2023 S.T. Yau High School Science Award

Research Report

The Team

Name of team member: Ximurong Zhai

School: Bishop Strachan School

City, Country: Toronto, Canada

Name of supervising teacher: Andrew McLeod

Job Title: Head of Science and Senior School Science Teacher

School/Institution: Bishop Strachan School

City, Country: Toronto, Canada

Title of Research Report

Intervention of Age-Related Memory Alteration by Integrative Musical Stimuli

Date

08/18/2023

Intervention of Age-Related Memory Alteration by Integrative Musical Stimuli

Ximurong (Jasmine) Zhai

Abstract

Age-related memory impairment often causes significant distress for the elderly, imposing substantial burdens on families and society. It is widely known that auditory stimulation can trigger certain emotions and brain activities, but studies about its effect on age-related memory impairment are lacking. This study aims to investigate the potential of music as an intervention for age-related memory impairment and the varying effects of different music genres on memory abilities. The research involves 22 participants and evaluates memory performance using the ten-word test (TWT) and digit span (DS) test. Additionally, raw electroencephalogram (EEG) data is analyzed to examine the impact of music on the cortical electrical activity of the brain. I created a random forest machine learning model to predict long-term memory scores based on short-term outcomes. We found that, on average, classical music stimuli were effective for improving long-term TWT and DS scores, and folk music stimuli effectively improved long-term TWT scores. Following folk music treatment, participants aged 65 and above yielded TWT score enhancements ranging from 25% to 66.7% and DS score enhancements ranging from 20% to 50%. I also discovered the effectiveness of utilizing the random forest machine learning model to predict long-term TWT and DS outcomes based on short-term results, potentially enhancing the efficiency of future treatment evaluation.

Keywords: *Musical stimuli, age-related memory impairment, electroencephalogram (EEG), Alzheimer's disease (AD), mild cognitive impairment (MCI), machine learning, random forest algorithm*

Declaration of Academic Integrity

The participating team declares that the paper submitted is comprised of original research and results obtained under the guidance of the instructor. To the team's best knowledge, the paper does not contain research results, published or not, from a person who is not a team member, except for the content listed in the references and the acknowledgment. If there is any misinformation, we are willing to take all the related responsibilities.

Names of team members Ximurong Zhai

Signatures of team members Ximurong Zhai

Name of the instructor Andrew McLeod

Signature of the instructor

Date Aug. 8, 2023

Table of Contents

1. Introduction

1.1 Background of Cognitive Impairment and Dementia

The morphology and function of the brain are significantly altered by aging[1] (**Figure 1**). Memory loss and shifts in behavior are all cognitive deficiencies that are often accompanied by these age-related changes in the brain[1]. It is found that about 40% of elders aged 65 or above experience age-related memory impairment, and about 10% exhibit mild cognitive impairment (MCI), which is more severe memory loss[2]. Each year, 10-20% of individuals with MCI transition to dementia, which is defined as acquired progressive cognitive impairment severe enough to affect daily activities[3,4]. Dementia is strongly associated with brain aging and is one of the leading causes of disability, dependency, and mortality[4], with elders aged 65 or above being the most vulnerable[5]. It is estimated that 55 million individuals worldwide currently suffer from dementia[2].

Among individuals with dementia, Alzheimer's disease (AD) accounts for over two-thirds of cases in adults aged 65 and above[6]. The symptoms of AD progresses gradually. Early learning and memory problems, as well as subsequent impairments in complex attention, executive function, spatial function, language, and social behavior, are characteristics of typical AD cases[7]. The prevalence of AD increases over age (**Figure 2**), and the rising cases of elders diagnosed with AD continue to raise the burden on families and public health, with AD-related healthcare costs reaching nearly $500 billion annually[8]. By 2050, it is estimated that the cost of treating AD and dementia in the United States will pose an even heavier burden on the healthcare system, reaching a total of $1 trillion[9].

