RELATIONSHIPS BETWEEN BACKGROUND MUSIC AND COGNITIVE CONTROL

By

Mari Hayashi

A capstone project submitted for Graduation with University Honors

May 6, 2021

University Honors
University of California, Riverside

APPROVED

Dr. Aaron Seitz
Department of Psychology

Dr. Richard Cardullo, Howard H Hays Jr. Chair
University Honors

ABSTRACT

When we observe students studying at a library, we see some students studying with their earbuds plugged in and other students without any audio. People have different preferences on the studying environment, but it remains unclear how listening to music affects cognitive control of human brains on average. Some research suggests that listening to music negatively affects cognition because audio prevents humans from focusing on their tasks. Other research predicts, however, that music will help individuals keep up with their task in the long term, because audio gives a variation to the environment and prevents our brains from being bored. Our study explored how the presence of background music affects students' brains working on school work. We studied this relationship by developing a game that gives a player three simple tasks: memorization, calculation, and attention. Participants played the game with and without music for 30 minutes divided into multiple sessions. By comparing their accuracy and response time for each session, we observed how their attentiveness changed as a function of condition and determined if music helps students perform better or not. One challenge I encountered was how we needed to run the experiments remotely because of the pandemic and campus closure. As a result, the recorded data indicates that music distracts humans from paying attention to objects, but listening to classical music encourages humans to solve calculation problems quickly. However, the research failed to derive a significant conclusion about the correlation between music and cognitive task performances.

ACKNOWLEDGEMENTS

I would like to thank my faculty mentor, Dr. Aaron Seitz, for his support on this project. While I had some experiences with game development, I did not have any experience conducting behavioral research using human subjects before starting working on my Capstone project. Dr. Seitz provided me guidance on how to design research-purpose game applications, how to create effective research protocols, and how to run a statistical analysis on collected data to derive a conclusion. I would not have been able to complete this research project without his help. Because of the experience, I was able to learn valuable knowledge outside of my major, including steps of conducting research and ways of effectively formatting a research paper.

Furthermore, I would like to thank Michael Gilbert, a graduate student at UCR, for reviewing this research paper and giving me meaningful feedback. I would also like to dedicate this to Audrey Anna Carrillo, a research coordinator at the Brain Game Center. During the data collection, she helped me a lot recruiting research participants and organizing collected data. Lastly, a special thank you to the research participants and the research assistants at the Brain Game Center, who ran the research with participants despite the campus closure due to the COVID-19 pandemic.

TABLE OF CONTENTS

INTRODUCTION

The presence of background music extremely affects our brain performance. A research paper by Thomas Schäfer states that, “the various surveys and interview studies clearly diverge with regard to the number of different musical functions” that affects emotion, friends, values, and culture (Schafer, 2013, p. 1). Schafer claims that music affects our brain in many different ways. Therefore, it is critical to study the effect of music on our brain if we want to maximize our brain capability and performance.

Many researchers have been investigating how the presence of background music affects people’s attentiveness on tasks. For example, a research study done by Demetriou et al. finds out that music produces flow state (“A mental state in which one’s complete attention is focused on a task”) in listeners (Demetriou, 2016, p. 295). The research suggests that music motivates humans to pay attention to their tasks. Research by Frances H Rauscher also suggests that listening to music improves our brain performance; Rauscher states, “subjects performed better on the abstract/spatial reasoning tests[, the tests that measure the participants’ IQ,] after listening to Mozart than after listening to either the relaxation tape or to nothing” (Rauscher, 1993, p. 1), meaning that participants achieved a higher IQ score when they are listening to a music composed by Mozart. While these research papers highlight the positive effect of music on our attentiveness, research by Lutz Jancke concluded that they did not find any significant influence of background music on verbal learning (Jancke, 2020, p. 1). The results from these research studies are contradicting because they used different tasks, conditions, and types of music. For example, Rauscher investigated the effect of music on our IQ score, but Jancke examined the effect of music on verbal learning. Therefore, in order to investigate the effect of background music on our brain activity, we need to specify the tasks, conditions, and the type of music.

In the current research study, we investigated the effect of two types of music: binaural beats and classical music. Past research studies have shown that these types of music improve our brain performance. For example, Rauscher's research measures the effect of music composed by Wolfgang Amadeus Mozart on subjects' performance on tasks and concludes that listening to classical music helps people perform better on tasks (Rauscher, 1993, p. 1). Similarly, a research study done by Verrusio et al. records that, "after listening to Mozart, an increase of alpha band and median frequency index of background alpha rhythm activity (a pattern of brain wave activity linked to memory, cognition and open mind to problem solving) was observed both in Adults and in Elderly" (Verrusio et al., 2015, p. 1). These research papers suggest that listening to the music composed by Mozart improves our brain performances. Another type of audio which is commonly said to improve our brain performance is binaural beats. A research paper written by Rosina Caterina Filimon explains the definition and some applications of binaural beats. Filimon defines binaural beats as the phenomenon which, "occurs when two sounds of different frequencies are released in a headset each in one ear and the brain makes the frequency difference between the two sounds" (Filimon, 2010, p. 104). In her paper, she introduced research done by Robert Monroe to highlight the positive effect of binaural beats on our brains. Filimon describes Monroe's research findings as, "[application on binaural beats into practice] has yielded significant results in the exploration of expanded consciousness states, creative capability development, [...] as well as in improved sleep, hypnosis induction, meditation and relaxation" (Filimon, 2020, p. 108). Just like the songs composed by Mozart, binaural beats also bring our brain to consciousness state and help us stay focused on our tasks.

Previous research suggests that having binaural beats or classical music as a background music enhances human attentiveness on tasks. Therefore, in the current research study, we examined if we would be able to use these two types of music as a tool for students to remain focused on their schoolwork. Past research studies measured the task performance of humans by giving subjects various tasks and comparing their performances in multiple conditions. Therefore, just like the past research, we also designed three tasks that measure the participants' attentiveness and had participants work on the tasks in multiple conditions to observe how the background music condition affected their performance. We designed the tasks based on three tasks used in the past research studies: the Behavioral Pattern Separation Task (BPS-O) (Stark, 2018), Symbolic Arithmetic Task (Au et al., 2018), and the Hearts and Flowers Task (Diamond et al., 2007). The Behavioral Pattern Separation Task is developed by Craig Stark at University of California, Irvine. The research measures subjects' memory performance for objects using not only unrelated objects, but also using objects that look similar to the targets. We used the images used in the Behavioral Pattern Separation Task to compare the subjects' abilities to memorize and distinguish the images while listening to different background music. We were also inspired by the Symbolic Arithmetic Task used in the research done by Au et al. The Symbolic Arithmetic Task gives subjects a set of simple calculation problems, such addition and subtraction, and measures how accurately and quickly subjects responded to the problems. Just like the task designed by Au, we also designed a task that gives participants a set of simple calculation problems and compared their performance between different background music conditions. We designed another task based upon a task called Hearts and Flowers designed by Adele Diamond. The Hearts and Flowers task measures the participants' attentiveness by asking them to differentiate the drawings of heart and flower and quickly react based on the drawing type. Our task also asked participants to quickly differentiate similar drawings and react to them in order to measure how focused the participants were during the training. We incorporated the features from these three tasks to design our own tasks that allow us to understand the effect of background music on human brain performance working on simple tasks.

METHODS

Participants

Fifty one undergraduate students (20.77 ± 0.36 years old) at the University of California, Riverside, were recruited for the research. The recruitment took place from February 2020 to May 2020. Participants were screened to ensure that they were proficient in English and could follow the instructions. Out of the fifty one participants, thirty four participants participated in the research remotely due to the COVID-19 spread and the campus closure. These thirty four participants were screened to ensure that they owned an Android device with 1.5 GB extra space to download the application used for the research and that they were able to take the data generated by the application out of their devices and send it to research assistants. The participants were numbered sequentially from 101 to 151. Three participants were excluded from the final analysis due to incomplete data. The final analysis considered the records of forty eight participants.

Procedure

Participants completed three training sessions, more than one day apart. Each session took 30 minutes. Before the day of the first session, the participants were asked to complete a short questionnaire for us to collect their background information including age, race, and if they usually listen to music while working on school assignment, exercising, or doing housework. During each session, the participants worked on three tasks, Attention Task, Calculation Task, and Memorization Task, on the mobile application using an iOS or Android device. Each task took 10 minutes. The participants were asked to work on the tasks with different background music each time. There were three different background music conditions: No Music, Binaural Beats, and Mozart. In “No Music” condition, participants worked on a task without listening to any music. In “Binaural Beats” condition, a participant worked on a task while listening to binaural beats from the video, “Accelerated Learning - Gamma Waves for Focus, Memory, Concentration - Binaural Beats - Focus Music,” which was uploaded on YouTube by Magnetic Minds (2017). In “Mozart” condition, a participant worked on the task while listening to the Piano Sonata No. 16 in C major composed by Wolfgang Amadeus Mozart and performed by Rousseau on a piano (2019). The participants completed each task listening to different music each session. For example, the following table shows the training session schedule of one participant.

Table 1

Training Schedule for Participant 101

	Attention Task (10 minutes)	**Memorization Task (10 minutes)**	**Calculation Task (10 minutes)**
Session 1	No music	Binaural beats	Mozart
Session 2	Mozart	No music	Binaural beats
Session 3	Binaural beats	Mozart	No music

This participant first worked on the Attention Task in No Music condition, then worked on the Memorization Task in Binaural Beats condition, and lastly worked on the Calculation Task in Mozart condition during the first session. The schedules were created based on Latin Squares to ensure the schedules are counterbalanced. We created three different schedules, and each participant was randomly assigned to one of the schedules. After each session, the participants were given an end-of-session survey to see how much each task stressed the participants. After the third session, the participants were given a post-training survey about their overall experience.

