Introduction

Listening to some music while studying or doing related tasks is quite common among college students, so we want to investigate the effects of different types of music on a person's concentration while doing different tasks, specifically reading and calculating.

A 2010 study [1] asked students to complete a reading comprehension task with either light classical music, hip hop music, or with no music in the background. The study showed that music with a higher intensity is more distracting and has a greater effect on task performance and concentration. One limitation of this study is that it didn't investigate the effect of music on the completion of quantitative tasks, which are different from reading comprehension.

To conduct this study, we planned on creating one set of both math and reading questions, each having 5 multiple choice questions, and asking participants to answer these questions in 15 minutes while listening to a type of music that was randomly assigned. Our measured response will be the rate of correct answers and our factors are different kinds of music – classical music, folk music, and rock music.

We encountered some uncontrollable factors. For example, we cannot control the types of music participants are already familiar with and the academic ability of the participants. This study was rather easy to conduct, we gathered a random sample of people to participate. We administered two different tests, one focused on reading and the other mathematics, with similar difficulty. Each person would receive one of the different treatments of music in an order assigned consistent with the design. Their results will be compared to see how the scores compare for each genre.

Some risks that could arise when trying to conduct this study is ensuring that each sample is truly random, and representative of the group in which we are trying to make a conclusion on. For example, if we were to run this experiment specifically on students at the

University of Wisconsin, we would have to make sure that our sample is representative of the student population at the school, otherwise any conclusions we could make would not be truly representative of the population.

Design

Hypothesis

H_0: The level of concentration will not be influenced by the different types of music, or the types of questions.

H_A: At least one of the factors is influential to the level of concentration.

We evaluate the level of concentration by calculating and comparing the scores participants earned on the test. Reading and mathematics cover a wide range of tasks people perform daily that require concentration, so by measuring these two tasks we can make the hypothesis representative. However, other tasks like communication and writing might also be an important part in evaluating the level of distraction, so not covering them might be a limitation of our experiment.

We ran ANOVA to compare the mean effect of different types of music. Then, we measured the variance of effects and conducted residual analysis to validate model assumptions.

Implementation

We decided to feature folk, classical, and rock music as our factors for music type. Additionally, we used two different questionnaires, one questionnaire featured math questions while the other featured reading questions. Both Math and Reading tests consist of 5 multiple choice questions. We assigned a type of music randomly and asked participants to listen to the assigned type of music when they are taking the test. Originally when we were

planning to collect our data, we wanted to block by familiarity with the type of music being played, but considered that familiarity is not a type of treatment and it is difficult to measure familiarity, we changed this variable to blocking by gender instead.

We then created the blocking of which kind of samples we would need to collect using a 2×3 factorial design with 2 replications. The exact blocking set up of our samples can be seen in Table 1. Summaries of the data collected can be seen in Figures 2 and 3.

Table 1

Gender	**Music**	**Test**
Female	Folk	English
Female	Folk	English
Female	Folk	Math
Female	Folk	Math
Female	Classical	English
Female	Classical	English
Female	Classical	Math
Female	Classical	Math
Female	Rock	English
Female	Rock	English
Female	Rock	Math
Female	Rock	Math

Gender	**Music**	**Test**
Male	Folk	English
Male	Folk	English
Male	Folk	Math
Male	Folk	Math
Male	Classical	English
Male	Classical	English
Male	Classical	Math
Male	Classical	Math
Male	Rock	English
Male	Rock	English
Male	Rock	Math
Male	Rock	Math

Results

The following model was created:

$$Y = a * X_1 + b * X_2 + c * X_3 + d * X_1 * X_2 + \varepsilon$$

- $Y = Score \_on \_test$
- $X_1 = Type \_of \_music \_listened \_to$
- $X_2 = Test \_taken$
- $X_3 = Gender$
- $\varepsilon = Random\ Error$

- a, b, c, and d are all constants which satisfy the model.

When ANOVA was run on this model, none of the variables were found to be significant, all of them having rather high p-values. The p-values for music, test, gender, and music/test interaction terms were p=0.699, p=0.6533, p=0.6533, and p=0.9298, respectively. We suspect that this is due to a lack of power caused by small sample size.