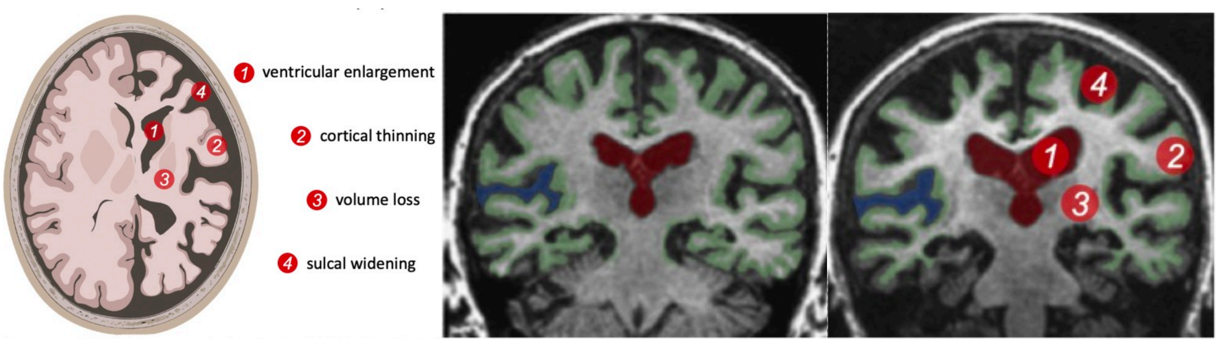


Figure 1. Hallmarks of age-related cerebral atrophy. Morphological changes include (1) ventricular enlargement, (2) cortical thinning, (3) gray and white matter volume loss, and (4) sulcal widening[1].

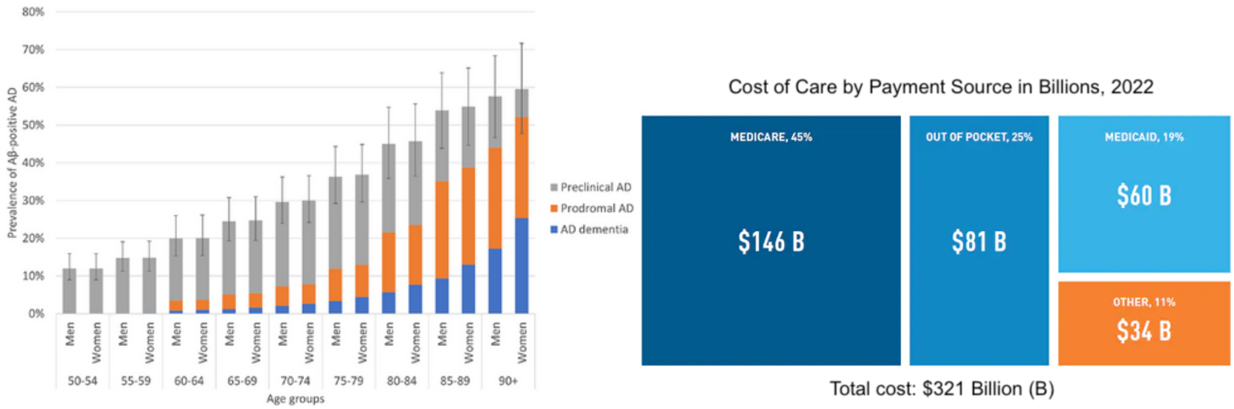


Figure 2. Age-specific prevalence of Aβ-positive AD (left) and economic burden (right). Data is estimated by age, sex, and stage, and whiskers represent uncertainty ranges[10]. The distribution of payment costs for AD and dementia for Americans aged 65 or above in 2022 (right)[9].

1.2 Current Treatment and Research of Age-Related Cognitive Impairment and Dementia

There is no cure for MCI or AD as of now. The most widely-used symptomatic treatment is acetylcholinesterase inhibitors (AChEIs), such as donepezil, galantamine, and rivastigmine, which prevent the breakdown of acetylcholine in the synapses. However, AChEIs should be administered with caution in individuals with heart conduction abnormalities due to the possibility of experiencing irregularly slow heart rates. On the other hand, anti-NMDA medications such as memantine have the potential to prevent glutamate-mediated neurotoxicity from developing as AD progresses with increased neuron death[7]. However, both AChEIs and anti-NMDA medications have not been found to alter the progression of the disease or the rate of memory and cognitive function deterioration[4] (**Figure 3**, line 3). Other therapies such as aducanumab and lecanemab directed Aβ amyloid plaques using anti-Aβ monoclonal antibodies[11]. Aducanumab is a recently approved, disease-modifying treatment for mild AD, but it has been controversial because its impact on AD-related symptoms remains unproven. Lecanemab is another recently approved, promising treatment, and clinical trials showed that it reduced the rate of memory and cognitive decline over 18 months while reducing brain amyloid markers[12]. However, lecanemab is only significantly effective in early AD cases.