Mobile Application Development

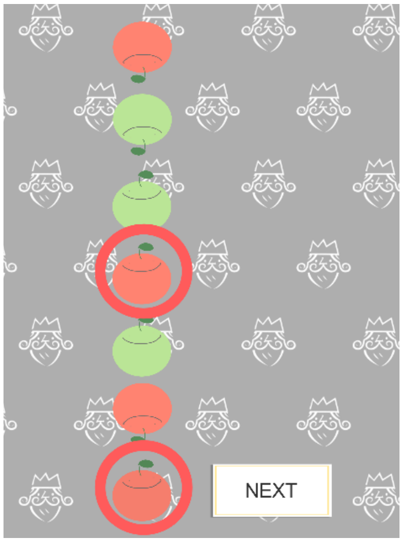
The mobile application used for this project was implemented by myself using Unity, a cross-platform game development software, and C#. We used Unity to build the application in order to make the mobile application compatible with both iOS and Android devices. The application was designed to give the player the following three tasks sequentially: Attention Task, Calculation Task, and Memorization Task. The research assistants specified the ordering of the tasks and the conditions before the beginning of the session,

Attention Task

The first task that was implemented in the mobile application was the Attention Task. The task consisted of multiple short trials. For each trial, the application displayed seven drawings of fruits on the screen (see Figure 1). The participants were asked to only tap on certain fruits with certain colors and orientations. The number of correct fruits displayed on the screen ranged from two to four out of seven. The participants were instructed to tap the “Next” button to move on to the next trial when they thought they caught all of the correct fruits on the screen.

Figure 1

The screenshot of the Attention Task



Note. Seven fruits were displayed for each trial. A red circle appeared around the fruits tapped by the participant. The participants were instructed to tap the "Next" button on the bottom right to move on to the next trial.

Participants were tasked to complete as many trials as possible within ten minutes. For each trial, the application recorded the response time, which was the time taken from when the fruits were displayed to when a participant tapped the "Next" button. The application also recorded the types of fruits displayed and which fruits were tapped for each trial. These records were used to count how many times a participant tapped or missed correct fruits. The participants worked on the Attention tasks in three different conditions on three different days. We compared the response time, the number of trials completed within one session, and the accuracy to observe how their performance changed as a function of condition.

Calculation Task

The second task was the Calculation Task. The task also consisted of multiple short trials. For each trial, the application displayed a simple calculation problem that was one of the following:

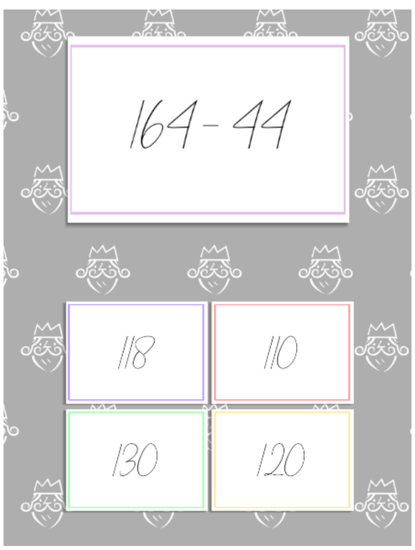
1. Addition with both operands ranging from 11 to 244

2. Subtraction with both operands ranging from 11 to 244 (All answers are positive)
3. Multiplication with both operands ranging from 2 to 9
4. Division with dividend ranging from 81 to 4 and divisor ranging from 2 to 9 (All answers are integers)

The calculation problem was displayed on the upper half of the screen (see Figure 2). The application displayed four choice buttons on the lower half of the screen. Each button had a different number written on it. The participants were tasked to solve a calculation problem displayed on screen and tap on the button with the correct answer written on it. Once one of the choice buttons were pressed, the application moved onto the next trial.

Figure 2

The screenshot of the Calculation Task



Note. The upper half of the screen displayed the calculation problem, and the lower half of the screen displayed four choice buttons.

Participants were asked to complete as many trials as possible within ten minutes. For each trial, the application recorded the response time, which was the time taken from when the calculation problem was displayed to when a participant tapped one of the choice buttons. The application also recorded the calculation problem and the answer chosen by the player for each trial. These records were used to calculate the accuracy for each session. The participants worked on the Calculation tasks in three different conditions on three different days. We compared the response time, the number of trials completed within one session, and the accuracy to observe how their performance changed as a function of condition.

Memorization Task

The last task was the Memorization Task. The task consisted of multiple short trials. For each trial, the application displayed 5 to 7 pictures sequentially, one at a time (see Figure 3). Each picture was displayed for 2 seconds. After all pictures were displayed, the application displayed a question such as “which picture is the 4th picture?,” varying the number for each question (see Figure 4). At the same time, the application displayed 4 choice buttons with a picture on it. One of the 4 buttons showed the correct picture. The participants were tasked to memorize the sequence of the pictures and correctly answer the question.

Figure 3

The screenshot of the Memorization Task - Displaying image sequence

Note. The application displayed 5 to 7 pictures sequentially.

Figure 4

The screenshot of the Memorization Task - Displaying question

Note. The application displayed a question such as "which picture is the 4th picture?," varying the number for each question, and 4 choice buttons with a picture on it.

Each trial contained pairs of pictures that look similar to each other. Based on the pairs, each trial was labeled with an integer value that represents the difficulty. The difficulty of each trial was determined by the pair of pictures a participant had to differentiate. For example, the following pair of pictures had a similarity value of 2.

Figure 5

An example of a pair of pictures with a similarity value of 2

If the problem displayed the picture on the left as a fourth picture in the sequence and asked the participant to recall the fourth picture, then the player had to differentiate the pictures above to select the correct answer. In this case, the difficulty for the trial was 2. If the trial did not contain any similar picture pairs, the difficulty was recorded as 0. The similarity value between two pictures were decided based on the Behavioral Pattern Separation Task developed by the Stark Lab at University of California, Irvine (Stark, 2018). Each pair of pictures was labeled by an integer ranging from 1 to 5 representing how difficult it is to differentiate the pictures, 1 being the hardest to differentiate. For our research, we only used the pairs of pictures with the difficulty from 1 to 3.

Participants were tasked to complete each trial as accurately as possible. For each trial, the application recorded the response time, which was the time taken from when the problem and choice buttons were displayed to when a participant tapped one of the choice buttons. The application also recorded the difficulty value and if the participant answered correctly for each trial. The participants worked on the Memorization tasks in three different conditions on three different days. We compared the response time and the accuracy to observe how their performance changed as a function of condition.

RESULTS

Attention Task

Number of Trials Completed, Accuracy, and Average Correct Response Time

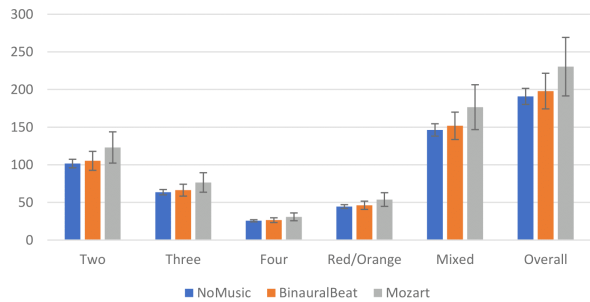
For the Attention Task, we measured the performance of the participants in three different ways: number of trials completed per session, accuracy, and average correct response time. Accuracy was calculated as the number of perfect trials (the trials where a participant tapped on all targeted fruits and ignored all non-targeted fruits) divided by the total number of trials. Average correct response time was the average of the time taken for a perfect trial in seconds. When a participant focused on the task, they would be able to quickly differentiate all target fruits from non-target fruits. Therefore, these three measures would successfully represent the participants' attentiveness during the session.

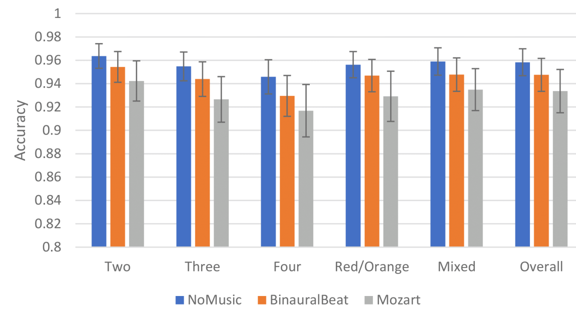
Figure 6 compares the average of the three performance measures among all participants by conditions. The bars labeled "Two," "Three," or "Four" compare the average accuracy by number of target fruits displayed on the screen. The bars labeled "Red/Orange" or "Mixed" compare the average accuracy among all participants by fruit color. The bar labeled "Red/Orange" represents the accuracy of the trials with only red or orange fruits. "Mixed" represents the accuracy of the trials with fruits of different colors, including green and yellow.

Figure 6

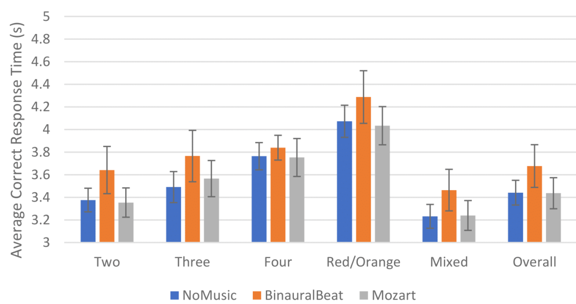
A Bar Graph for the three primary performance measures

A. Number of Trials Completed **B. Accuracy**





C. Average Correct Response Time



According to Figure 6 B and C, the average performance for the trials with two targets was better than the average performance for the trials with four targets. (Higher accuracy and lower response time.) This is likely to be because the participants had a higher chance of missing some of the target fruits when there were more targets on the screen. Also, the average performance for the trials with fruits of mixed color was better than the average performance for the trials with only red or orange fruits for all conditions. (Lower response time.) The most likely explanation is that all fruits colored in yellow or green were not target fruits, so the participants were able to locate the target fruits more effortlessly and quickly when the trial contains fruits of mixed colors. Hence, the average accuracy and response time successfully represents the performance measure of the participants during the sessions.

