Background Music Will not Distract Students' Attention*

Hong Pan

April 22 2021

Abstract

Music has gradually become an essential part of our lives, and it has become a common trend for most students to choose to listen to music while studying. This paper examines whether the background music affects students' attention level by conducting a within-subjects experiment to compare the time of finishing the Trail Making Test with and without music. From the liner mixed mode, the result shows that background music does not significantly impact attention level. Apart from that, male students tend to be more easily distracted by the background music.

Keywords: background music, attention, trail making test, executive functioning skills

Contents

1 Introduction **2**

2 Data **3**

2.1 Participants 3

2.2 Experiment 3

2.3 Mearues 4

2.4 Descriptvie Analysis 4

3 Model **7**

3.1 Linear Mixed Model 8

3.2 Modeling Procedure 9

3.3 Model Diagnostics 9

4 Results **12**

4.1 Random Effect Result 12

4.2 Full Model Result 13

4.3 Model Assessment Result 14

*Code and data are available at: https://github.com/HongTrista/Music_effect

5 Discussion **14**

5.1 Bias and ethical concerns 14

5.2 What is done in this this paper? 15

5.3 What is something that we learn about the world? 15

5.4 What are some weaknesses of what was done? 15

5.5 What is left to learn or how should we proceed in the future? 15

6 Appendix **16**

References **19**

1 Introduction

Should students listen to music while studying? In contemporary times, music has become an indispensable part of human life. With the development of technology, students can more easily and conveniently access music through various devices and channels. According to Nielsen's report, teens use an average of 3.8 devices in a week to engage with music, such as smartphones, laptops, and tablets, and on average, teens spend over 32 hours per week listening to music(Nielsen 2017). Meanwhile, the survey took by Jackson Gutknecht illustrates that around 75 of 80 students said they listen to music while studying(Gutknecht 2018). Listening to music during the learning process has become a common trend for many students, therefore, it is important to understand the influence of background music on students' attention level.

Trail Making Test, which is used to measure executive functioning skills, was utilized to study whether background music would influence students' attention level. Executive functions refer to the mental processes needed when people have to concentrate or pay attention, which is "a specific set of attention-regulation skills involved in conscious goal-directed problem-solving.(Philip David Zelazo 2016)" Defectiveness of executive functioning skills will make students have a hard time paying attention(Team 2021). In the Trail Making Test, each round consists of several dots numbered from 1 to 12 and lettered A to J, randomly distributed on a page. The participants are required to connect the circles as soon as possible.[1]

The within-subjects experiment is conducted by randomly selected students from one specific school, and each participant needs to complete two rounds of tests, with and without music. The dataset created by Arriken Worsley(Worsley 2020) records the time of accomplishing the test, gender of the student, music favourability, and the order of test which is starting with the music or without the music. The descriptive analysis results show that students performed better without music. In addition, whatever the order of the test, the second round will be faster on average. Furthermore, background music has the biggest influence on male's executive functioning skills. Because the dataset is longitudinal, the linear mixed model is applied to do the statistical analysis. The result from the model illustrates that the students' attention level is not significantly impacted by listening to the music. In this study, the practice effect[2] may exist since the students need to do the test twice continuously.

The rest of this article is structured as follows. In the data section, the study design, including the experiment process, sampling method, and the dataset variables, will be explained. In the model section, the model selection will be discussed, which aims to investigate how background music impacts students' attention level. In the result section, the descriptive and statistical analysis detail will be presented. Finally, in the discussion section, the main findings, limitations and direction of future research will be discussed. In the Appendix, the screenshot for the survey and the real Trail Making Test used for the experiment will be shown.

[1] The order should alternate between numbers and letters, i.e. 1, A, 2, B,…

[2] People improve their scores simply by taking the same test multiple times. (Kendra Cherry,2019)

2 Data

The data used to investigate whether background music will effect students' attention level. The data analyzed through R Statistical Programming Language(R Core Team 2020). The Tidyverse package(Wickham et al. 2019) used for data processing, kableExtra package(Zhu 2021) is applied to generate tables, ggplot2(Wickham 2016) is used to draw diagrams, here(Müller 2020) used for import dataset and nlme(Pinheiro et al. 2021) utilized to create model. The performance pacakge(Lüdecke et al. 2020) and lattice package(Sarkar 2008) are used to draw diagram help to check the model assumptions. The package stargazer(Hlavac 2018) is used to visualize the table.

2.1 Participants

The experiment was conducted in a high school in 2020, and the sample of the study was comprised of 47 students between 13 and 17 years old randomly chosen from the school. Both male and female students participated in the research on whether background music affects people's attention. The pie chart shown in Figure 1 illustrates that 51% are female, 47% are male, and the rest 2% are transexual.

2.2 Experiment

2.2.1 Intervention

The purpose of the experiment is to examine the background music effects on students' attention level. The experimental design plan randomly selects some students from a school, completes the Trail Making Test with and without background music, and records the completion time. In this experiment, the intervention is the existence of background music. This experiment wants to determine whether the background music will affect the concentration level of students by studying whether the time for students to complete the test with background music is longer than the time for completing the test without background music. If the result is yes, it means that music will affect the student's attention to a certain extent; otherwise, it will not. The experiment can help the target audience have some basic understanding of whether music has an impact on attention and can make the right decision whether to listen to music in future studies or work to avoid inefficiency.

2.2.2 Procedure

As the paper aims to understand whether background music would affect students' attention level, all students in the world would be considered as the target population. The frame population is all students from the school. 47 students being randomly selected from the school as a sampled population to finish the experiment. The detailed procedure will be explained as followed.

The 47 students chosen randomly from the school will be randomly separated into two groups to start the Trail Making Test in a different order. 27 students started the experiment in a quiet environment, and the remaining 20 students started with background music. The Trail Making Test utilized during the experiment are shown in figure 2 and 3 in the appendix. Each test contains 23 randomly distributed dots, including 12 numeric points from 1 to 12 and 11 letter points from A to K. Participants must start connecting from 1 and connect these points to the final point in alternating ascending order, such as 1, A, 2, B, 2, C, etc. During the process, the line should not cross over, making a "trail". The background music, the intervention, while playing during the experiment will be the song Panda by Desiigener because, during the experiment period, it was at the top of the billboard and was the most popular music.

The experiment is conducted by students in the form of a class. Students selected to the same group will start the test uniformly, and a timer is placed in front of each student to record the completion time. In order to counteract the learning effect gained from the first test to the second one to some extent, changing

the order of starting the experiment was applied. The students in the first group will start the Trail Making Test without background music, then will do another test while playing the background music, Panda. The second group of students will start to do the test with background music and do the second test in silence for control. After each round, the time for accomplishing the test will be recorded by seconds. Meanwhile, when participants finish all the two tests, they will be required to fill in an anonymous survey shown in the appendix that will ask for feedback, gender, and personal enjoyment of the song played, on a scale from one to five.

