Determining the Effects of Music Genres on Short Term Memory in Elderly Islanders

Johan Chua, Ryan Largo, Masato Ishizuka

September 21, 2023

Contents

1 Abstract

In recent years, particular attention has been paid to research related to music and memory-related illnesses common among the elderly, such as Alzheimer's disease. Accordingly, studies have long been conducted in the United States to examine how memory retention, particularly in the elderly, is altered by various genres of music. We believe that if certain genres of music can be shown to help older adults retain or improve their memory, it could be effective in preventing the progression or onset of conditions such as the aforementioned Alzheimer's disease. Our goal is to determine whether listening to specific music is an effective modulator of memory retention in the elderly. Our study includes a representative sample of islanders that was analyzed using a two-way randomized block design. Two-way analysis of variance with blocking was used to analyze the results. These results suggest that specific genres of music have significant effects on memory ability, and thus may be a potential treatment to control memory decline.

2 Introduction

Memory is a fundamental cognitive process of humans that plays a crucial role in shaping our worldview and ensuring quality of life. The human memory is a rather complex and dynamic process, involving multiple levels of processing and functions—from short-term to long-term to visual and sensory. Memory impairment is becoming an ever-increasing concern that can significantly hinder daily function, especially for the elderly (Ridel et. al). In recent research, music has proven to be a promising method of improving memory and cognitive ability; however, a majority of research has been only done on the Classical genre (Mammarella et. al). Thus, this paper aims to extend this ongoing investigation of music and memory by studying how various music genres affect short-term memory in elderly individuals.

Our study focuses on people 65 years of age and older because numerous studies with older adults have demonstrated the potential benefits of music, suggesting that the cognitive abilities of the elderly are particularly sensitive to situational factors (Ward et. al). Therefore, if music can create situations in which older adults' performance is optimized, this should have important practical implications for aging well. In particular, a study found that listening to unfamiliar classical music increased the number of words recalled compared to the no music condition (Proverbio et. al and Bugter et.al). This raised the question of whether music other than classical music would have a similar positive effect on memory cognition. We therefore wanted to extend this to listening to dance and heavy metal music for 0-20 minutes, and to perform vocabulary tests before and after to see if there were any negative or positive effects on cognitive memory associated with the different genres.

Because we believe that music could be a preventative measure of memory loss in elderly as opposed to a cure, we chose to block by gender and avoid the dementia population altogether. One of the best-known genes that affect the risk of Alzheimer's disease is the apolipoprotein E (APOE) gene, which is involved in making a protein that helps transport cholesterol and other types of fats into the bloodstream (Loprinzi et. al). Problems with this process are thought to lead to the development of Alzheimer's disease; there are several forms of APOE, called alleles (epsilon2, epsilon3, etc). Women tend to have greater memory impairment than men, possibly due to the presence of the APOE-4 allele (i.e., possession of at least one APOE-4 allele), which leads to a faster rate of atrophy in the whole brain and temporal lobes (Loprinzi et. al). However, after age 65, the incidence of episodic memory loss in men tends to be higher than in women (Theofilidis et. al). Although there are a variety of studies on these gender-based factors, we did not expect them to have much effect on short-term memory, which is what we are examining in this study.

3 Methods

3.1 Design

The participants studied were 198 islanders aged 65+ from the islands of Ironbard, Providence, and Bonne Santé. Our sampling population was completely random, as we randomly selected sample observations from all islands and cities over a period of two weeks. We adopted a balanced design and collected data from both male and female elderly. Additionally, we ensured that none of our subjects had dementia in accordance with our research goals.