On the other hand, it is difficult to determine if the number of trials completed reflects the participants' performance just by looking at Figure 6 A. It is because completing a high number of trials for a certain trial group does not mean the participants performed better at the certain trial group; the game was designed to give the participants more number of trials on a certain trial group than other trial groups. Therefore, it is unclear whether a high number of trials completed means better performance or not.

Figure 6 A shows that the participants completed the most number of trials in Mozart condition than in two other conditions. On the other hand, as shown in Figure 6 B, the average accuracy was the highest in No Music condition and the lowest in Mozart condition for all groups of trials. Figure 6 B suggests that the presence of background music, especially those composed by Mozart, impaired the participants' accuracy. If we look at Figure 6 C, however, the average response time was the lowest in No Music and Mozart condition and the highest in Binaural Beats condition for all groups of trials. Figure 6 C suggests that the presence of binaural beats slowed down the participants' response to the stimulus. These results were contradictory and we could determine which condition improved or impaired the participants' overall performance.

For each measure, we conducted ANOVA Test (Analysis of Variance Test, a statistical test used to compare the scores of two groups), within-subject ANOVA Test (ANOVA test that specifically compares the differences in scores under multiple conditions), and Paired Sample T-tests (a statistical test that compares the participants's performances between two conditions) to determine which condition affected the participants' performance the most. We obtained the most significant results from a Paired Sample T-test. Also, a Paired Sample T-test is the easiest and the most direct statistical approach for the data because the test is specifically designed to compute the significance of difference in the participant's performance scores in multiple conditions. Therefore, we only presented the results of a Paired Sample T-test on this paper. The result of a Paired Sample T-test is summarized in the Table 2.

Table 2

Result of Paired Samples T-Test on Number of Trials, Accuracy, and Response Time

			Number of Trials			Accuracy			RT		
			t	df	p	t	df	p	t	df	p
Two(NoMusic)	-	Two(BinauralBeat)	-0.453	47	0.653	1.011	47	0.317	-1.631	47	0.110
Two(NoMusic)	-	Two(Mozart)	-1.327	47	0.191	1.479	47	0.146	0.210	47	0.835
Three(NoMusic)	-	Three(BinauralBeat)	-0.543	47	0.590	1.313	47	0.196	**-2.091**	**42**	**0.043**
Three(NoMusic)	-	Three(Mozart)	-1.283	47	0.206	**1.784**	**47**	**0.081**	-0.822	42	0.416
Four(NoMusic)	-	Four(BinauralBeat)	-0.412	47	0.682	1.576	47	0.122	-0.571	41	0.571
Four(NoMusic)	-	Four(Mozart)	-1.315	47	0.195	1.678	47	0.100	-0.556	41	0.581
Red/Orange(NoMusic)	-	Red/Orange(BinauralBeat)	-0.474	47	0.638	0.915	47	0.365	-1.190	47	0.240
Red/Orange(NoMusic)	-	Red/Orange(Mozart)	-1.311	47	0.196	1.360	47	0.180	0.298	47	0.767
MixedColor(NoMusic)	-	MixedColor(BinauralBeat)	-0.479	47	0.634	1.300	47	0.200	**-1.993**	**47**	**0.052**
MixedColor(NoMusic)	-	MixedColor(Mozart)	-1.311	47	0.196	**1.728**	**47**	**0.091**	-0.088	47	0.930
Overall(NoMusic)	-	Overall(BinauralBeat)	-0.478	47	0.635	1.209	47	0.233	**-1.884**	**47**	**0.066**
Overall(NoMusic)	-	Overall(Mozart)	-1.311	47	0.196	1.630	47	0.110	0.055	47	0.956

Note. Student's t-test. The row with p value less than 0.1 is bolded.

While Figure 6 A suggests that the participants completed a higher number of trials in Mozart condition than in the two other conditions, the result of the Paired Sample T-Test on the number of trials failed to reach significance because all p values are high, as shown in the second column of Table 2.

The participants in No Music condition produced significantly more accurate responses than the participants in Mozart condition. When the trials contained three target fruits, the difference in accuracy between No Music (M = 0.955, SD = 0.085) and Mozart condition (M = 0.927, SD = 0.135) was marginally significant; $t(47) = 1.784$, $p = 0.081$. The table above also indicates that when the participants were exposed to fruits of mixed color, the difference between No Music (M = 0.959, SD = 0.730) and Mozart (M = 0.935, SD = 0.913) condition was slightly significant; $t(47) = 1.728$, $p = 0.091$.

For response time, participants produced significantly faster responses in No Music condition than in Binaural Beats condition. Especially, when trials contained three target fruits, the difference in response time between No Music condition (M = 3.490, SD = 0.911) and Binaural Beats condition (M = 3.765, SD = 1.491) was significant when trials contained three target fruits; $t(42) = -2.091$, $p = 0.043$. Overall, the difference in average correct response time between No Music (M = 3.441, SD = 0.758) and binaural beats (M = 3.677, SD = 1.308) condition was more significant, $t(47) = -1.884$, $p = 0.066$, than the difference between No Music and Mozart condition.

Figure 6 and Table 2 suggest that the presence of Mozart music in the background impaired the participants' accuracy, but the presence of binaural beats in the background impaired the participants' response time. The effect of condition on the number of trials completed failed to reach significance. The result is contradictory and we cannot conclude which of two background music, binaural beats or Mozart, affected the participants' performance more negatively.

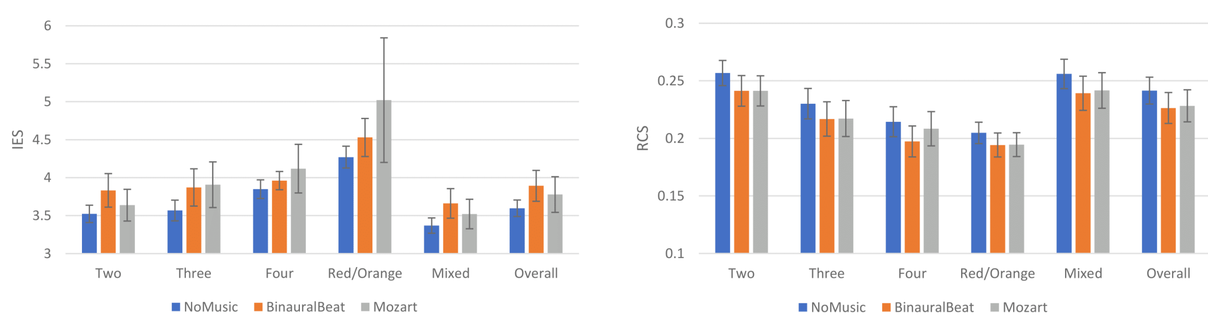
IES and RCS

Analyzing accuracy and response time separately sometimes does not provide a great summary. Some participants might have gone through the trials accurately but slowly, while others might have tried to go through trials as quickly as possible. The speed and accuracy are sometimes inversely proportional to each other. Therefore, we combined the accuracy and the response time value into one value. A research done by André Vandierendonck introduces several ways to combine accuracy and response time. According to Vandierendonck, inverse efficiency score (IES) and rate correct score (RCS) are the two effective scores that combine both accuracy and response time as one value. IES is calculated as the average correct response time divided by accuracy. As a participant performs more accurately and quickly, the participant has a smaller IES. RCS is calculated as the number of perfect trials divided by the sum of response time for all trials. RCS represents the number of perfect trials per second on average. We calculated IES and RCS for each participant using the recorded measures. Figure 7 compares the average of the two additional performance measures among all participants by conditions.

Figure 7

Average IES and RCS

A. IES **B. RCS**



According to Figure 7 A and B, the average performance for the trials with two targets is better than the average performance for the trials with four fruits. (Lower IES, and higher RCS.) Also, the average performance for the trials with fruits of mixed color is better than the average performance for the trials with only red or orange fruits for all conditions. (Lower IES and higher RCS.) Hence, the average IES and RCS successfully represent the participants' performance during the session.

If we look at Figure 7 A, we see that the effect of background music on IES differs by the trial groups. For example, the average IES is the highest in Binaural Beats condition when participants had two target fruits, but the average IES is the highest in Mozart condition when participants had four target fruits. We cannot derive a solid conclusion from Figure 7 A alone. If we look at Figure 7 B, the participants produced higher RCS in No Music condition than in two other conditions for all trial groups. It means the participants produced more perfect trials per second in No Music condition than in the other conditions on average.

We also conducted a paired Sample T-Test on IES and RCS. Table 3 displays the result.

