	-28.2914	-12.1451	0.0000
2:HeavyMetal-2:Classical	-23.8635	-31.9366	-15.7903	0.0000
3:HeavyMetal-2:Classical	-26.3968	-34.4700	-18.3236	0.0000
1:HeavyMetal-3:Classical	-20.3183	-28.3914	-12.2451	0.0000
2:HeavyMetal-3:Classical	-23.9635	-32.0366	-15.8903	0.0000
3:HeavyMetal-3:Classical	-26.4968	-34.5700	-18.4236	0.0000
2:HeavyMetal-1:HeavyMetal	-3.6452	-11.7184	4.4280	0.7840
3:HeavyMetal-1:HeavyMetal	-6.1785	-14.2517	1.8946	0.2405
3:HeavyMetal-2:HeavyMetal	-2.5333	-10.6065	5.5398	0.9449

Figure 4: TukeyHSD Post-Hoc Analysis

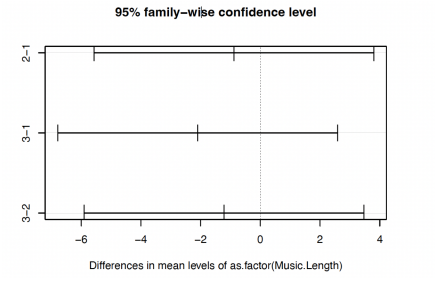

Figure 5: Pairwise Comparison 1

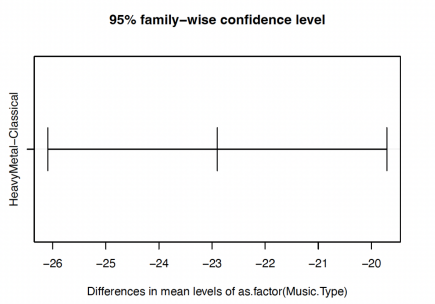

Figure 6: Pairwise Comparison 2

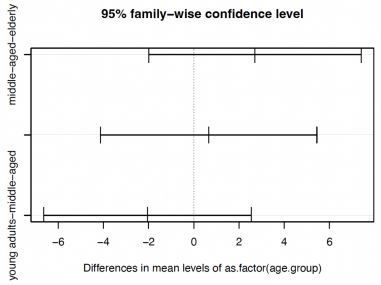

Figure 7: Pairwise Comparison 3

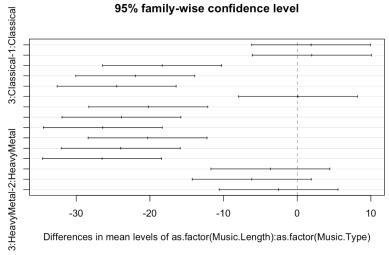

Figure 8: Pairwise Comparison 4

The Tukey's HSD results show a significant difference in Test Score Differences between Classical and Heavy Metal music across all levels of Music Length ($p < 0.0001$). This confirms that participants who listened to Classical music showed a significantly different (and likely better) performance in reading comprehension than those who listened to Heavy Metal. The post-hoc comparisons strengthen the conclusion that music type has a robust effect on test performance, favoring Classical music over Heavy Metal.

However, the pairwise comparisons between different levels of Music Length (10, 20, 30 minutes) did not reveal any significant differences (p-values above 0.94). This reinforces the earlier finding from the ANOVA that the duration of music exposure did not significantly affect reading comprehension.

5.3 Residual Diagnostics Plots

The residual diagnostic plots were used to evaluate whether the assumptions of ANOVA were met:

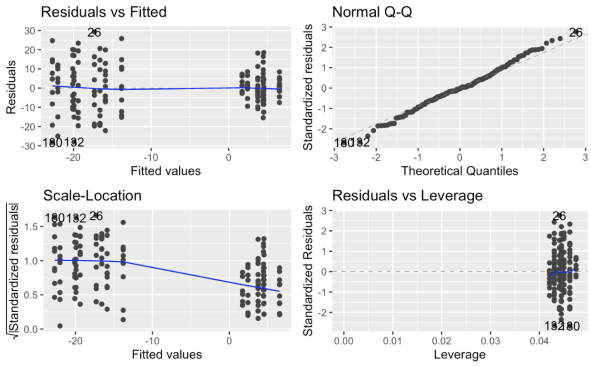

Figure 9: Residuals Analysis

1. Residuals vs Fitted: This plot showed that the residuals were fairly evenly distributed around the fitted values. Although the distribution of the points seems to be in 2 distinct clusters, this is likely due to the strong association between the types of music the subjects listened to with the reading comprehension performance. Since the points are still evenly distributed along the two sides of the line in both Residuals vs Fitted and Scale-Location plot, this suggests that the assumption of linearity holds, meaning the relationship between the predictor variables and the response is appropriately modeled.