2.3 Mearues

The paper intends to use the dataset collected from the experiment set up by Arriken Worsley(Worsley 2020) to analyze whether the background music will influence students' attention level. The dataset records 94 observations with 13 variables. The data set contains some empty columns and variables that are not related to the research topic[3]; therefore, only 8 variables and 94 observations are left after changing the dataset into a longitudinal.

The **Id** represents a student, and a number represents a student. **Test_Type** illustrates whether the students finished the test with background music or without, the control group or the treatment group. **Time_Sec** shows how many seconds does the student to accomplish the Trail Making Test. The column, **Gender**, shows the gender of the student. **Rating** indicates the student's preference for background music. After completing the experiment, students need to fill out a questionnaire anonymously, including rating background music from 1 to 5. The higher the score, the more they like the background music. The `Feedback` column shows the students' feelings and evaluations of the entire experiment. **Change** shows test 2 compares with test 1, by what percentage increase or decrease the time. **Order** records the order of the test, which helps to illustrate whether the test is the first or second round.

2.4 Descriptvie Analysis

2.4.1 Gender distribution

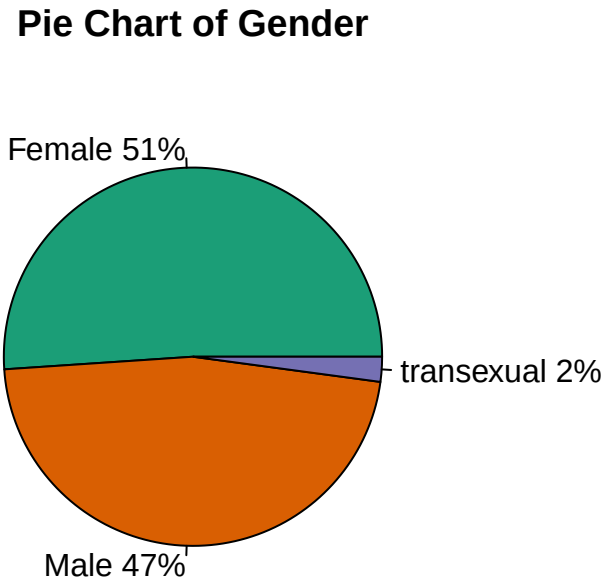


Figure 1: Gender ratio chart

[3]The variables are "question" (the experimenter's comment) as well as "Diff-AVG" and "DIF_AVG". The raw datasets before cleaning could be found in the Input folder, music_first.csv and no_music_first.csv.

According to the experiment, students are randomly selected from the school. Students need to fill out a questionnaire after completing two "Trail Making Test" to inform the experiment initiator of their gender and other information. The pie chart(Figure 1) shows that among the students selected for the experiment, the proportions of men and women are almost the same, which leads to the conclusion that the distribution of men and women is almost the same. Female accounted for 51% of the total number, male accounted for 47% of the total number, and the rest were transexual.

2.4.2 Mean time by test type

Table 1: Average time to complete a test with and without background music

Test_Type	Average Time to Complete a Test	Standard Deviation
Music	47.19	13.53
No Music	44.26	13.30

Table 1 shows the average time to complete a Trail Making Test with or without background music. The results in Table 1 show that, compared with no music, it takes longer for students to complete the test with background music. The average time to complete a test with background music is 47.19 seconds(SD= 13.53), while the average time to complete a test in silence is 44.26 seconds(SD= 13.30), which is approximate 3 seconds faster than with background music.

2.4.3 Average time of completing a test by type and order

Table 2: Average time of completing a test by type and order

order	Test_Type	Average Time to Complete a Test	Standard Deviation
1	Music	52.56	12.14
1	No Music	49.25	15.52
2	Music	39.95	12.06
2	No Music	40.56	10.18

Since the experiment is a within-subject experimental design, each student is both in a control group and a treatment group. In order to study whether the practice effect exists, Table 2 shows that in addition to studying the average time for students to complete the test in each experiment condition, the factor of the order of completing the test is also added. Overall, whether the students completed the first round of the test with background music or first completed the test in a quiet environment, the average completion time of the second round was shorter than that of the first round. The Table 2 shows that when students' first round is added intervention, the average time to complete the test is around 52.56 seconds(SD= 12.14). The second round, which is without background music, takes around 40.56 seconds(SD= 15.52) to accomplish the test. For the students who finish the test without background music in the first round, the average time to complete the test is 49.52 seconds(SD= 12.06). When added intervention, the background music, to the second round, the average time for them to complete the test not only did not increase, but it decreased by one second, about 39.95 seconds(SD= 10.18).