Table 3

Result of Paired Sample T-Test on IES and RCS

			IES			RCS		
			t	df	p	t	df	p
Two(NoMusic)	-	Two(BinauralBeat)	-1.607	47	0.115	1.398	47	0.169
Two(NoMusic)	-	Two(Mozart)	-0.537	47	0.594	1.366	47	0.178
Three(NoMusic)	-	Three(BinauralBeat)	**-2.009**	**42**	**0.051**	1.377	47	0.175
Three(NoMusic)	-	Three(Mozart)	-1.279	42	0.208	1.188	47	0.241
Four(NoMusic)	-	Four(BinauralBeat)	-0.815	41	0.420	**1.763**	**47**	**0.084**
Four(NoMusic)	-	Four(Mozart)	-1.156	41	0.254	0.564	47	0.575
Red/Orange(NoMusic)	-	Red/Orange(BinauralBeat)	-1.183	47	0.243	1.252	47	0.217
Red/Orange(NoMusic)	-	Red/Orange(Mozart)	-0.925	47	0.360	1.129	47	0.265
MixedColor(NoMusic)	-	MixedColor(BinauralBeat)	**-2.071**	**47**	**0.044**	1.544	47	0.129
MixedColor(NoMusic)	-	MixedColor(Mozart)	-0.860	47	0.394	1.220	47	0.229
Overall(NoMusic)	-	Overall(BinauralBeat)	**-1.896**	**47**	**0.064**	1.494	47	0.142
Overall(NoMusic)	-	Overall(Mozart)	-0.819	47	0.417	1.213	47	0.231

Note. Student's t-test. The row with p value less than 0.1 is bolded

According to Table 3, the presence of binaural beats affected participants' IES and RCS negatively. If we look at the data for IES from overall trials, there is a marginally significant difference between No Music (M = 3.597, SD = 0.751) and Binaural Beats (M = 3.892, SD = 1.408) conditions; t(47) = -1.896, p = 0.064. The result proves that a large portion of the participants had lower IES in No Music condition than in Binaural Beats condition, meaning that they had less adjusted response time in No Music condition than in Binaural Beats condition. Moreover, we observed a significant difference in RCS values between No Music condition (M = 0.214, SD = 0.0903) and Binaural Beats condition (M = 0.197, SD = 0.0931) when there were four target fruits; $t(47) = 1.763$, $p = 0.084$. This data suggests that the participants produced more perfect trials per second in No Music condition than in Binaural Beats condition. The result of Paired Sample T-Test on average IES and RCS suggests that the participants managed to pay attention to the task throughout the session the most in No Music condition and the least in the Binaural Beats condition.

However, the effect of condition on performance still did not reach significance. The p values were rarely below 0.05, meaning that no statistically significant difference was observed between the conditions. Also, we obtained smaller p-values for only some of the trial groups. When we analyzed RCS, we observed a significant difference in RCS values between No Music condition and Binaural Beats condition when there were four target fruits, but not for other trial groups and overall trials. If we have had a larger sample size, we would be able to see more consistent results.

Calculation Task

For the Calculation Task, one participant was excluded from the final analysis due to an unusual performance. While most participants solved around 100 - 150 trials per session, the excluded participant completed more than 1,000 trials within one session. He also had significantly low accuracy.

Number of Trials Completed, Accuracy, and Average Correct Response Time

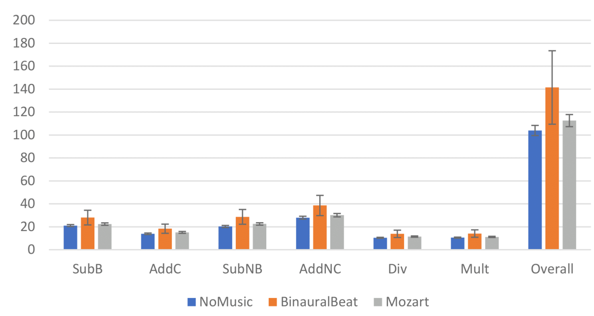
We recorded the three primary performance measures for the Calculation Task. Accuracy is calculated as the number of trials where a participant chose a correct answer to the calculation divided by the total number of trials. Average correct response time is the average of the time taken for a correct trial in seconds. These three measures would successfully represent the calculation performance of the participants during the training.

Figure 8 compares the average of the three performance measures among all participants by conditions. Figure 8 also compares the average value by the calculation problem type. We have six types of calculation problems: addition with carry operation (Labeled as "AddC" on Figure 8), addition without carry operation (Labeled as "AddNC"), subtraction with borrow operation (Labeled as "SubB"), subtraction without borrow operation (Labeled as "SubNB"), multiplication (Labeled as "Mult"), and division (Labeled as "Div").

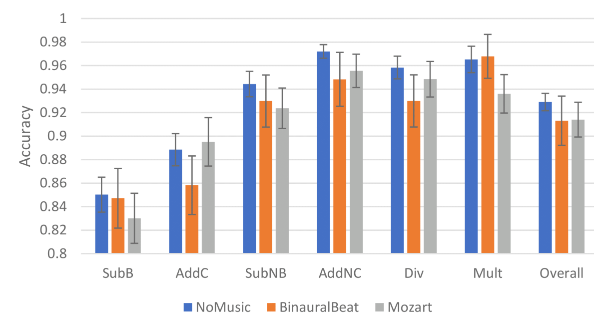
Figure 8

A Bar Graph for the three primary performance measures

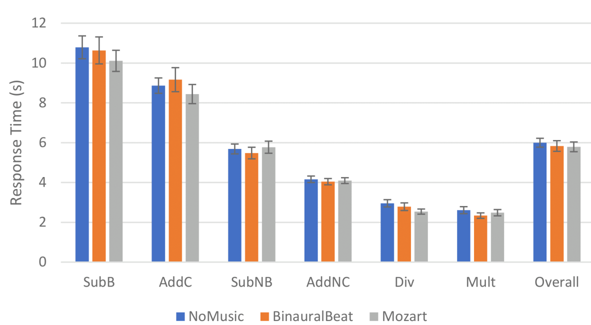
A. Number of Trials Completed



B. Accuracy



C. Average Correct Response Time



By looking at Figure 8 B and C, we can observe that the participants performed better at addition without carry operation than at addition with carry operation because they had higher accuracy =and lower response time. It is likely because addition without carry operation is easier to solve. Similarly, the participants performed better at subtraction without borrow operation than at subtraction with borrow operation. Therefore, we can assume that accuracy and response time successfully represents how well the participants performed on calculation problems on average.

On the other hand, it is difficult to determine if the number of trials completed reflects the participants' performance just by looking at Figure 8 A. Completing a high number of trials on a certain trial group does not mean the participants performed better at the certain trial group; the game was designed to give the participants more number of trials on a certain trial group. Therefore, it is unclear whether a high number of trials completed means better performance or not.

As shown in Figure 8 B, the average accuracy is the highest in No Music condition for all calculation types. However, the difference in accuracy between the conditions is not large, because most error bars are overlapping. Figure 8 A shows that the participants completed the most number of trials in Binaural Beats condition, but Figure 8 C shows that the participants had about the same response time for all conditions. The results are contradictory, and we cannot determine which condition affected the participants' overall performance the most.

We also conducted ANOVA, within-subject ANOVA, and Paired Sample T-tests to compare performance in each condition for the Calculation task. The result of a Paired Sample T-test on the three primary performance measures is presented in the Table 4.

Table 4

Result of Paired Samples T-Test on Number of Questions, Accuracy, and Response Time

	Number of Questions			Accuracy			Response Time		
	t	df	p	t	df	p	t	df	p

SubB(NoMusic)	-	SubB(BinauralBeat)	-1.058	46	0.295	0.144	46	0.886	0.256	46	0.799
SubB(NoMusic)	-	SubB(Mozart)	**-1.801**	**46**	**0.078**	0.957	46	0.343	**1.990**	**46**	**0.053**
AddC(NoMusic)	-	AddC(BinauralBeat)	-1.108	46	0.273	1.218	46	0.230	-0.550	46	0.585
AddC(NoMusic)	-	AddC(Mozart)	**-1.814**	**46**	**0.076**	-0.351	46	0.727	1.023	46	0.312
SubNB(NoMusic)	-	SubNB(BinauralBeat)	-1.285	46	0.205	0.751	46	0.457	0.798	46	0.429
SubNB(NoMusic)	-	SubNB(Mozart)	**-3.211**	**46**	**0.002**	1.342	46	0.186	-0.348	46	0.729
AddNC(NoMusic)	-	AddNC(BinauralBeat)	-1.184	46	0.243	1.068	46	0.291	0.691	46	0.493
AddNC(NoMusic)	-	AddNC(Mozart)	**-1.843**	**46**	**0.072**	1.135	46	0.262	0.543	46	0.590
Div(NoMusic)	-	Div(BinauralBeat)	-1.056	46	0.296	1.313	46	0.196	0.950	46	0.347
Div(NoMusic)	-	Div(Mozart)	**-2.562**	**46**	**0.014**	0.620	46	0.538	**2.727**	**46**	**0.009**
Mult(NoMusic)	-	Mult(BinauralBeat)	-1.073	46	0.289	-0.142	46	0.888	1.405	46	0.167
Mult(NoMusic)	-	Mult(Mozart)	-1.494	46	0.142	1.465	46	0.150	0.716	46	0.478
Overall(NoMusic)	-	Overall(BinauralBeat)	-1.148	46	0.257	0.870	46	0.389	0.774	46	0.443
Overall(NoMusic)	-	Overall(Mozart)	**-2.477**	**46**	**0.017**	1.132	46	0.263	1.329	46	0.190

Note. Student's t-test. The row with p value less than 0.1 is bolded.

While Figure 8 B suggests that the participants had the highest accuracy in No Music condition on average, the result failed to reach significance, as shown in Table 4. Thus, Table 4 suggests that the presence of background music did not significantly affect how accurate the participants solved the calculations.

On the other hand, we found significant differences on the number of problems completed between No Music condition and Mozart condition for multiple types of calculation. Overall, the participants in Mozart condition (M = 112.532, SD = 36.148) completed significantly more trials than the participants in No Music condition (M = 103.93, SD = 30.29); $t(46) = -2.477$, $p = 0.017$. This result suggests that the presence of Mozart music as a background helped people go through a large number of tasks.