2. Normal Q-Q Plot: The points in the Q-Q plot mostly followed the 45-degree line, indicating that the residuals are approximately normally distributed. This supports the assumption of normality for ANOVA residuals.

3. Scale-Location: The Scale-Location plot demonstrated that the variance of the residuals is generally constant across the range of fitted values. Although there seems to be a slight downward change in the trend from the left to the right cluster, the fluctuation is within an acceptable range, and the overall plot affirms the homoscedasticity (constant variance) assumption.

4. Residuals vs Leverage: The Residuals vs Leverage plot did not highlight any particularly influential points that would distort the model, meaning that no single observation is unduly influencing the results.

5.4 Interaction Plot

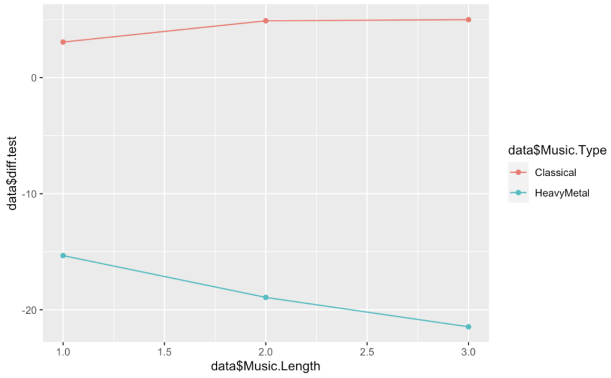

Figure 10: Interaction Plot

The interaction plot illustrates the relationship between Music Length, Music Type, and the Test Score Difference. The plot reveals no significant interaction between Music Length and Music Type, as the lines for Classical and Heavy Metal music remain fairly parallel. This lack of interaction implies that the impact of music type on reading performance is consistent across all durations of music exposure, rather than varying depending on how long participants listened to the music.

5.5 Boxplot Comparisons

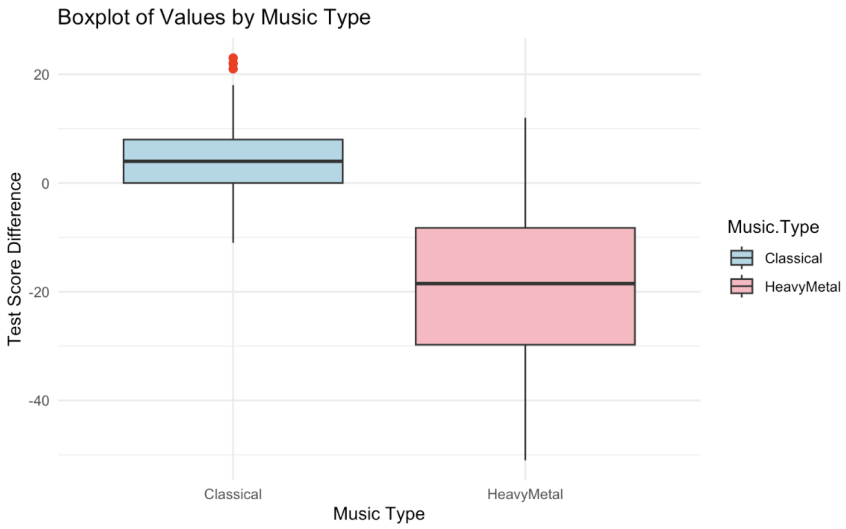

Figure 11: Boxplot of Values by Music Type

The boxplot comparing Test Score Differences across Music Types highlights that participants in the Classical music group generally had a higher median score compared to those in the Heavy Metal group. The boxes for Classical music are narrower, suggesting less variability in test score differences compared to Heavy Metal, where the spread of the data is more pronounced, indicating more variability in how participants responded to the heavy metal treatment.

!