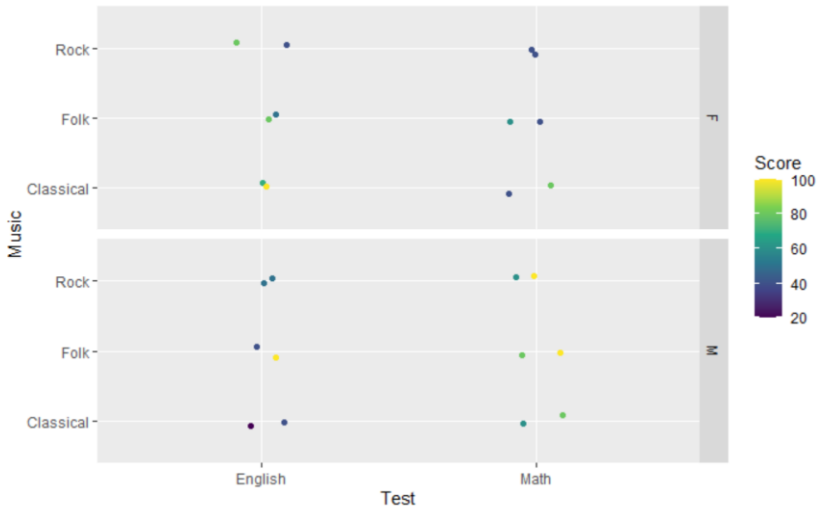


Figure 1: Test results based on type of test, music, and gender (score is based on the percentage of correct answer out of 5 questions)

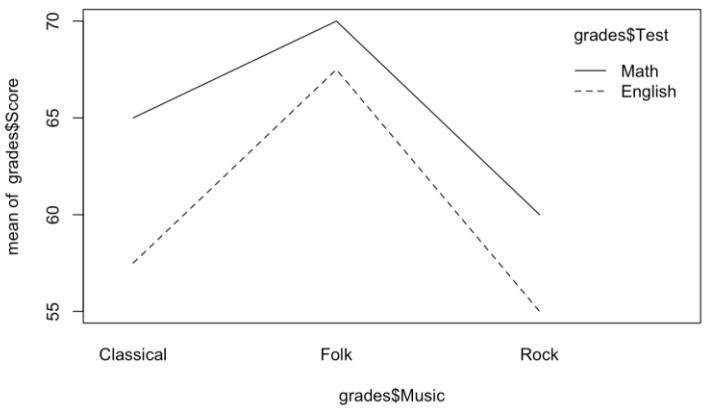


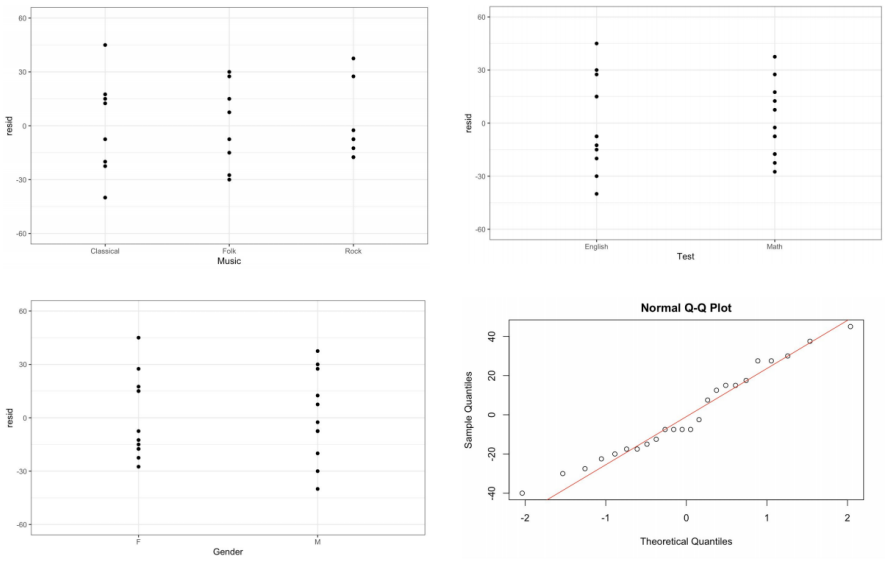
Figure 2: The mean of the test grades for two-way combinations of the two factors, type of test and type of music.