2.4.4 Average time for different genders to complete the test by test type

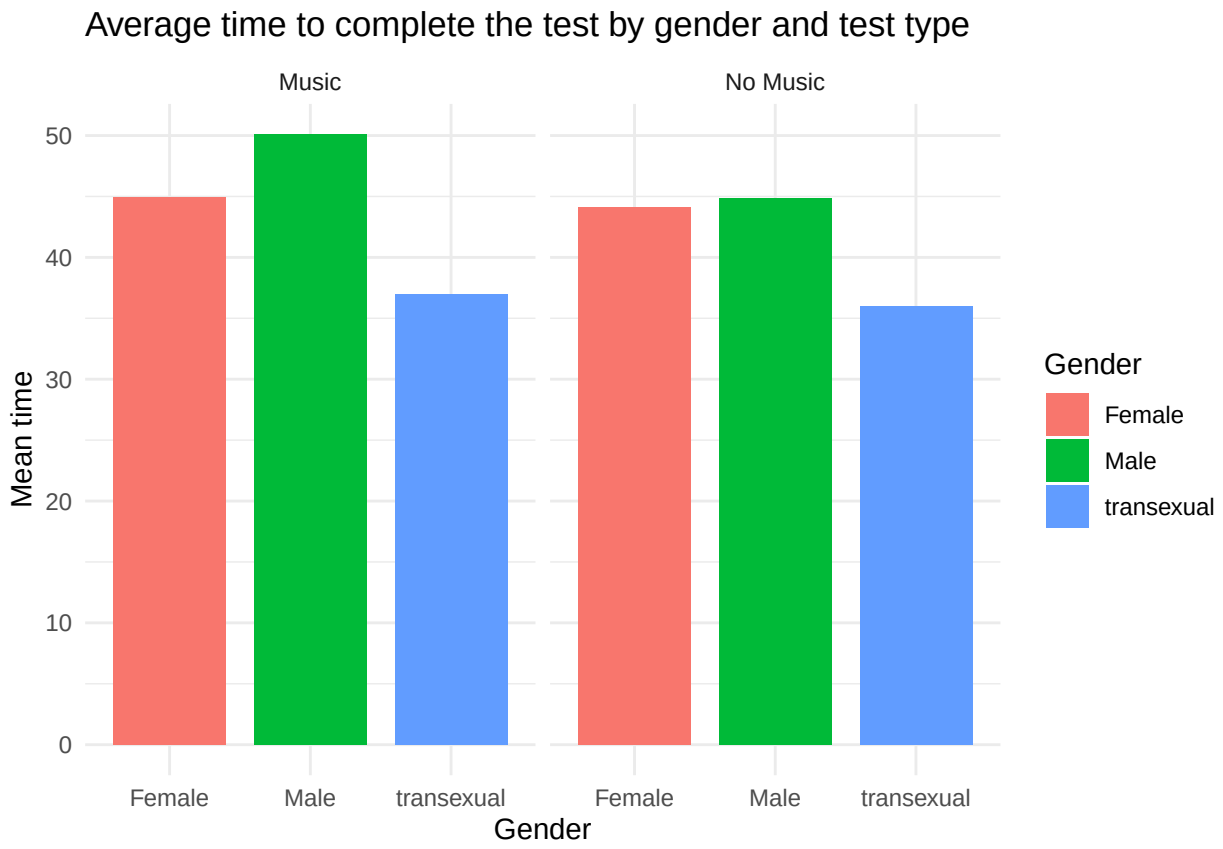
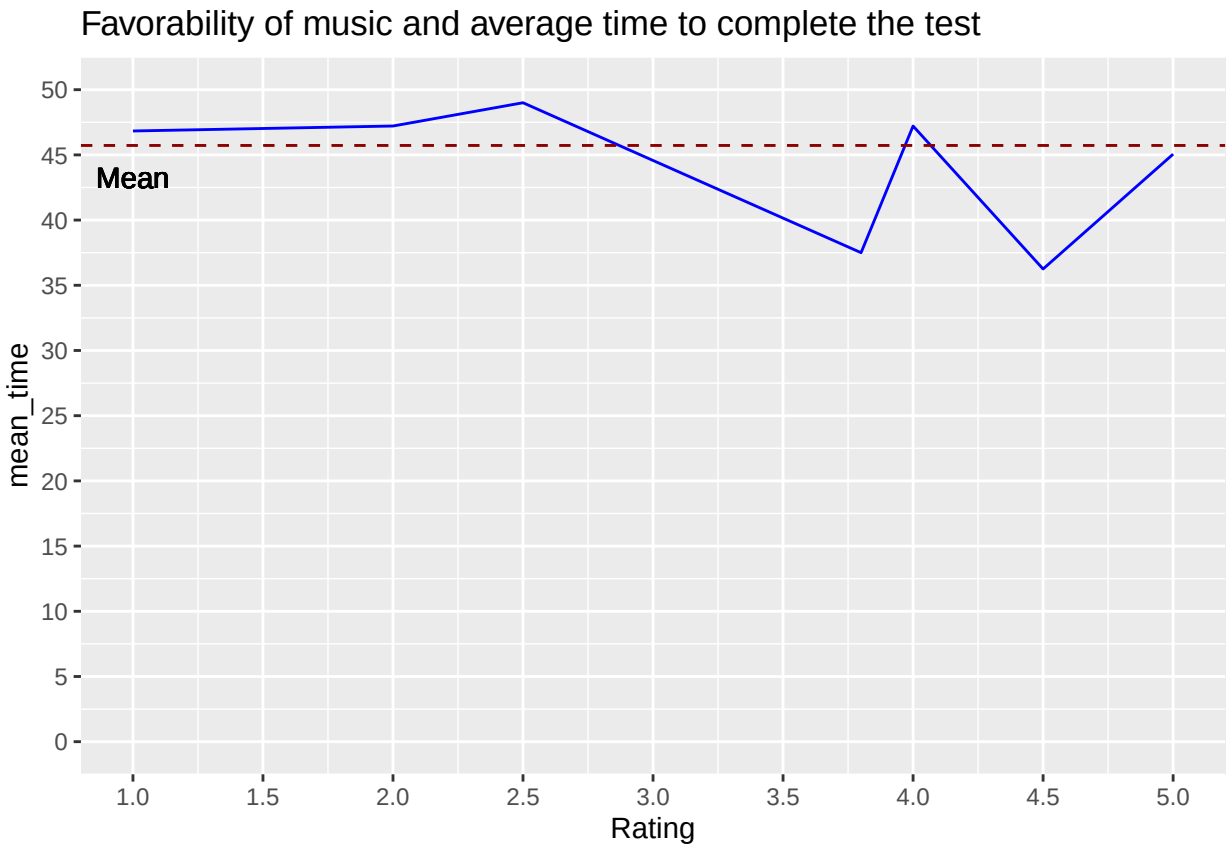


Figure 2: Average time for different genders to complete the test by test type

The paper(Goldman 2017) illustrates females are good at concentration and not easily being distracted compared with males. In order to study whether background music will affect students' attention levels differently due to different genders, Figure 2 is drawn. Overall, the conclusion from the graph goes along with the conclusion the Bruce Goldman(Goldman 2017) made, which is females are less affected by the intervention, while males' attention levels are easily affected. Background music increased the average time for female and transexual students to complete the test by about one second. In contrast, background music influences male students' average time to complete the test increased by 5 seconds. And whether with or without background music, the average completion time for men is the longest, while the average time for transexuals is the shortest.

2.4.5 Music favorability and average time to complete the test



The line chart shown in Figure **??** is used to visualize the relationship between the degree of enjoyment of music from students and their attention levels. The blue line indicates the average time to complete the test at each rating level, and the brown dash line represents the total average time for accomplishing the test. As we can see from the chart, when the rating is approximately less than 3, the average time for students to complete the test is all greater than the overall average. When the rating is greater than 3, the average time for all students is significantly less than the overall average except for level 4.0. According to the findings in Figure **??**, when students listen to songs they don't like while studying or doing other things that require concentration, their attention level is easily affected.

3 Model

The dataset collected from the within-subject experimental design mentioned in the paper is longitudinal data, which means one participant provides multiple data points[4]. These data will be related to each other because they are from the same participant. In this case, the basic linear regression model cannot be applied as it assumes that data points are independent of one another. Therefore, a linear mixed model provides an elegant model to handle the correlated observations within a group and suitable for the study with multiple trials per participant.