The result of the test on average correct response time has a similar trend. For subtraction with borrow operation, participants produced significantly faster responses in Mozart condition (M = 10.108, SD = 3.63) than in No Music condition (M = 10.79, SD = 3.933); $t(46) = 1.990$, $p = 0.053$. The participants also solved division significantly faster in Mozart condition (M = 2.537, SD = 0.884) than in No Music condition (M = 2.953, SD = 1.259); $t(46) = 2.727$, p =

0.009. The result of a Paired Sample T-test on both the number of trials and average correct response time suggests that the presence of Mozart music helped participants to pay attention to and perform better at the calculation tasks.

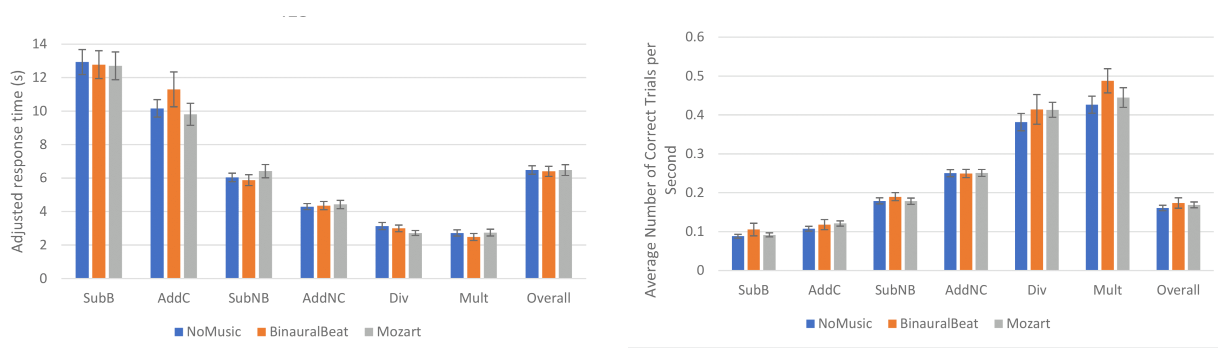
IES and RCS

We also analyzed IES and RCS for the calculation task in order to combine the accuracy and response time statistics. Figure 9 compares the average of the two additional performance measures among all participants by conditions.

Figure 9

Average IES and RCS

A. IES **B. RCS**



According to Figure 9 A and B, the average performance for the trials with complex calculation types (subtraction with borrow operation and addition with carry operation) is worse than the average performance for simple calculation types (division and multiplication). Hence, the average IES and RCS successfully represent the participants' performance during the sessions.

However, we cannot determine which condition affected the participants' IES the most by looking at Figure 9 A. The average IES is the highest in Binaural Beats condition for addition with carry operation, but the average IES is the highest in Mozart condition for subtraction without borrow operation. Figure 9 B suggests that the participants had the highest RCS in Binaural Beats condition in all trial groups. Therefore, the participants produced more correct trials per second in Binaural Beats condition than in two other conditions on average. We also conducted a Paired Sample T-Test on IES and RCS. Table 5 displays the result.

Table 5

Result of Paired Samples T-Test on IES and RCS

			IES			RCS		
			t	df	p	t	df	p
SubB(NoMusic)	-	SubB(BinauralBeat)	0.172	46	0.865	-1.019	46	0.314
SubB(NoMusic)	-	SubB(Mozart)	0.277	46	0.783	-0.844	46	0.403
AddC(NoMusic)	-	AddC(BinauralBeat)	-1.198	46	0.237	-0.775	46	0.442
AddC(NoMusic)	-	AddC(Mozart)	0.863	46	0.392	**-3.170**	**46**	**0.003**
SubNB(NoMusic)	-	SubNB(BinauralBeat)	0.512	46	0.611	-1.021	46	0.313
SubNB(NoMusic)	-	SubNB(Mozart)	-1.152	46	0.255	0.106	46	0.916
AddNC(NoMusic)	-	AddNC(BinauralBeat)	-0.218	46	0.828	0.055	46	0.957
AddNC(NoMusic)	-	AddNC(Mozart)	-0.569	46	0.572	-0.145	46	0.885
Div(NoMusic)	-	Div(BinauralBeat)	0.714	46	0.479	-0.846	46	0.402
Div(NoMusic)	-	Div(Mozart)	**2.326**	**46**	**0.025**	**-1.885**	**46**	**0.066**
Mult(NoMusic)	-	Mult(BinauralBeat)	0.970	46	0.337	**-1.875**	**46**	**0.067**
Mult(NoMusic)	-	Mult(Mozart)	-0.114	46	0.910	-0.927	46	0.359
Overall(NoMusic)	-	Overall(BinauralBeat)	0.273	46	0.786	-0.955	46	0.345
Overall(NoMusic)	-	Overall(Mozart)	0.042	46	0.967	**-1.791**	**46**	**0.080**

Note. Student's t-test. The row with p value less than 0.1 is bolded

According to Table 5, the participants' IES was larger in No Music condition (M = 3.131, SD = 1.477) than in Mozart condition (M = 2.722, SD = 1.018) for divisions; t(46) = 2.326, p = 0.025. The result of a Paired Sample T-test on RCS suggests the same idea. Overall, the participants produced slightly more correct responses per second in Mozart condition (M = 0.169, SD = 0.052) than in No Music condition (M = 0.161, SD = 0.048); t(46) = -1.791, p = 0.080. The presence of Mozart music especially affected participants' RCS positively for additions with carry operation; there is a significant difference in RCS between No Music (M = 0.108, SD = 0.043) and Mozart (M = 0.121, SD = 0.046) conditions; t(46) = -3.170, p = 0.003. The statistical analysis of IES and RCS proves that a large portion of the participants performed better when listening to a music composed by Mozart in the background than when not listening to any music; hence, the data suggests that listening to a music composed by Mozart improves the simple calculation performance.

The effect of condition on the participants' performance was more significant in the Calculation Task than in the Attention Task because we observed more p values below 0.05. For the Attention Task, we only observed two trial groups where a p-value produced by a Paired Sample T-test fell below 0.05: response time when participants had three target fruits and IES when participants had a trial with fruits of mixed colors. However, for the Calculation Task, we observed a total of six trial groups and performance measures where the p-value fell below 0.05: (1-3) number of completed trials for subtraction without borrow operation, division, and overall, (4) response time for division, (5) IES for division, and (6) RCS for addition with carry operation. The result of the Paired Sample T-test shown in Table 4 and Table 5 highlights the significance of classical music on the calculation performance. Therefore, we conclude that the presence of Mozart music might help people focus on the task involving simple calculations.

Memorization Task

For the Memorization Task, one participant was excluded from the final analysis due to unusually low accuracy.

Number of Trials Completed, Accuracy, and Average Correct Response Time

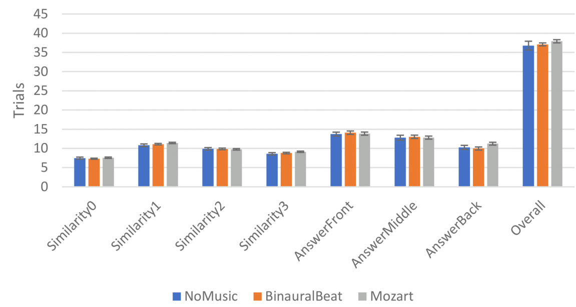
For the Memorization Task, we recorded accuracy, which is computed as the number of trials where a participant selected a correct answer out of four choices divided by the total number of trials. Accuracy represents the participants' visual memorization ability during the session. We also recorded average correct response time and number of trials. Average correct response time is the time taken from when the choice buttons were displayed to when a participant tapped the correct choice button. However, this task focuses on how accurately the participants memorized the images, not how quickly they responded to the problems. While a quicker response time might represent a confidence in their answer, response time does not directly represent how well a participant memorized the images. Therefore, we focus on analyzing accuracy more than the two other performance measures.

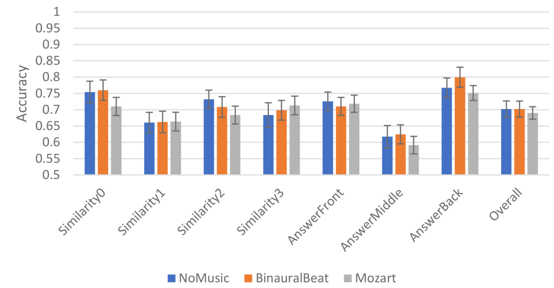
Figure 10 compares the average of the three performance measures among all participants by conditions. Figure 10 also compares the average value by the image similarity. (Labeled as "SimilarityX," where X is replaced by the corresponding similarity value.) We also analyzed the difference in performance by the index of the answer. The bars labeled "AnswerFront" represent the average performance for trials when the participants were asked to recall the image displayed in the first one third of the sequence. The bars labeled "AnswerMiddle" represent the average performance for trials when the participants were asked to recall the image displayed in the second one third of the sequence. Lastly, the bars labeled "AnswerEnd" represent the average performance for trials when the participants were asked to recall the image displayed in the first one third of the sequence.