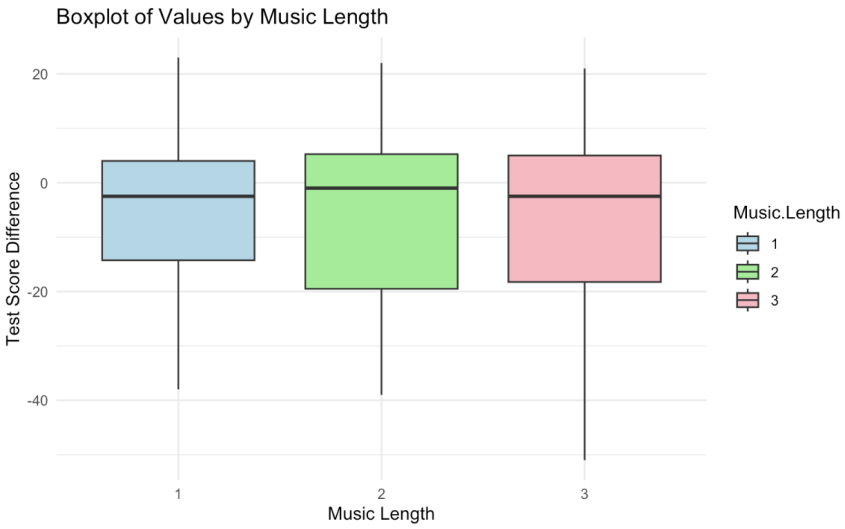

Figure 12: Boxplot of Values by Music Type

The side-by-side boxplot for Music Length confirms that there is no strong pattern or trend across the different durations (10, 20, or 30 minutes). The median values for each boxplot are similar, and there is substantial overlap between the interquartile ranges of each music length. This further supports the finding that music length, on its own, did not have a significant effect on reading comprehension performance.

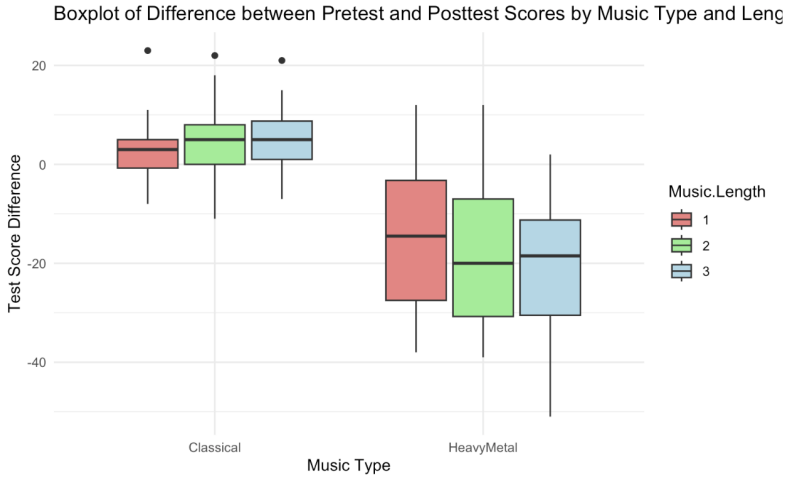

Figure 13: Boxplot of Values by Music Type

The side-by-side boxplot for Music Type and Music Length illustrates the difference in Test Score Differences between participants who listened to Classical or Heavy Metal music, with further grouping by the three levels of Music Length (10, 20, or 30 minutes).

For participants in the Classical music group, the median test score difference is generally higher across all levels of music length, with relatively narrow interquartile ranges (IQRs), indicating less variability in responses. This suggests that participants listening to classical music had more consistent performance improvements in reading comprehension compared to those who listened to heavy metal. The presence of a few outliers (denoted by black dots) suggests that some participants had test score differences that were unusually high or low compared to the rest of the group.

In contrast, participants in the Heavy Metal group exhibit a wider spread in test score differences, as shown by the larger IQRs, especially for shorter music durations. The Heavy Metal group's median scores are generally lower, with more variability, indicating that this group's performance was less consistent. The spread of the boxes suggests that while some participants improved their scores after listening to heavy metal, others saw little or no improvement.

Overall, this boxplot indicates that Music Type may have a significant effect on reading comprehension performance, with Classical music leading to more consistent and possibly higher improvements compared to Heavy Metal. However, there does not appear to be a clear trend regarding Music Length, as the differences between the three lengths are relatively small within each music type.

6. Discussion

Our study's objective was to determine the effect of listening to different types of music and listening time on reading comprehension among adults of varying age groups. The results determined that music type was the only significant factor affecting reading comprehension while listening length and the blocking factor of age were not important. Additionally, there were no significant interactions.

Our data revealed that music type causes a significant difference in reading comprehension scores. As shown in the boxplots, those who listened to classical music had a positive score difference (indicating that their scores increased), whereas heavy metal listeners' scores decreased. The Tukey HSD confidence interval determines a range of considerable differences among music types and further demonstrates their statistical difference. This is in contrast to previous studies which showed no significant effect of classical music on general intelligence (Rauscher et al 1993; Hetland 2000; Pietschnig et al 2010).

Delving further into the boxplots reveals that the factor of listening length did not significantly impact the difference in reading scores. Variability within music-type groups appears relatively similar, however, there is no significant difference between listening to music for 10, 20, or 30 minutes before the second reading comprehension test. Some possible reasons that duration may have not played a role include listeners' quick adaptability to the pre-test music or that the times had no incremental impact.

The interaction plot shows the relationship between music length and listening time with heavy metal music having more diminishing changes in test score differences as the time increases. There is no interaction

between music length and listening time. The ceiling effect may be a likely limiter to interaction between music type and listening time. The participants may have reached their maximum performance in reading comprehension early in the study, so there would be little room for additional influences from the listening time or music type.