The analysis in this paper will be done entirely through R(R Core Team 2020). From Figure 3 below, each line is for 2 Trail Making Test performed by each student and shows the measured completion time of each student who performed the test under different order conditions. As clearly illustrated, whether there is background music or no background music first, different individuals will have different levels of

[4]In the experiment, each student made the Trail Making test twice, one with background music, the other one without background music.

initial attention level, which suggests the existence of the random effects, specifically random intercepts. In addition, the slope for each line is different, which imply that random slope also exist. So, the linear mixed model is applicable. It quantified how background music played in the background could impact the students' attention level compared to the quiet background.

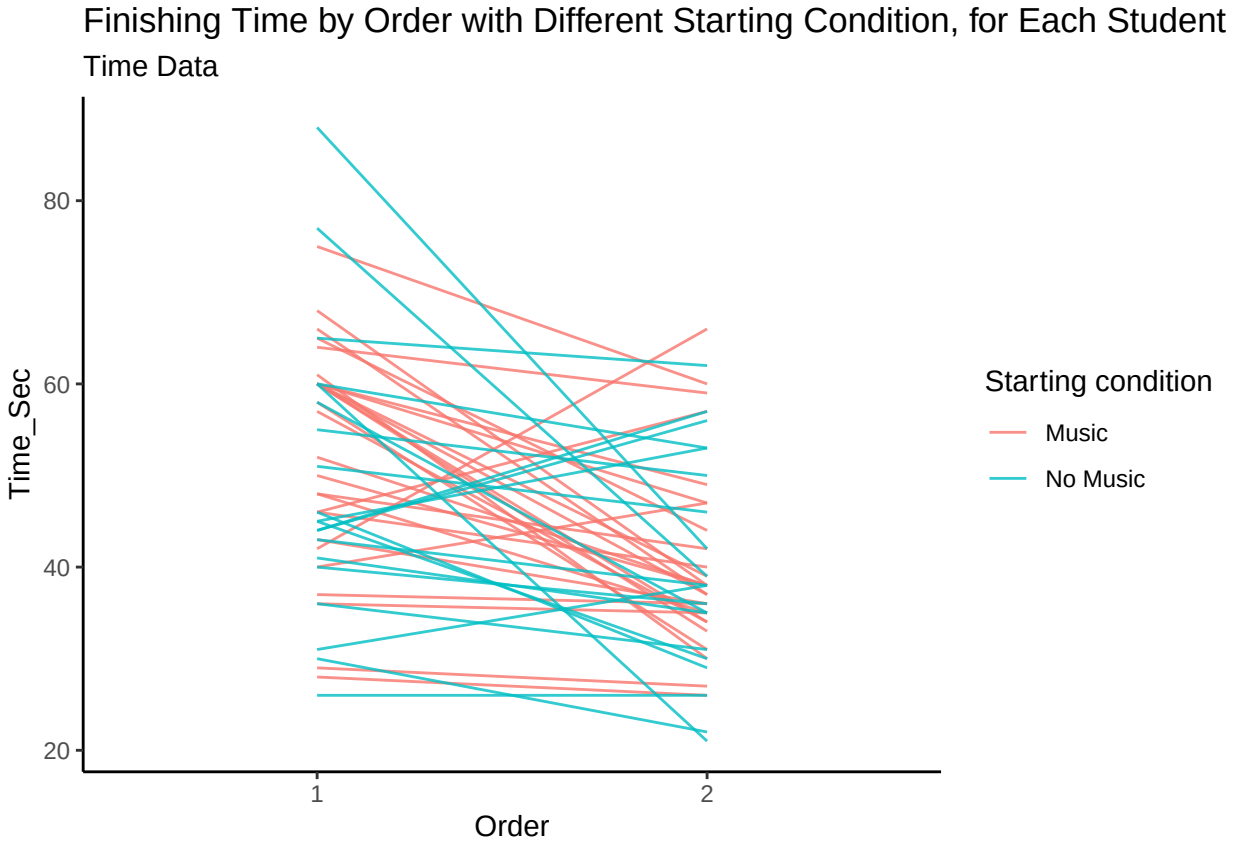


Figure 3: Finishing Time by Order with Different Starting Condition, for Each Student

3.1 Linear Mixed Model

$$\hat{Y}_{TimeSec} = \hat{\beta}_0 + \hat{\beta}_1 * I_{No\ Music_{ij}} + \sum_{a=Male}^{Transexual} \hat{\beta}_2 * Gender_{aij} + \hat{\beta}_3 * I_{order=2ij} + \hat{\beta}_4 * Rating_{ij} + \alpha_{ij} + \epsilon_{ij}$$

Time is the time of the i-th students at the j round to complete the Trail Making Test. **The beta** shown in the formula are the overall fixed intercept and fixed-effect coefficients. **The alpha** is a vector of the random effect for group identity, which includes random intercept and random slope effect for the i-th student. **Epsilon** is the random error for the observation j in group i. **The variables** are the fixed-effect regressors for observation j in group i.

According to the dataset, the response variable and the explanatory variable, Rating, are continuous variables in the model. The reason is that they all can take any value between their minimum value and their maximum value(contributors 2021a). The other variables, type of test, gender and order, are all categorical variables because all of the data within the variables can be stored into groups or categories with names or labels(Blog 2019).

3.2 Modeling Procedure

A linear mixed model consisted of two-part, fixed effect and random effect. The modeling procedure will start to specify the random effects term that best fits the data in the null model, then building up hierarchically. ANOVA, p-value, AIC and RMSE help to find the best-fitted model for the data and check the model validation.

3.2.1 Random Effects

Three null models with different random effects terms will be set up, and the ANOVA function will be applied to compare the models to find the best fitting random effects structure by comparing their p-value to see whether fit has improved with additional items.

The three different random effects term are as followed: *(1| Id): Random intercepts and slopes for each Id.* (1 + Test_Type |Id) = The effect of the background music will vary between each student Id. Random intercepts for test type, random slopes for students influenced by test type. * (1 + Test_Type * order | Id) = Each student Id can have their intercept, random slopes influenced by Test Type and order, and their interaction between Test Type and order.

3.2.2 Fixed Effects

After finding the best-fitting random effect term, ANOVA will be applied to select the significant variables and compare different models. The model for the study starts with only focused variable, test type. Then, other variables will be included in the model one at a time. From the ANOVA output, if the p-value were smaller than 0.1, the variable would be added to the final model. In the process of determining the best model, ANOVA will be applied and the smaller AIC^[AIC(Akaike information criterion) is a mathematical method for evaluating how well a model fits the data it was generated from. AIC is used to compare different possible models and determine which one is the best fit for the data(Bevans 2020.) suggests the better model.