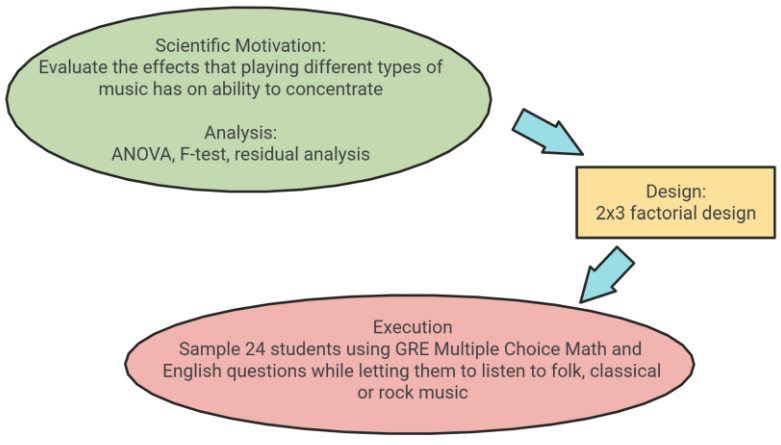
Based on Figure 1, we see that there is no difference in the effects of different types of music between male and female.

Discussion

```
Analysis of Variance Table

Response: Score
           Df Sum Sq Mean Sq F value Pr(>F)
Music       2    525  262.50  0.3658 0.6990
Test        1    150  150.00  0.2090 0.6533
Gender      1    150  150.00  0.2090 0.6533
Music:Test  2     25   12.50  0.0174 0.9828
Residuals  17  12200  717.65
```





When analyzing the ANOVA results, the F-test indicates that none of the main effects in the data are significant. Based on the residual plots, the normality, independence, and additivity assumptions are all satisfied. The equal variance assumption might be mildly violated and requires further investigation.

Initially, when looking at Figure 2, the participants who listened to folk music achieved a better mean score than participants who listened to classical or rock music while doing the test. However, all these differences become insignificant after seeing the ANOVA table, which gives very high P-value. It should also be noted that we got our samples from students studying at UW-Madison and majoring in Math or Statistics, so there is an underlying assumption that their abilities are fairly similar.

To conclude, based on the ANOVA test, different types of music does not affect the person's concentration on doing different tasks. The interaction between type of music and test does not have significant influence on the grades participants achieved.

References

[1] Tze P, ChouM. Attention drainage effect: How background music affects concentration in Taiwanese college students. Journal of the Scholarship of Teaching and Learning 2010; 10(1): 36 – 46.

Appendix

Example of GRE questions:

Math:

Question #1 20 分

If the integer m is a multiple of both 10 and 12, then all of the following must be factors of m EXCEPT

A. 15

B. 12

C. 10

D. 8

E. 6

- A
- B
- C
- D
- E

English:

Conflict between generations may be a problem that has persisted for centuries, but the nature and intensity of the conflict obviously _____ in response to changes in social and economic conditions

- A. increase
- B. disappear
- C. decline
- D. vary
- E. wane

Names of Music Tracks used:

Rock Music: Classic Rock Greatest Hit 60s-70s-80s

Classical Music: The Best of Classical music

Folk Music: Indie/Indie-Folk Compilation - Winter 2020

Code

```
rm(list = ls())
library(readr)
library(ggplot2)
grades <- read_csv("grades.csv")
grades
grades$Music <- as.factor(grades$Music)
grades$Test <- as.factor(grades$Test)
grades$Gender <- as.factor(grades$Gender)
# Visualization
ggplot(grades) +
  geom_point(
    aes(x = Test, y = Music, col = Score),
    position = position_jitter(w = 0.1, h = 0.1)
  ) +
  facet_grid(Gender ~ .) +
  scale_color_viridis_c() +
  theme(legend.position = "bottom")

ggplot(grades) +
  geom_point(aes(x  = Test, y = Gender, col  = Score))  +
  facet_wrap(~Music)

interaction.plot(grades$Music, grades$Test, grades$Score)
# ANOVA
fit.grades <- lm(Score ~ Music + Test + Gender + Music:Test, data =
grades)
anova(fit.grades)
# Model diagnostics
resid <- resid(fit.grades)
y_hat <- predict(fit.grades)
ggplot(grades) + geom_point(aes(x = y_hat, y = resid)) + ylim(-60, 60)
ggplot(grades) + geom_point(aes(x = Music, y = resid)) + ylim(-60, 60)
ggplot(grades) + geom_point(aes(x = Test, y = resid)) + ylim(-60, 60)
ggplot(grades) + geom_point(aes(x = Gender, y = resid)) + ylim(-60, 60)
qqnorm(resid); qqline(resid, col = "red")
```