Figure 10

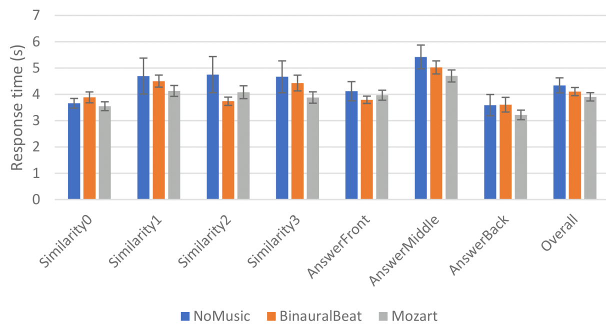
A Bar Graph for the three primary performance measures

A. Number of Trials Completed **B. Accuracy**





C. Average Correct Response Time



By looking at Figure 10 B, we can observe that the participants were likely to memorize the image when the index of the image asked in the problem was located in the first one third of the image sequence or the last one third of the image sequence. The most likely explanation is that a participant is able to memorize the image that appeared first or second better than the fourth or fifth image. Also, a participant is likely to recall the image that appeared just before the question is asked. Therefore, they would perform better when the problem asked to recall one of the first few or last few images. Accuracy seems to represent how well the participants memorized the images.

For image similarity, participants had lower accuracy for the trials with the image similarity value of 1 than for the trials with the image similarity value of 3 on average. This is because lower the similarity value, it is harder to differentiate similar pictures and pick the correct image among other choices. Therefore, we can tell that accuracy successfully represents how well the participants memorized the images from the difference in the measured accuracy between similarity values.

The conditions affected participants' accuracy differently for each trial group. They responded the most accurately in the Binaural Beats condition when the problem asked the participants to recall the image that was shown at the end of the image sequence, but they responded the most accurately in the No Music condition when the problem asked the participants to recall the image that was shown at the beginning of the image sequence. The results are contradictory, and we cannot determine which condition improved or impaired the participants' overall performance.

We also conducted ANOVA, within-subject ANOVA, and Paired Sample T-tests to compare performance in each condition for the Memorization task. The result of a Paired Sample T-test on the three primary performance measures is summarized in the Table 6.

Table 6

Result of Paired Samples T-Test on Number of Trials, Accuracy, and Response Time

			Number of Trials			Accuracy			Response Time		
			t	df	p	t	df	p	t	df	p
Similarity0 (NoMusic)	-	Similarity0 (BinauralBeat)	0.476	46	0.637	-0.352	47	0.727	-0.923	45	0.361
Similarity0 (NoMusic)	-	Similarity0 (Mozart)	-0.333	46	0.741	1.169	47	0.248	0.432	45	0.668
Similarity1 (NoMusic)	-	Similarity1 (BinauralBeat)	-0.876	46	0.385	-0.261	47	0.795	0.293	45	0.771
Similarity1 (NoMusic)	-	Similarity1 (Mozart)	-1.544	46	0.129	-0.336	47	0.739	0.792	46	0.432
Similarity2 (NoMusic)	-	Similarity2 (BinauralBeat)	0.051	46	0.959	0.404	47	0.688	1.373	45	0.177
Similarity2 (NoMusic)	-	Similarity2 (Mozart)	0.354	46	0.725	1.663	47	0.103	1.013	45	0.316
Similarity3 (NoMusic)	-	Similarity3 (BinauralBeat)	-0.551	46	0.584	-0.631	47	0.531	0.345	45	0.731
Similarity3 (NoMusic)	-	Similarity3 (Mozart)	-1.523	46	0.135	-0.702	47	0.486	1.160	45	0.252

AnswerFront (NoMusic)	-	AnswerFront (BinauralBeat)	-0.546	46	0.588	0.169	47	0.867	0.830	46	0.411
AnswerFront (NoMusic)	-	AnswerFront (Mozart)	-0.221	46	0.826	0.159	47	0.874	0.385	46	0.702
AnswerMiddle (NoMusic)	-	AnswerMiddle (BinauralBeat)	-0.318	46	0.752	-0.534	47	0.596	0.837	45	0.407
AnswerMiddle (NoMusic)	-	AnswerMiddle (Mozart)	-0.028	46	0.978	0.674	47	0.503	1.521	45	0.135
AnswerEnd (NoMusic)	-	AnswerEnd (BinauralBeat)	0.389	46	0.699	-1.196	47	0.238	-0.034	45	0.973
AnswerEnd (NoMusic)	-	AnswerEnd (Mozart)	-1.514	46	0.137	0.464	47	0.645	0.927	45	0.359
Overall (NoMusic)	-	Overall (BinauralBeat)	-0.281	46	0.780	-0.339	47	0.736	0.789	46	0.434
Overall (NoMusic)	-	Overall (Mozart)	-0.971	46	0.337	0.487	47	0.628	1.396	46	0.169

Note. Student's t-test.

We did not observe p-values being lower than 0.1 for any of the trial groups. Thus, Table 6 suggests that the presence of background music did not significantly affect how accurate the participants solved the calculations.

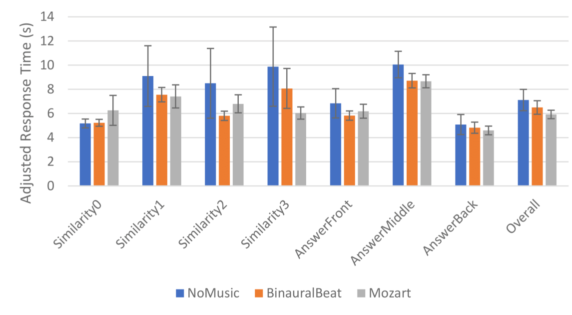
IES and RCS

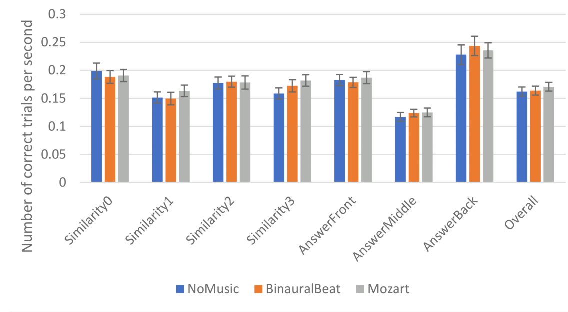
Although we focus on analyzing accuracy for the Memorization Task, we also computed the IES and RCS for all participants for each condition. Figure 11 compares the average of the two additional performance measures among all participants by conditions.

Figure 11

Average IES and RCS

A. IES **B. RCS**





The average score of the participants reflects the difficulty of each trial group. For example, they had a lower IES and higher RCS when the problem asked to recall one of the first few or last few images, meaning they performed better on these tasks. Figure 11 also shows that they performed the best when the similarity value is 0, or when the player did not have to distinguish any similar images. Therefore, IES and RCS successfully represent how well the participants memorized the images.

According to Figure 11 A, the participants seem to have high IES in No Music condition than in two other conditions. However, the error bars for No Music condition are very large, so it is hard to come up with a strong conclusion from the graph.

Similarly, Figure 11 B does not provide strong insight about the data. The participants have the highest RCS in No Music condition for the trials with the similarity value of 0, but the participants have the highest RCS in Mozart condition for the trials with the similarity value of 1. We cannot derive a solid conclusion from Figure 11. We also conducted a paired Sample T-Test on IES and RCS. Table 7 displays the result.

Table 7

Result of Paired Samples T-Test on IES and RCS

	IES			RCS		
	t	df	p	t	df	p

Similarity0(NoMusic)	-	Similarity0(BinauralBeat)	-0.121	45	0.904	0.638	46	0.526
Similarity0(NoMusic)	-	Similarity0(Mozart)	-0.852	45	0.399	0.601	46	0.551
Similarity1(NoMusic)	-	Similarity1(BinauralBeat)	0.616	45	0.541	0.155	46	0.877
Similarity1(NoMusic)	-	Similarity1(Mozart)	0.634	46	0.529	-1.051	46	0.299
Similarity2(NoMusic)	-	Similarity2(BinauralBeat)	0.898	45	0.374	-0.185	46	0.854
Similarity2(NoMusic)	-	Similarity2(Mozart)	0.664	45	0.510	-0.070	46	0.944
Similarity3(NoMusic)	-	Similarity3(BinauralBeat)	0.472	45	0.639	-1.062	46	0.294
Similarity3(NoMusic)	-	Similarity3(Mozart)	1.144	45	0.259	-1.818	46	0.076
AnswerFront(NoMusic)	-	AnswerFront(BinauralBeat)	0.850	46	0.400	0.403	46	0.689
AnswerFront(NoMusic)	-	AnswerFront(Mozart)	0.525	46	0.602	-0.353	46	0.726
AnswerMiddle(NoMusic)	-	AnswerMiddle(BinauralBeat)	1.443	45	0.156	-0.912	46	0.367
AnswerMiddle(NoMusic)	-	AnswerMiddle(Mozart)	1.257	45	0.215	-0.993	46	0.326
AnswerEnd(NoMusic)	-	AnswerEnd(BinauralBeat)	0.272	45	0.787	-0.952	46	0.346
AnswerEnd(NoMusic)	-	AnswerEnd(Mozart)	0.860	45	0.394	-0.483	46	0.631
Overall(NoMusic)	-	Overall(BinauralBeat)	0.688	46	0.495	-0.220	46	0.827
Overall(NoMusic)	-	Overall(Mozart)	1.490	46	0.143	-1.270	46	0.211

Note. Student's t-test.

According to Table 7, there is no statistically significant difference between conditions for any of the trial groups for IES and RCS. We failed to find a significant effect of music in the participants' memorization ability.