3.3 Model Diagnostics

Before conducting the linear mixed model, the following assumptions need to be checked: whether the variance of the residuals is equal across groups, whether the residuals of the model are normally distributed, and whether there does not exist multicollinearity.

Figure 4 helps to check the assumption of the variance of the residuals is equal across groups. From the chart, there seems to be an evenly spread around the central line and no patterns evident, which means the assumption of homogeneity of variance being satisfied.

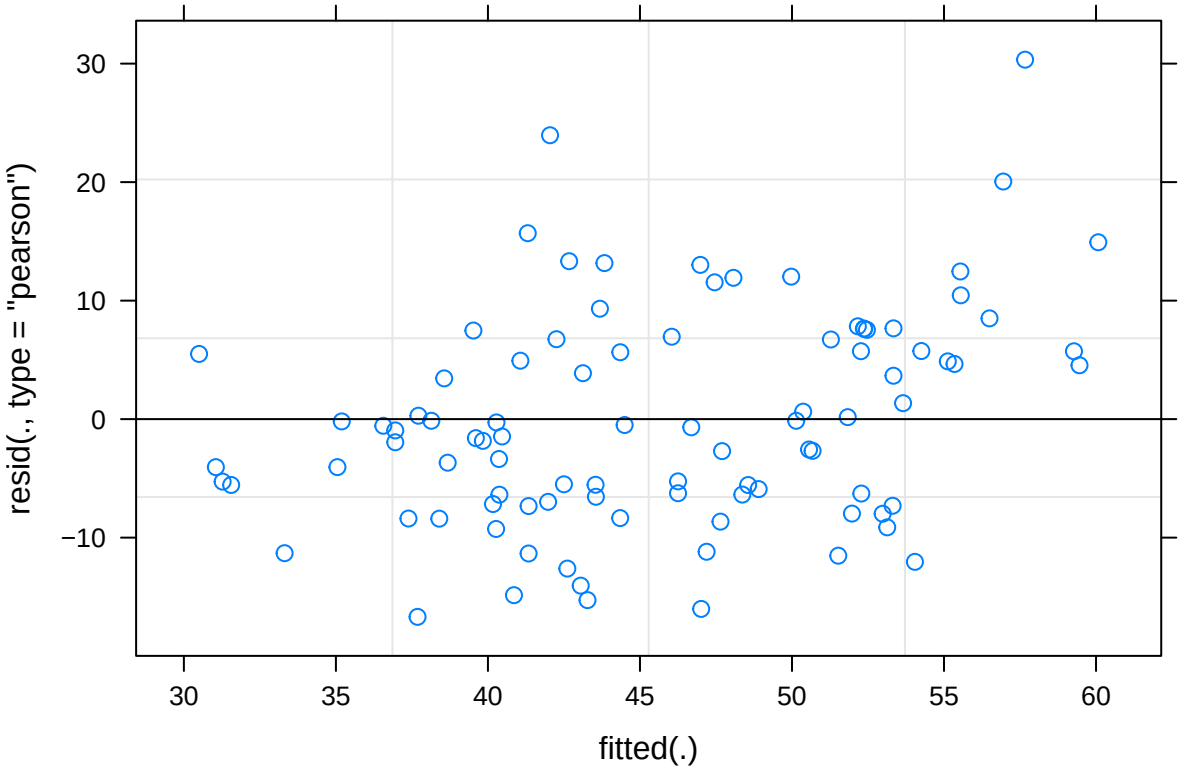


Figure 4: The variance of the residuals is equal across groups

QQ plots(Figure 5) can provide an estimation of where the standardized residuals lie with respect to normal quantiles. Strong deviation from the provided line indicates that the residuals themselves are not normally distributed. From Figure 5, there is some deviation from eh expected normal line towards the two-side tails, but overall all the line looks straight and therefore pretty normal and suggests that the assumption is not violated.