Response Variable	Memory Test Vocabulary		
Treatment 1 (Music Genre)	Classical	Heavy Metal	Dance
Treatment 2 (Listening Duration)	0 Listens	1 Listen	2 Listens
Block (Gender M/F)	Male	Female	

Figure 1: Design Table

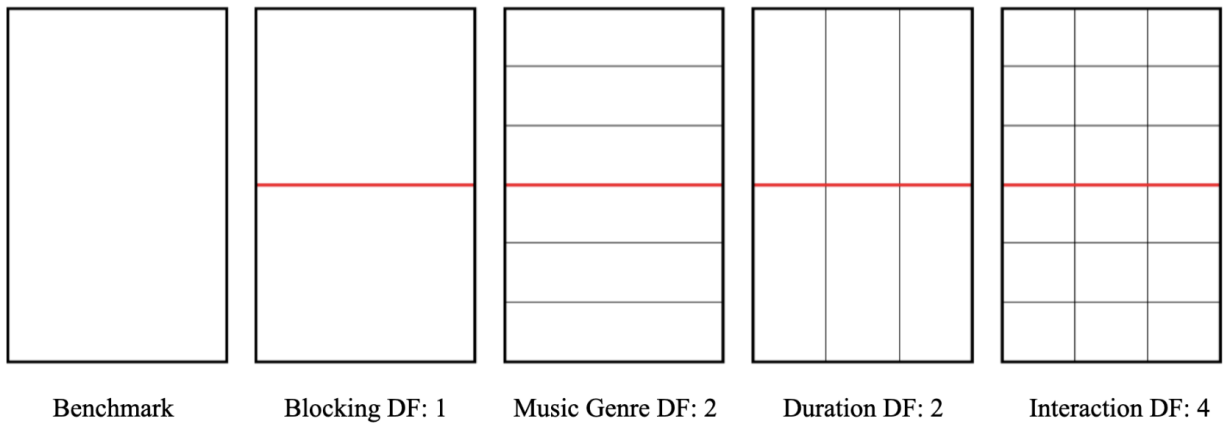


Figure 2: Factor Diagram

Our response variable was the before and after difference of the Island's Memory Test Vocabulary score. It is a score out of 20 marks that tests the ability of an individual to recall a list of 20 given words in 30 seconds, after waiting a minute from the time the words were shown. Our two design factors are music genre (a genre from Classical, Dance and Heavy Metal) and listening duration (measured by the number of times the music was played, with each play lasting 10 minutes). Our blocking factor was gender, which was split between male and female.

3.2 Instruments

The "Memory Test Vocabulary" task on the Island has each observation:

1) Look at a list of 20 words for one minute
2) Waiting for one minute
3) Seeing how many words they can recall in 30 seconds

Using the memory vocabulary test, we are able to get a score out of 20 which we then are able to determine if there is change depending on the music genre and the listening duration our observations are randomly assigned to.

3.3 Procedure

Step 1: Randomly select consented participants, specifically age 65+, willing to take part in our experiment. We would need an equal amount of observations for each of our treatment groups and blocking combinations.

Step 2: Randomly assign the groups into their designated treatment groups

1) Zero listens to Classical music, One listen to Classical music, Two listens to Classical music
2) Zero listens to Heavy Metal music, One listen to Heavy Metal music, Two listens to Heavy Metal music
3) Zero listens to Dance music, One listen to Dance music, Two listens to Dance music

Step 3: For each islander studied, measure their memory vocabulary test out of 20 prior to any treatments

Step 4: For each islander, apply the treatments of having them listen to Classical, Heavy Metal, or Dance music, and check their memory vocabulary test score after the required listening durations the islanders have been specifically assigned to.

Step 5: Measure the memory vocabulary test again after the correct amount of listening durations have occurred.

4 Data Analysis

4.1 Method of Data Analysis

We will use Analysis of Variance (ANOVA) to statistically analyze our data. In order to determine if the genre of music and its dosage have an effect on vocabulary memory test scores, we will use the F-test to identify if there is a statistically significant difference between scores for each treatment group. Through F-testing, we will conclude whether or not the genre of music, the duration of music and their interaction have any effect on memory test scores. For each variable, we will see whether or not the variation between treatment groups is greater than the variation within groups, which indicates that one of the group means is significantly different from the others. As part of the study, we also have control groups in each genre of music (duration is 0 listens) that we will compare our treatment groups to.