Some of the reasons why we could not see a consistent result for the Memorization Task are because of the small number of samples, small number of trials per session, and difficulty of the task. While the participants went through more than 100 trials on average in the Attention Task and Calculation Task, they only completed around 40 trials in the Memorization Task. The Memorization Task required the participants to memorize the image sequences to answer the question, meaning that they had to wait for 10 - 14 seconds for each trial going through the image sequence. This "memorizing time" slowed down the speed that the participants went through the trial, and as a result, the participants could not demonstrate their memorization ability with the few number of trials. The difficulty of the task also takes account of the inconsistent result. On the Post-Training Survey, the participants answered how difficult each of the three tasks was. 2.1 % of the participants answered the Attention Task was “Very Hard” or “Somewhat Hard,” and 29.2 % of the participants answered that the Calculation Task was “Very Hard” or “Somewhat Hard.” On the other hand, 81.2 % of the participants answered that the Memorization Task was “Very Hard” or “Somewhat Hard.” Since the task was hard to achieve high accuracy for most participants, we could not observe the effect of background music on their performance. If we have had a larger sample size, larger number of trials per session, and easier tasks, we would be able to see more consistent results.

DISCUSSION

In the current research study, we investigated the effect of background music on human brain performance by measuring participants' performance on several tasks in multiple conditions using a smartphone application. We designed and developed three tasks: Attention Task, measuring the participants' ability to pay attention to similar objects, Calculation Task, measuring the participants' calculation ability, and Memorization Task, measuring the participants' memorization ability. First, the statistical analysis on the participants' Attention Task performance suggested that there were statistically significant differences in the participants' attentiveness between No Music condition and Binaural Beats condition; the participants took longer response time and IES when they were listening to binaural beats than when they were not listening to any music on average. Secondly, for the calculation ability, the result indicated that there are statistically significant differences in the participants' performance between No Music condition and Mozart condition. The participants solved simple calculations taking less response time when listening to classical music than when not listening to any music. Lastly, we failed to find a significant effect of music in the participants' memorization ability. The research result indicates that the effect of background music on performance depends on the type of activity. Humans appear to pay more attention when not listening to any music. However, humans tend to solve calculation problems quickly when listening to classical music. While there is some evidence of effects of music on performance, the results are inconsistent that more research will be required to both confirm and to better understand these findings.

The effect of music on the participants' performance observed in this research study is somewhat consistent with the findings from past research studies. We discovered that there is a significant effect of background music on the participants' performance on solving simple calculation problems or paying attention to multiple objects. These findings support the conclusion derived from multiple research studies mentioned previously, including Thomas Schäfer's (Schafer, 2013), Demetriou et al.'s (Demetriou, 2016), and Frances H Rauscher's (Rauscher, 1993) research studies. In Particular, we discovered that humans tend to solve simple calculations more quickly when listening to classical music than when not listening to any music. The positive effect of classical music is also supported by a research study conducted by Rauscher (Rauscher, 1993). On the other hand, we also identified the negative effect of binaural beats on the participants' attention ability. This result contradicts with Monroe's research that discovered the benefits of binaural beats on human brain cognition (Filimon, 2020). However, our p-values produced from the Paired-Sample T tests are not very large and thus while our research supports some claims, our conclusion is not as strong as the claims supported by other studies.

Additionally, we also investigated if the participants' performance is affected by their music preference. After each session, the participants completed a survey with the following questions:

1. When you were listening to classical music, did you find music distracting?
2. When you were listening to binaural beats, did you find music distracting?
3. If you could redo the task without any music, would you choose to listen to music?

Based on the responses to the above questions, we analyzed whether each participant found each type of music distracting or not for each task. Then we ran a MANOVA test (Multivariate Analysis of Variance test, a variance of ANOVA test that handles multiple dependent variables) to study the correlation between their music preference and their performance in No Music condition. As a result, the statistical analysis did not provide evidence of a consistent effect of music preference on performance, because the collected data produced relatively high p-values. Therefore, the result failed to reach significance and was not able to support the correlation between the participants' performance and whether they found a certain type of music distracting or not. Graphs showing performance of each participant in relationship to their music preference are displayed in Appendix A for reference.

We successfully designed our smartphone application to observe the participants' performance during the game play. All tasks are straight forward, so each task accurately reflects how well the participants paid attention, solved calculation problems, or memorized the images during the training. The application also effectively gives each participant the tasks of equal difficulty. For example, during the Calculation Task, the application gives a certain number of additions, subtractions, multiplications, and divisions per session, making the difficulty of the Calculation Task similar for all participants. Therefore, we succeeded to obtain an accurate performance measure from each participant for each condition. One thing we could have changed is that, as mentioned in the Result section, we could have designed the Memorization Task easier. Many participants found the task harder than the other two tasks, so their performance did not reflect the effect of the background music condition as successfully as other tasks.

The way we designed the research procedure effectively measured the participants' performance. We were especially glad that we were able to collect data and make progress on the research during the Spring 2020 campus closure. The training sessions did not involve any in-person assistance, so we were able to run research remotely despite the campus closure. We also effectively designed the training schedule for each participant based on Latin Squares in order to minimize the order effect. On the other hand, there are few things we could have improved about the research protocol. First, as mentioned above, we could have run the research with the larger number of samples in order to have more accurate mean values and smaller errors. Second, if all training was done in person, we could have obtained more accurate performance measures. Research assistants were unable to detect any unwanted noises during the remote training, so there may be an environmental factor that affected the performances. Also, for the Calculation Task, we had to exclude one participant from the final analysis due to an unusual performance; while most participants solved around 100 - 150 trials per session, the excluded participant completed more than 1,000 trials within one session with significantly low accuracy. We assumed that he guessed all calculation questions instead of solving them. If a research assistant was with him in person during the training, the research assistant could have noticed his strategy and asked him to work on the tasks property. However, overall, we succeeded in running the experiment smoothly.

If we were to run this experiment again, we could improve the task design and research procedure as mentioned above. We could adjust the task difficulty so the performance accurately reflects the participants' ability at a moment. We could have more subjects to obtain more accurate mean values. We could also have all participants go through training in person to ensure they play the game in the same environment. Although the conclusions did not retain as much significance as other research studies, this research study still provided us a good insight into the method to investigate the effect of music on the human brain cognition.

CONCLUSION

This research study is conducted to examine the effect of background music on human brain performance by designing and developing a smartphone application using a game development engine called Unity and writing code in C#, that gives the player three simple tasks: Attention Task, Calculation Task, and Memorization Task. We recorded the participants' performance in three different conditions: when not listening to any music, when listening to a classical music, and when listening to a binaural beat sound. We ran a statistical analysis on the recorded data to investigate if the background music helped them perform better or not. As a result, the recorded data indicates that music distracts humans from paying attention to objects, but listening to classical music encourages humans to solve calculation problems quickly. Despite this, the research failed to derive a significant conclusion about the correlation between music and cognitive task performances. Future studies should solidify these results by adjusting the task difficulty and incorporating larger sample size. More conclusive findings, especially those that address individual differences, could help our community create an environment that allows people to effectively work on their tasks at the office or library.

REFERENCES

Au, Jacky, et al. “Effects of Non-Symbolic Arithmetic Training on Symbolic Arithmetic and the Approximate Number System.” Acta Psychologica, https://www.journals.elsevier.com/acta-psychologica.

Demetriou, Andrew M, et al. “Go With the Flow : When Listeners Use Music As Technology.” *ResearchGate*, Jan. 2016, https://www.researchgate.net/publication/326711518_Go_With_the_Flow_When_Listeners_Use_Music_As_Technology.

Diamond, A., Barnett, W.S., Thomas,J. & Munro, S. (2007). Preschool Program Improves Cognitive Control. SCIENCE, 318, 1387-1388.

Filimon, Rosina Caterina. “Beneficial subliminal music: binaural beats, hemi-sync and metamusic.” *Semantic Scholar*, 2010, http://www.wseas.us/e-library/conferences/2010/Iasi/AMTA/AMTA-18.pdf.

Jäncke, Lutz, and Pascale Sandmann. “Music Listening While You Learn: No Influence of Background Music on Verbal Learning.” *Behavioral and Brain Functions*, vol. 6, no. 1, 7 Jan. 2010, p. 3., doi:10.1186/1744-9081-6-3.

Magnetic Minds. Accelerated Learning - Gamma Waves for Focus, Memory, Concentration - Binaural Beats - Focus Music. YouTube, 23 May 2017, www.youtube.com/watch?v=vLEek3I3wac.

Rauscher, Frances H., et al. “Music and Spatial Task Performance.” *Nature*, vol. 365, no. 6447, 14 Oct. 1993, pp. 611–611., doi:10.1038/365611a0.

Rousseau. Mozart - Piano Sonata No. 16 in C Major, K.545 (1st Mvt). YouTube, 24 Jun 2019, https://www.youtube.com/watch?v=qjk-YRuQZDE.

Schäfer, Thomas, et al. "The Psychological Functions of Music Listening." *Frontiers in Psychology*, vol. 4, 13 Aug. 2013, https://doi.org/10.3389/fpsyg.2013.00511.

Stark, Craig. "MST (BPSO)." *Stark Lab*, UCI BioSci, 12 Jan. 2018, https://faculty.sites.uci.edu/starklab/mnemonic-similarity-task-mst/.

Verrusio, Walter, et al. "The Mozart Effect: A quantitative EEG study." *Science Direct*, Feb. 2015, https://www.sciencedirect.com/science/article/abs/pii/S1053810015001130?via%3Dihub.