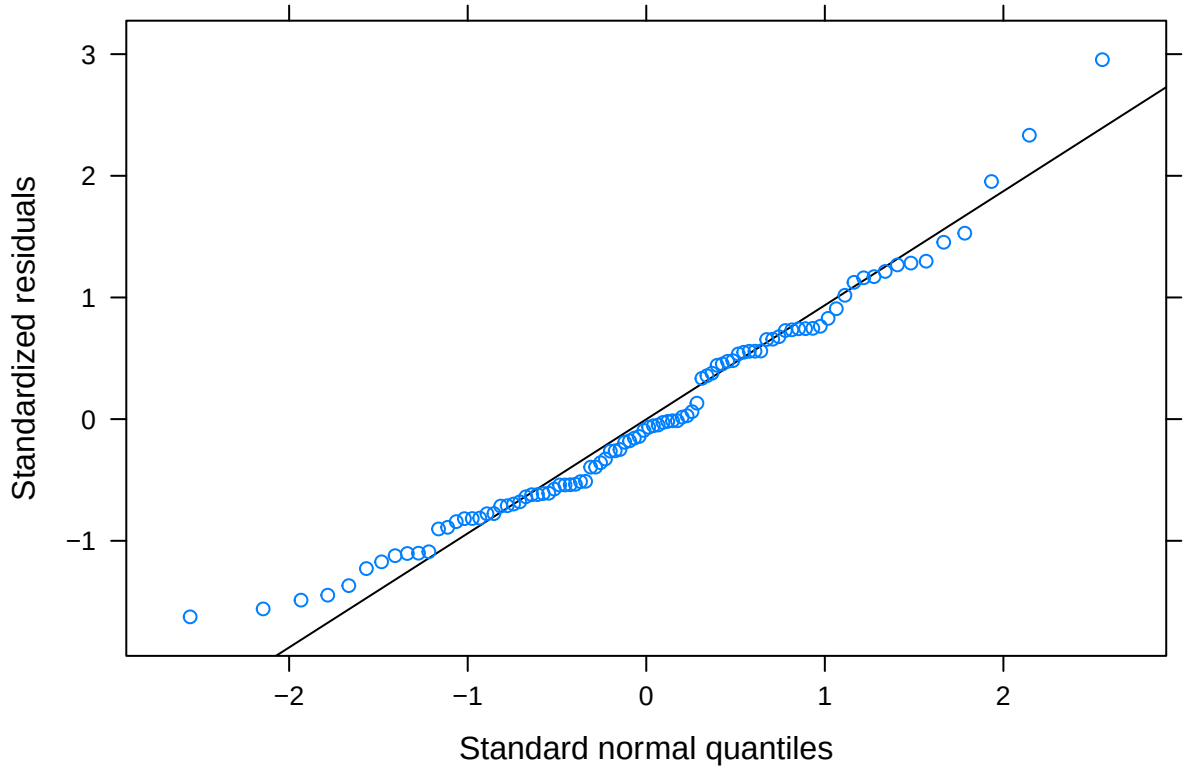
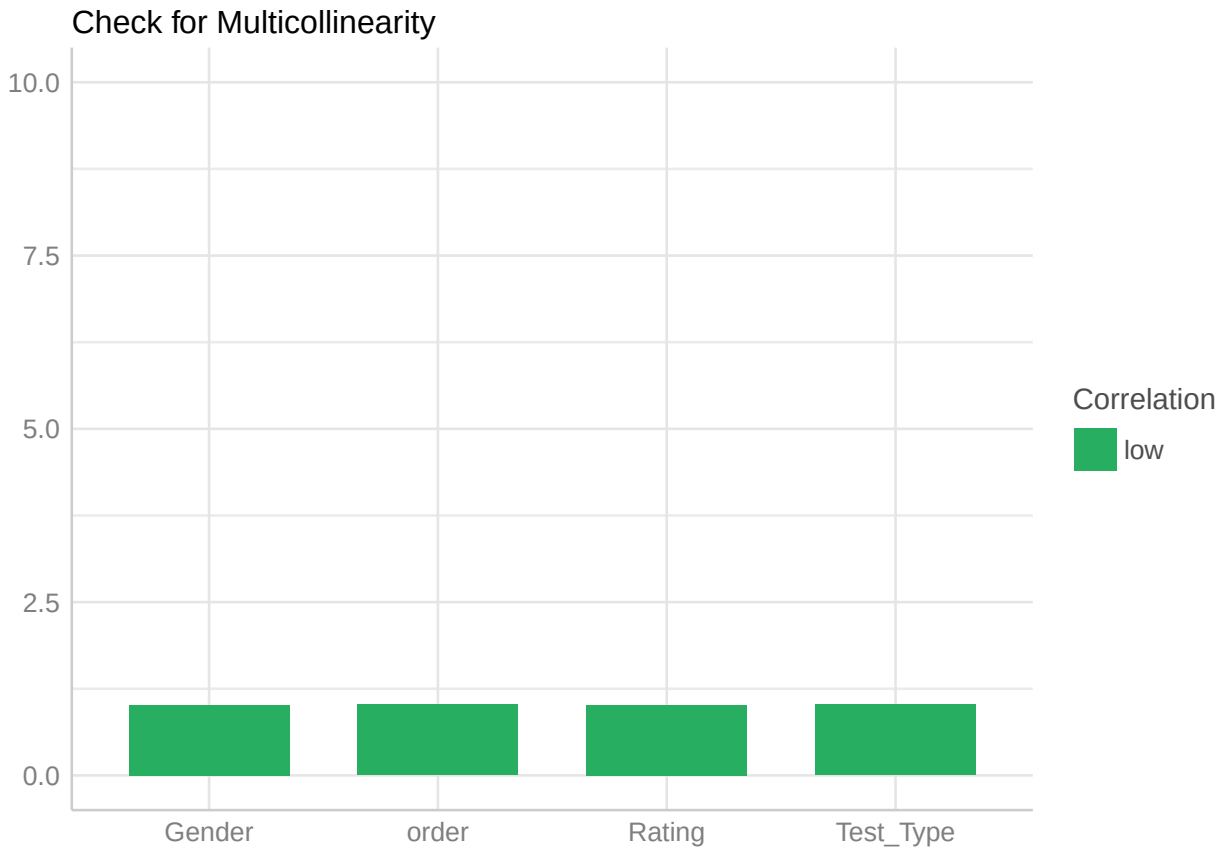


Figure 5: The residuals of the model are normally distributed

Variance inflation factor (VIF) is a measure of the amount of multicollinearity, which means one predictor variable can be linearly predicted from the others with a substantial degree of accuracy(contributors 2021b). The bar chart in Figure **??** illustrates that all the independent variables in the model have a VIF value less than 2.5, which means there are low, and even no multicollinearity exists.



4 Results

The purpose of the study was to explore whether background music has a distracting effect on students' attention levels. There were a total of 47 students participated in the study. A linear mixed model was used to analyze the results for the research question for the within-subject design. An alpha level of 0.1 was used for the analysis.

4.1 Random Effect Result

The output table(Table 3) using ANOVA to compare null models with different random effect terms shows that nullmodel1 is the best because the p-value for comparing nullmodel1 and nullmodel2 is larger than 0.1. There is no improvement from nullmodel2 to nullmodel3 due to a high p-value,0.994. As a result, the best random effect term is (1|Id).

Table 3: The result of comparing model by ANOVA

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
nullmodel1	3	759.7853	767.4152	-376.8927	753.7853	NA	NA	NA
nullmodel2	5	761.2598	773.9763	-375.6299	751.2598	2.525539	2	0.2828696
nullmodel3	12	774.2202	804.7397	-375.1101	750.2202	1.039605	7	0.9941641

4.2 Full Model Result

After finding the best-fit random effect term, the predictor variables will be added to the model one by one. Model 1 in the table only includes the most important variable, Test_Type. The result in the table(Table **??**) shows that the coefficient of Test_Type is -2.936 with a 90% confidence level is -8.064 to 2.192, which illustrates that the average time for completing the Trail Making Test with or without background music is not significantly different between groups. Then, the variable order was added to model 2. The result from the Table **??** shows that the coefficient for the variable order is -10.65 and the 90% confidence interval is -14.848 to -6.452, which does not cover value 0, suggesting that the average time to complete the test in order 2 is 10.65 seconds faster than the order 1, holding other variables constant. Meanwhile, for model 3, the interaction between test type and order being considered for testing whether the effect of the order variable depending on the level of test type. The output from model 3 shows that the interaction term is not significant, p-value larger than 0.1 and the 90% confidence interval cover value 0. Furthermore, variable Rating was added to model 4 and showed not significantly related to completing the test. Finally, the same logic applied to model 5, which plus variable Gender and illustrate that the different Gender would not be a factor to influence student's time of accomplishing the test.

In order to find out the best-fit model for the data, ANOVA and AIC were applied to compare models. In Table **??**, there is an AIC value in each model. The smaller the AIC value, the better the model. From Model 1 to Model 5, their AIC values are 760.54, 742.68, 744.21, 746.02, and 747.68 respectively. Among them, the AIC value of Model 2 is the smallest. Therefore, it is deduced that Model 2 should be the best model among the five models.To further prove that model two is the best model among the five, ANOVA is used to compare the five models. In Table 4, analyze which model is the best by looking at the p-value. When comparing model 1 and model 2, the p-value is much smaller than 0.01, which shows that model two is significantly better than model one, so model two is the best at this time. When comparing model two and model three, the p-value is 0.497, which is greater than 0.1, so adding predictor variables does not increase the fittness of the model. The same logic applies to the comparison between Model 3 and Model 4 and Model 4 and Model 5. In conclusion, through ANOVA's analysis, the conclusion is that Model 2 is still the best.