4.2 Sample Size Determination

We will use the software GPower to determine our minimum required sample size. In statistical experiments, it is a generally accepted practice to have a power of at least 0.8 (ensuring a 80% or greater probability of correctly rejecting a false null hypothesis). We also used a 0.5 significance level (probability of incorrectly rejecting a null hypothesis when it's true) to keep Type I error to a minimum, and used the standard 0.25 effect size (measures the intensity of the difference in groups). With this information, G*Power has determined that our Two-Way Randomized Complete Block Design will require a minimum sample size of 197 to achieve at least 0.8 power. However, since our design is balanced and we have 18 treatment groups, we increase the sample size to 198 to achieve a balanced 11 samples per treatment group ($197/18 = 10.94$, rounded up to 11).

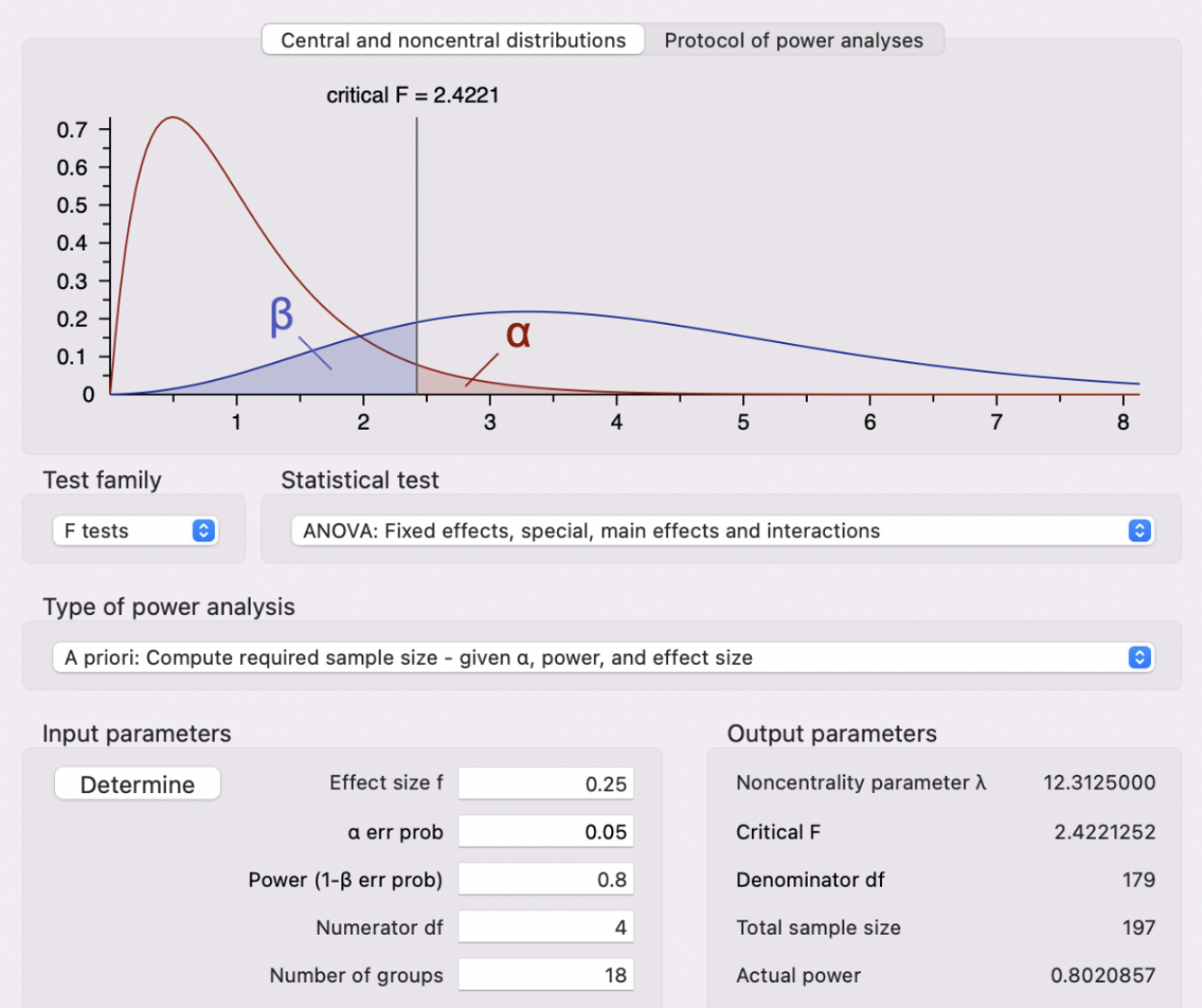
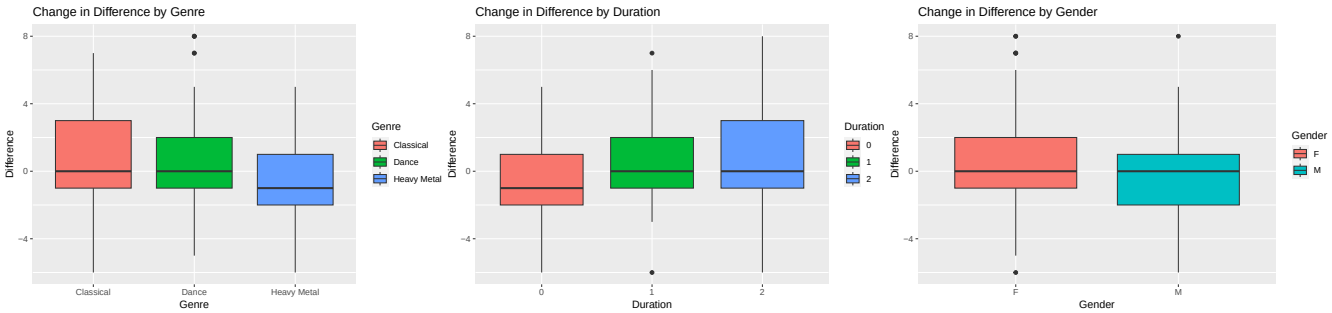