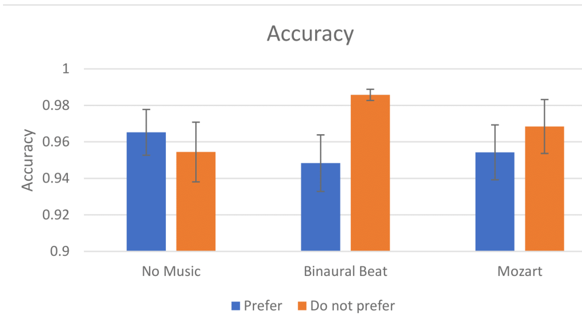
APPENDICES

Appendix A: The average performance of the participants by their music preference

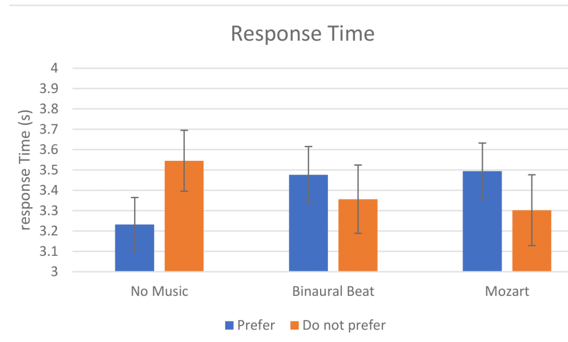
Figure 12

A Bar Graph for Attention Task performance by music preference

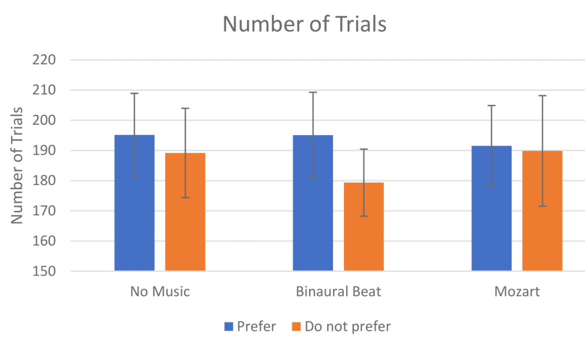
A. Average Accuracy



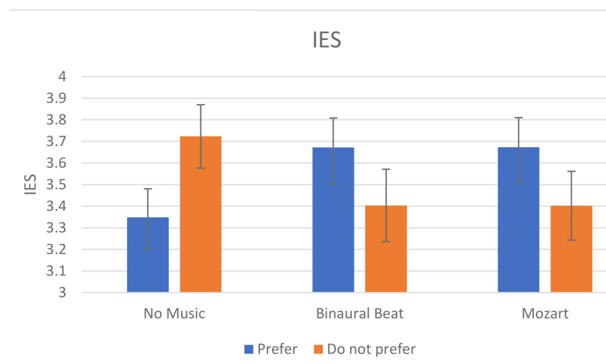
B. Average Response Time



C. Average Number of Trials Completed



D. Average IES



E. Average RCS

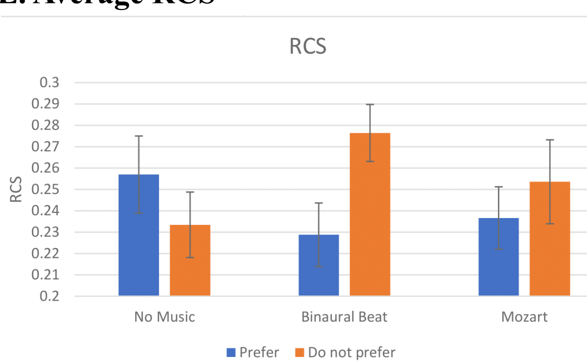
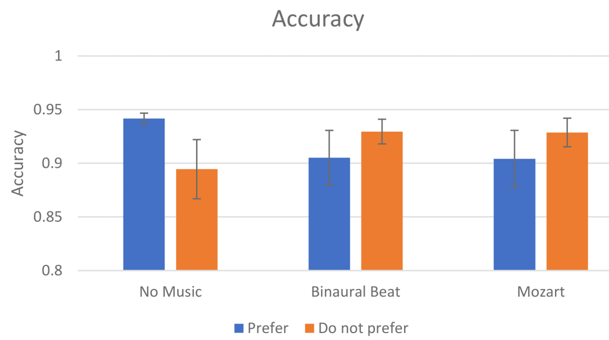


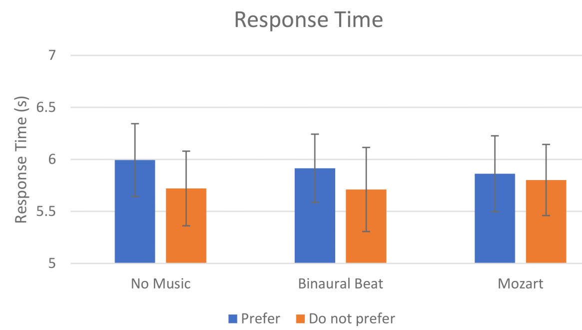
Figure 13

A Bar Graph for Calculation Task performance by music preference

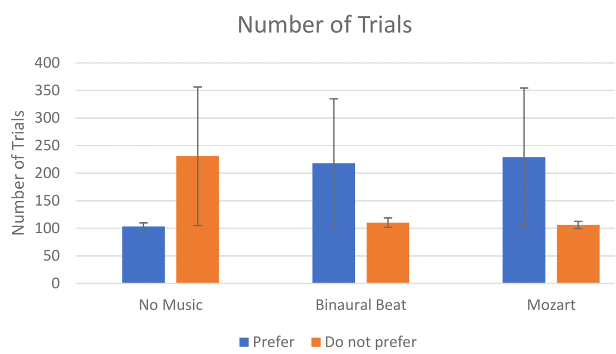
A. Average Accuracy



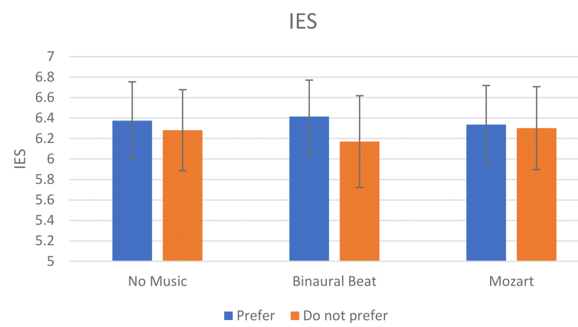
B. Response Time



C. Average Number of Trials Completed



D. Average IES



E. Average RCS

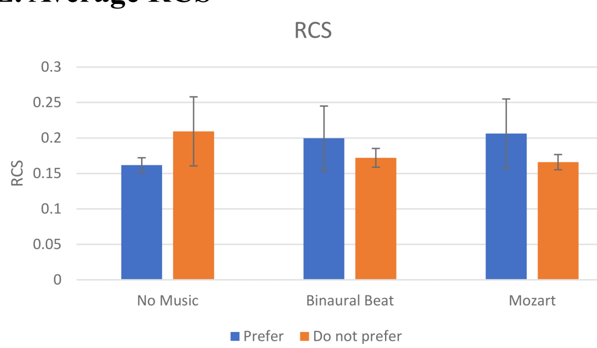
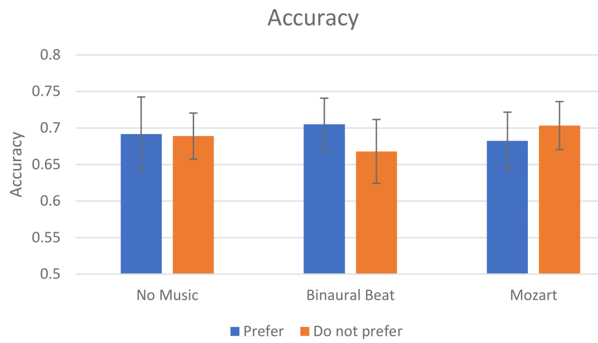


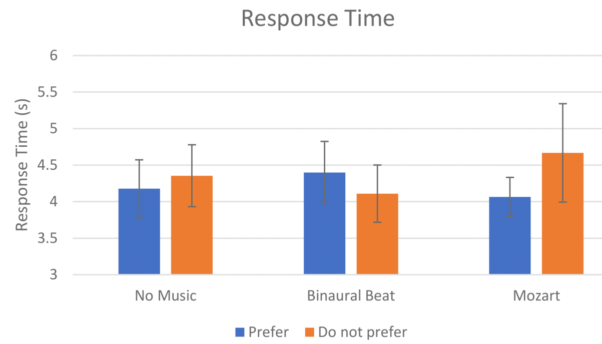
Figure 14

A Bar Graph for Memorization Task performance by music preference

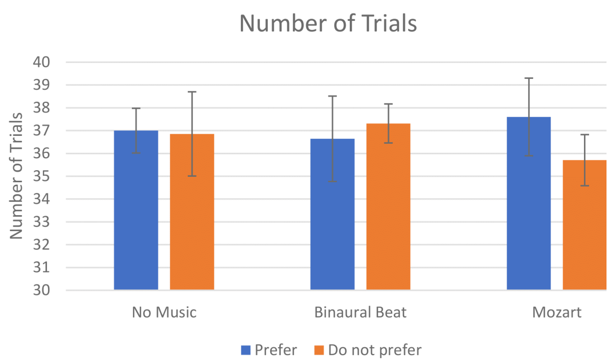
A. Average Accuracy

B. Average Response Time

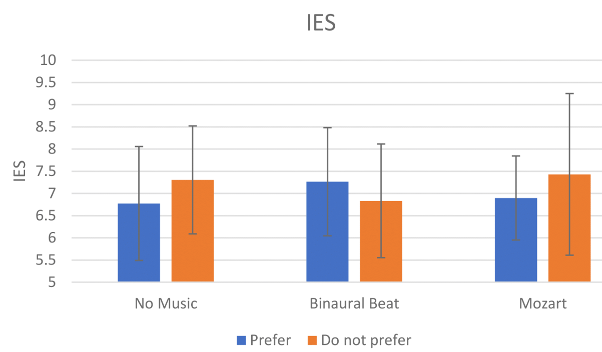




C. Average Number of Trials Completed



D. Average IES



E. Average RCS