Furthermore, the blocking factor of the age group did not significantly impact reading comprehension when listening to music. A potential explanation for insignificance may be similar adaptability to distractions across age groups. While age group may affect general cognitive functions, previous studies show stable reading comprehension in increasing ages in lab environments supporting our findings (De Beni et al., 2007; Schroeder, 2014).

Our design resulted in significant results for music types' effect on reading comprehension, however, we acknowledge the limitations of our study. Since sampling was limited to the region of Ironbard, this sample may not be fully representative of the entire island. Additionally, we did not collect educational background for the subjects. Misrepresented Educational level would bring different outcomes in the relationship between music listening and reading comprehension. For example, receiving consistent training in reading comprehension at school might strengthen one's ability to focus on reading tasks after other activities. Similarly, individual differences were also not accounted for, as some people had little to no difference in test scores based on treatment. Some people may have preferred one music type over the other which would affect their scores. We also limited music types to only classical and heavy metal. Additionally, we would have preferred to administer music during the reading comprehension test to determine the effect while testing. However, the listening and comprehension tasks must be done one at a time. This may have limited the effect of the music on the scores.

For future research directions, we would further study the topic by introducing more types of music and more distinct characteristics within music, like tempo or lyrics. We would also like to include a new blocking factor of either pre-test score or educational background. Additionally, we would account for individual differences in music preferences and learning styles with pre-screening surveys. In our original experiment, we had to administer the music listening tasks and reading comprehension one at a time. In further research, we would like to administer music during the reading comprehension test to determine the effect of music while testing. While listening time was not significant, we would still like to include music exposure in the study by changing the time frame to a more longitudinal study. Subjects would listen to music for a set amount of time per week to investigate cumulative music exposure. Overall, the findings from this research could benefit educational settings, study habits, and workplace productivity with music to enhance cognitive performance.

7. Conclusion

The ANOVA results provide compelling evidence that Music Type plays a critical role in influencing Test Score Differences between the pre-test and post-test reading comprehension scores. Participants exposed to Classical music performed significantly better than those exposed to Heavy Metal music, regardless of the length of exposure. While Music Length did not appear to have a significant effect, the interaction between Music Length and Music Type also showed no substantial influence. The block factor of Age Group did not significantly affect the results, suggesting that music's impact on reading comprehension is consistent across different age groups. Post-hoc analysis reinforced these findings, confirming the robust effect of Music Type on cognitive performance. Overall, the study provides new insights into how music influences reading comprehension, especially highlighting the benefits of classical music in enhancing cognitive function in reading tasks.

8. References

1) De Beni, R., Palladino, P., Borella, E. et al. 2003.Reading comprehension and aging: Does an age-related difference necessarily mean impairment?. Aging Clin Exp Res 15, 67–76. https://doi.org/10.1007/BF03324482

2) Suzanne Norman, Susan Kemper, Donna Kynette, Adults' Reading Comprehension: Effects of Syntactic Complexity and Working Memory, Journal of Gerontology, Volume 47, Issue 4, July 1992, Pages P258–P265, https://doi.org/10.1093/geronj/47.4.P258

3) Schroeder, P. J. 2014. The Effects of Age on Processing and Storage in Working Memory Span Tasks and Reading Comprehension. Experimental Aging Research, 40(3), 308–331. https://doi.org/10.1080/0361073X.2014.896666

4) De Beni, R., Borella, E., & Carretti, B. 2007. Reading Comprehension in Aging: The Role of Working Memory and Metacomprehension. Aging, Neuropsychology, and Cognition, 14(2), 189–212. https://doi.org/10.1080/13825580500229213

5) Hannon, B., & Daneman, M. 2009. Age-Related Changes in Reading Comprehension: An Individual-Differences Perspective. Experimental Aging Research, 35(4), 432–456. https://doi.org/10.1080/03610730903175808

6) Dewar. "The Mozart effect: The truth behind the claims." 2024. https://parentingscience.com/mozart-effect/. Accessed 16th Sept, 2024.

7) Rauscher, F., Shaw, G. & Ky, C. 1993. Music and spatial task performance. Nature 365, 611. https://doi.org/10.1038/365611a0

8) Hetland, L. 2000. Listening to Music Enhances Spatial-Temporal Reasoning: Evidence for the "Mozart Effect." Journal of Aesthetic Education, 34(3/4), 105–148. https://doi.org/10.2307/3333640

9) Pietschnig J, Voracek M, and Formann AK. 2010. Mozart effect–Shmozart effect: A meta-analysis. Intelligence 38(3): 314-323.https://doi.org/10.1016/j.intell.2010.03.001