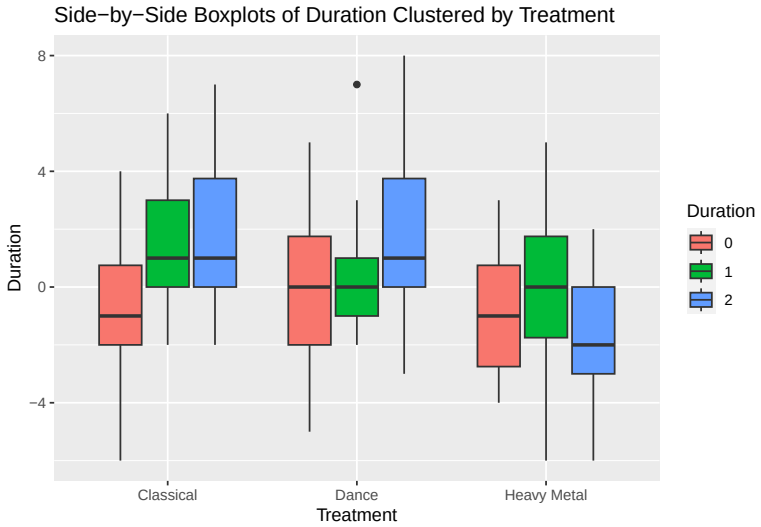
Figure 3: Sample Size Determination

5 Results

5.1 Box Plots

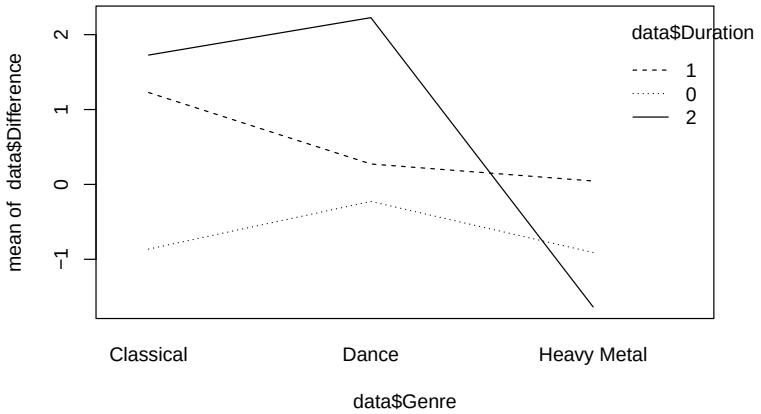


Preliminary data analysis through boxplots reveal multiple characteristics of our data. Firstly, the Genre boxplots show that the Classical and Dance genres have similar effects with regards to changes in memory test score, but Classical seems to have a greater positive effect. Heavy Metal appears to hurt test scores more than it improves them, with a negative median. Secondly, the Duration boxplot shows that 1 and 2 listens have similar effects, while 0 listens has the worst effect. Also interesting to note is that the medians of both 1 and 2 listens are 0. Laslty, from the Gender boxplot, we see that music has a greater effect on Females than Men in terms of memory score.



Analyzing the duration and treatment boxplots to gauge the interaction effect, we find that there is a positive trend between memory score improvements and listening durations for Classical and Dance music, but a reverse effect for Heavy Metal. It is interesting how 2 listens of Heavy Metal deviates significantly from the overall trend, which is something worth exploring furher in our ANOVA/Tukey analysis.

5.2 Interaction Plot



The interaction plot indicates that different durations follow different trends, suggesting that the interaction effect is signficant and should be accounted for in the model. Specifically, there is minimal interaction observed among the various genres when listeners engaged in a single session, as compared to the control group, which abstained from music altogether. However, an interaction becomes evident, specifically within the Heavy Metal category, where the score declines as the listening duration is extended to twice the original duration.