Linear Mixed Model Result

	Dependent variable:				
	Time_Sec				
	(1)	(2)	(3)	(4)	
Test_TypeNo Music	-2.936	-1.350	-3.306	-2.992	-3
	(-8.064, 2.192)	(-5.548, 2.848)	(-10.329, 3.718)	(-10.145, 4.161)	(-11.02
order2		-10.650***	-12.606***	-12.292***	-13.
		(-14.848, -6.452)	(-19.629, -5.582)	(-19.445, -5.139)	(-20.32
Rating				-0.489	-0
				(-2.673, 1.696)	(-2.47
GenderMale					3
					(-2.43
Gendertransexual					-9
					(-28.21
Test_TypeNo Music:order2			3.911	3.285	5

			(-7.352, 15.174)	(-8.299, 14.868)	(-6.447
Constant	47.191***	51.723***	52.556***	54.058***	52.5
	(43.396, 50.987)	(47.808, 55.638)	(47.974, 57.137)	(45.930, 62.185)	(43.913
Observations	94	94	94	94	9
Log Likelihood	-376.271	-366.338	-366.107	-366.011	-364
Akaike Inf. Crit.	760.542	742.675	744.214	746.022	747
Bayesian Inf. Crit.	770.716	755.392	759.474	763.826	770
Note:					*p<0.05; **p<0.01; *

Table 4: To Find Out the Best Final Model by ANOVA

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
m1	4	760.5425	770.7156	-376.2712	752.5425	NA	NA	NA
m2	5	742.6751	755.3916	-366.3376	732.6751	19.8673298	1	0.0000083
m3	6	744.2142	759.4739	-366.1071	732.2142	0.4609627	1	0.4971744
m4	7	746.0225	763.8255	-366.0112	732.0225	0.1917094	1	0.6614977
m5	9	747.6820	770.5716	-364.8410	729.6820	2.3405008	2	0.3102892

4.3 Model Assessment Result

Root Mean Square Error(RMSE) is used to evaluate the quality of the model. RMSE stands for the standard deviation of the residuals and it helps to tell how concentrated the data is around the line of best fit(To 2021). In Table 5, the RMSE value of the final model, model 2, is 0.135 which a good sign for the model and illustrates that the model have good ability to accurately predict the data. The R-square is applied to assess the model as well. R-square is a statistical measure of how close the data are to the fitted regression line(Blog 2013). For the linear mixed model, R-square consisted two types, which are marginal R-square and conditional R-square. The paper will mainly focused on marginal R-square values, which provides the variance explained by only fixed effect. The marginal R-square for the final model, model 2, shown in Table 5 is 0.169, which means 16.9% of the variation in the dependent variable can be explained by the final linear mixed model.

Table 5: Assessment Result

RMSE	Marginal R-square	Conditional R-square
0.134	0.169	0.41

5 Discussion

5.1 Bias and ethical concerns

. no ethical concerns . learning effect . hawthorn effect .

5.2 What is done in this this paper?

. Descriptive analysis: 1) similar number of students for each gender; students finshed the test with background music in a longer time and the learning effect exist. 2)male easily be impacted by music than other genders 3)favouriblity of the background would influence the time

. Statistical Analysis: 1) random intercept and random slop occurs 2) each students whatever which group they belong to have different attention level. 3) background music is not significantly impact studnets' attention level, but order did influnce the time of completing the test.

5.3 What is something that we learn about the world?

.

5.4 What are some weaknesses of what was done?

. internal validity: justified by the procedures of sampling, randomization in the experiment intervention, controlling or confounders and having sufficient sample size in the model. 1) sampling: 2) Randomization in intervention 3) Controlling for confounders 4) Sufficient sample size

.External validity: cannot be representative

5.5 What is left to learn or how should we proceed in the future?

. add more students to do the experiments . change the setting to different types of experiments with Stroop Test.