5.3 ANOVA Table

Df	Sum Sq	Mean Sq	F Value	Pr > F
2	107.28283	53.641414	8.696486	0.0002442
2	77.76768	38.883838	6.303949	0.0022394
1	11.15657	11.156566	1.808732	0.1802807
4	110.68687	27.671717	4.486211	0.0017442
188	1159.61616	6.168171	NA	NA

With p-values of 0.0002 and 0.0022 (less than 0.05) for the variables Genre and Duration respectively, we conclude that both variables are statistically significant and have an effect on "Memory Test Vocabulary" scores. The large F-ratios and low p-values for both variables suggest that the variation between groups for both variables are greater than the variation within groups, and that at least one of the treatment group means for both variables are different from the others. Additionally, the interaction between our two treatment variables Genre and Duration are also significant, with a low p-value of 0.0017 (less than 0.05), suggesting that the interaction effect between Genre and Duration has a statistically significant effect on the change in memory test scores. However, the blocking factor proved to be insignificant. With a high p-value of 0.1802 (greater than 0.05), we conclude that the variation between groups is less than the variation within groups and that Gender does not have a significant effect on changes in memory test scores. Contrary to literature, our blocking was ineffective.

5.4 Tukey Analysis

Table 1: Genre

	diff	lwr	upr	p adj
Dance-Classical	0.0606061	-0.9607554	1.0819675	0.9892253
Heavy Metal-Classical	-1.5303030	-2.5516645	-0.5089416	0.0014598
Heavy Metal-Dance	-1.5909091	-2.6122705	-0.5695477	0.0008861

Table 2: Duration

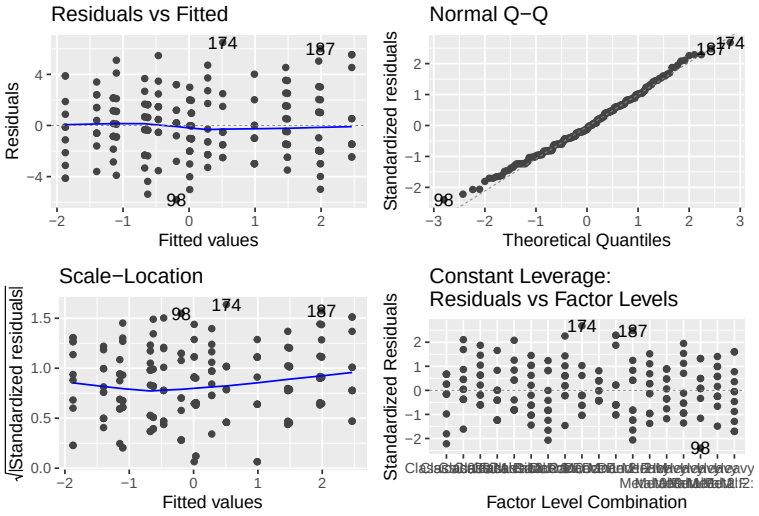
	diff	lwr	upr	p adj
1-0	1.1818182	0.1604567	2.203180	0.0187362
2-0	1.4393939	0.4180325	2.460755	0.0029960
2-1	0.2575758	-0.7637857	1.278937	0.8225518

Table 3: Gender

	diff	lwr	upr	p adj
M-F	-0.4747475	-1.171099	0.2216039	0.1802807

Using the Tukey Multiple Comparison Method to determine the specific differences between treatment groups reveals interesting nuances between our Genre and Duration treatment groups. Specifically, we find that Heavy Metal music is statistically different from both Classical (p-value of 0.0015) and Dance (p-value of 0.0009) music. However, the difference in memory test scores from Dance and Classical music are not (p-value 0.99). For the Duration variable, we see a significant difference between our control group 0 plays and 1 plays (p-value of 0.0187) and between 0 plays and 2 plays (p-value of 0.0030); but, there is no statistically significant difference between 1 and 2 plays (p-value of 0.8226). This suggests that, as far as duration is concerned, the number of plays doesn't matter—the crucial aspect is if the subjects listened to music at all.

5.5 Diagnostic Plots



From observing our four plots, “Residuals vs Fitted”, “Normal Q-Q”, “Scale-Location”, and “Residuals vs Leverage”, we can see that our plots are generally flat and residuals follow a normal distribution, as the points on the plot are dispersed relatively evenly along the straight line. The spread of data does not seem to lack constant variance and the lines appear to follow homoscedasticity with minimal outliers along the tail ends.

6 Discussion

The purpose of this paper was to investigate the effect of different genres of music on the short-term memory of individuals 65 years and older. In doing so, we hope to determine whether or not listening to music is a viable option of memory enhancement for the elderly in the short-term, and if pursuing further research on music's effect on long-term memory would be worthwhile.

In order to achieve a minimum power of 0.8, we determined that we needed a minimum sample size of 197. We used a sample size of 198 to ensure a balanced design of 11 observations per treatment group, of which there were 18.

The results from our ANOVA analysis indicate that some factors in our model are more significant than others. Our two design factors, music genre and listening duration, were both statistically significant. This suggests that each factor alone has an effect on the change in memory score.

Specifically, looking at the Tukey analysis, we see that the two treatment level pairings under the Genre factor that have statistically significant differences are Heavy Metal-Classical and Heavy Metal-Dance. Thus, from the negative differences, we find that Heavy Metal, relative to the other two genres, is the least effective genre in improving memory test scores. Also, the Classical-Dance pairing was not significant, indicating that these two genres are equally as effective in changing memory test scores.

Analyzing our Duration treatment factor, we find that significant differences only arise between the 1-0 and 2-0 pairings. This implies that changes in memory test scores are only truly contingent on whether the individual listens to the music or not. Since the 0 duration acted as our control group, this confirms our hypothesis that listening to music in general does have an effect on short term memory abilities. Therefore, in terms of key takeaways from the Duration treatment factor, it is important to note that—in the specific context of this experiment, with 0 listens being a treatment level option—the listening duration is not a very important variable, despite it having a statistically significant p-value on the ANOVA table.

Returning to the ANOVA table, we also see that the Genre:Duration interaction factor had a significant effect on our response variable. Therefore, our ANOVA testing concludes that there are statistically significant differences in the vocabulary memory test scores of elderly Islanders based on the different treatment combinations (of both type and dosage levels) that they were administered. Lastly, the blocking factor of Gender was ultimately deemed not significant. We chose to block by gender to account for the scientifically proven differences in memory impairment between females and males, as well as differences in autobiographical memory (females outperform males on random word recall tasks). However, the results of the ANOVA reveal that these concerns did not have a significant effect on the experiment.