6 Appendix

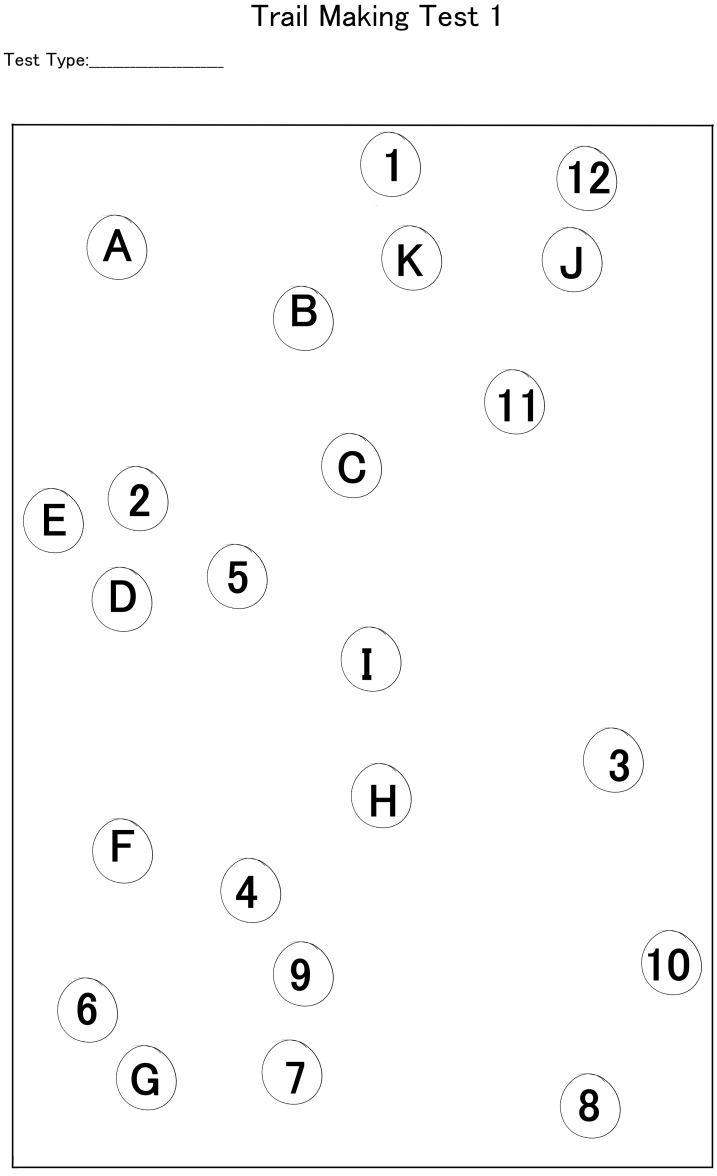


Figure 6: First Trail Making Test

Trail Making Test 2

Test Type:_____

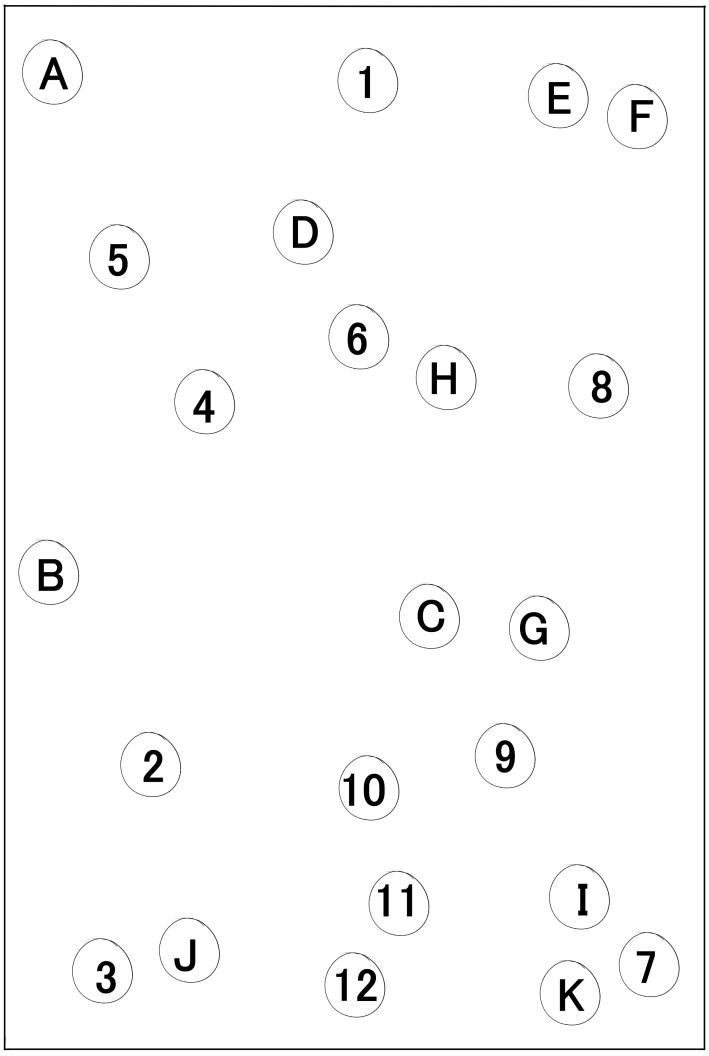


Figure 7: First Trail Making Test

Elizabeth Blackburn School of Sciences
Extended Investigation Research Project

Post-experiment Survey

Please state your gender: (write)

_____

Rate the music you listen to on a scale from 1-5: (circle)

Heavily Disliked 1 ---------- 2 ---------- 3 ---------- 4 ---------- 5 Heavily Enjoyed

Any Feedback for the experiment: (optional)

_____

_____

_____

_____

Figure 8: First Trail Making Test

References

Bevans, R. 2020. "An Introduction to the Akaike Information Criterion." https://www.scribbr.com/statistics/akaike-information-criterion/#:%7E:text=The%20Akaike%20information%20criterion%20(AIC,best%20fit%20for%20the%20data.

Blog, F. 2019. "Categorical Vs Numerical Data: 15 Key Differences & Similarities." https://www.formpl.us/blog/categorical-numerical-data.

Blog, Minitab. 2013. "Regression Analysis: How Do I Interpret R-Squared and Assess the Goodness-of-Fit?" https://blog.minitab.com/en/adventures-in-statistics-2/regression-analysis-how-do-i-interpret-r-squared-and-assess-the-goodness-of-fit.

contributors, Wikipedia. 2021a. "Continuous or Discrete Variable." https://en.wikipedia.org/wiki/Continuous_or_discrete_variable.

———. 2021b. "Multicollinearity." https://en.wikipedia.org/wiki/Multicollinearity.

Goldman, Bruce. 2017. "The Cognitive Differences Between Men and Women." https://stanmed.stanford.edu/2017spring/how-mens-and-womens-brains-are-different.html#:~:text=In%20adulthood%2C%20women%20remain%20more,tend%20not%20to%20be%20gigantic.

Gutknecht, Jackson. 2018. "In a Tiger Hi-Line Survey, over Half of Students Listen to Music When Studying." http://www.hiline.cfschools.org/2018/11/in-a-tiger-hi-line-survey-over-half-of-students-listen-to-music-when-studying/.

Hlavac, Marek. 2018. *Stargazer: Well-Formatted Regression and Summary Statistics Tables*. Bratislava, Slovakia: Central European Labour Studies Institute (CELSI). https://CRAN.R-project.org/package=stargazer.

Lüdecke, Daniel, Dominique Makowski, Philip Waggoner, and Indrajeet Patil. 2020. "Performance: Assessment of Regression Models Performance." *CRAN*. https://doi.org/10.5281/zenodo.3952174.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

Nielsen. 2017. "TIME with Tunes: HOW Technology Is Driving Music Consumption." https://www.nielsen.com/us/en/insights/article/2017/time-with-tunes-how-technology-is-driving-music-consumption/.

Philip David Zelazo, Michael T. Willoughby, Clancy B. Blair. 2016. "Executive Function: Implications for Education." Institute of Education Sciences. https://doi.org/NCER%202017-2000.

Pinheiro, Jose, Douglas Bates, Saikat DebRoy, Deepayan Sarkar, and R Core Team. 2021. *nlme: Linear and Nonlinear Mixed Effects Models*. https://CRAN.R-project.org/package=nlme.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Sarkar, Deepayan. 2008. *Lattice: Multivariate Data Visualization with R*. New York: Springer. http://lmdvr.r-forge.r-project.org.

Team, The Understood. 2021. "What Is Executive Function?" https://www.understood.org/en/learning-thinking-differences/child-learning-disabilities/executive-functioning-issues/what-is-executive-function.

To, Statistics How. 2021. "RMSE: Root Mean Square Error." https://www.statisticshowto.com/probability-and-statistics/regression-analysis/rmse-root-mean-square-error/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Worsley, Arriken. 2020. "OSF." https://osf.io/9xh8e/.

Zhu, Hao. 2021. *KableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.