Finally, in terms of model validity, we find that our model satisfies the linearity, constant variance and normality conditions, and does not have any overly concerning potential bad leverage.

7 Limitations and Further Investigations

It is important to acknowledge the limitations of our study and suggest improvements that could be applied to or inspire potential future investigations.

During data collection, we noticed that some individuals scored near perfect scores (19-20 out of 20) on their 'before' memory score tests. This was problematic because it meant that there was no room to fully capture any improvement in score. For example, for an individual who scored 19, an improvement of +1 ('after' score of 20) and an improvement of +6 (hypothetical 'after' score of 25) would both be reflected as the same thing (an 'after' score of 20).

We could improve this by changing the way we measure memory. Our experiment used the "Memory Vocabulary Test" on the Island because we determined that it was the best option available to us based on our research. However, since there are many types of memory functions, we could use the other memory tests that the Island provides (Memory Game, Memory Test Cards) to create a composite score that provides a more holistic picture of an individual's memory abilities.

Another limitation was that our blocking factor of gender did not have the same effect that we were anticipating. Contrary to the research papers we studied, our ANOVA test concluded that gender did not have an effect on memory scores and that our blocking design was ineffective. Understanding this, a potential further investigation could extend the study's population to include elderly with dementia, and repeat our experiment using dementia (with levels of present/not present) as a blocking factor instead of gender. This would provide insight on the effectiveness of music in affecting memory abilities in memory-impaired elderly.

Lastly, knowing that it was the control group that caused the Duration design factor to be significant, we could alter the experiment design to truly determine the effect of duration on memory test scores. Our Tukey analysis deemed that the difference between 1 plays and 2 plays is negligible; however, one pairing is not sufficient enough to determine the overall effect of duration. Thus, we could extend the variable to include more treatment levels and exclude the control group: we would add another level to the genre called "Control" and change the Duration treatment levels to 1-5 instead of the current 0-2.

Through these improvements and further studies, we can better understand the relationship between music and short-term memory abilities; but more importantly, we can extend this research to long-term memory in order to determine if listening to music is a viable method of hindering memory loss in elderly.

8 References

1) Bugter D., Carden R. (2012). The effect of music genre on memory task. Modern Psychological Studies. https://scholar.utc.edu/cgi/viewcontent.cgi?article=1214&context=mps

2) Loprinzi, P. D., & Frith, E. (2018, May 31). The role of sex in memory function: Considerations and recommendations in the context of exercise. Journal of Clinical Medicine. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6028920/

3) Mammarella, N., Fairfield, B., & Cornoldi, C. (2013, July 25). Does music enhance cognitive performance in healthy older adults? the vivaldi effect - aging clinical and experimental research. SpringerLink. https://link.springer.com/article/10.1007/BF03324720

4) McDougall, G. J., Pituch, K. A., Stanton, M. P., & Chang, W. (2014, August). Memory performance and affect: Are there gender differences in community-residing older adults?. Issues in Mental Health Nursing. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6339525/

5) Proverbio A. M., Valentina L. N., Arcari L. A. (2015, October 15). The effect of background music on episodic memory and autonomic responses: listening to emotionally touching music enhances facial memory capacity. Scientific Reports. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4606564/

6) Riedel, B. C., Thompson, P. M., & Brinton, R. D. (2016, June). Age, APOE and sex: Triad of risk of alzheimer's disease. The Journal of Steroid Biochemistry and Molecular Biology. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4905558/#:~:text=Evidence%20indicates%20that%20the%20APOE,mild%20cognitive%20impairment%20and%20AD.

7) Theofilidis A., Karakasi M., Kevrekidis D. P. (2020, November 10). Gender Differences in Short-term Memory Related to Music Genres. National Library of Medicine. https://pubmed.ncbi.nlm.nih.gov/32891706/

8) Ward E. V., Berry C. J., Shanks D. R. (2013, September 23). Age effects on explicit and implicit memory. Frontiers in Psychology. https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00639/full.