In addition to medication, physical therapy can mitigate both cognitive and motor impairments in individuals experiencing cognitive impairment and AD[13,14]. Specifically, it brings improvement in gait, balance, and cognition[15]. Music stimulates the brain to provoke a response, and it is a non-invasive, accessible technique that could be frequently utilized. It has been found that music-based therapy improved the rehabilitation compliance of stroke patients while alleviating negative emotions at the same time[16]. Other studies that investigate the effect of musical stimuli on stroke recovery suggest that music could often activate the auditory and motor systems, which indicates its effectiveness in assisting with stroke rehabilitation[17]. Studies also suggest that memory traces built through music are more deeply established against neurodegenerative memory impairments such as AD[18]. Since music-associated memories are better preserved in AD[19], music therapy can be useful in helping individuals retrieve autobiographical memories and emotions associated with music[20]. It is also found that the caudal anterior cingulate cortex and the ventral pre-supplementary motor area (pre-SMA) are involved in musical memory and music-evoked emotions, and they are one of the last brain regions to experience impairment in AD. Since these regions remain for a longer time, it might explain why AD patients usually have well-preserved musical memory. Other studies used personalized musical training to investigate the effectiveness of favored music on brain functional connectivity[21]. With the use of functional magnetic resonance imaging, it was found that preferred musical stimuli activated the SMA, along with increasing functional connectivity.

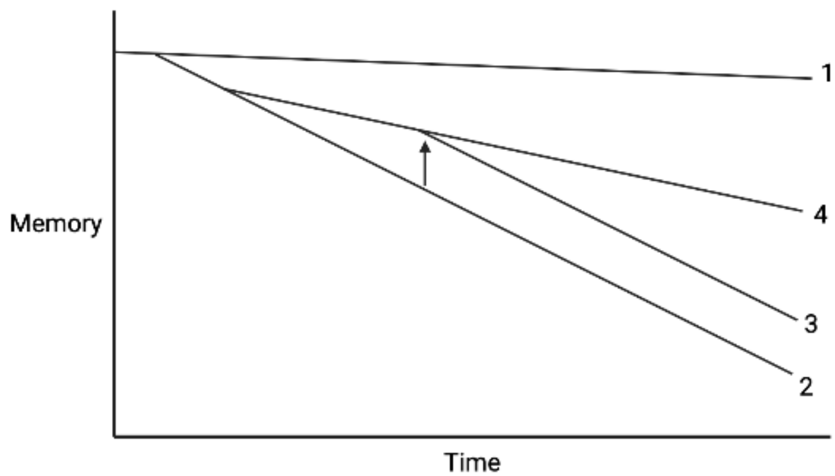


Figure 3. Rate of memory decline under different conditions over time. (1) Slow memory decline in normal aging. (2) Rapid, earlier memory decline in AD. (3) Most current therapies that do not change the course of memory decline, only capable of enhancing cognition. (4) Reduced rate of memory decline as an anticipated effect of therapies[8]. Prepared by biorender.

Aside from musical stimuli, certain auditory frequencies were found to have an impact on the pathological markers of AD. Several published studies by the research group led by Li-Huei Tsai found that 40 Hz stimulation could decrease the harmful Aß amyloid plaques and tau tangles in mice, which are hallmarks of AD[22,23]. 40 Hz stimulation also alters the state of microglia, which are brain cells involved in synapse loss while exacerbating tau pathology and harming neurons by releasing inflammatory factors[22-25].

Although music therapy does not directly target a pathological marker, it has certain advantages. In addition to its non-invasive characteristic, the low cost and high accessibility of music therapy have made it become increasingly accepted by individuals seeking to improve neurological conditions. It is also well-known that different types of music can trigger distinctive emotions and affect our emotional well-being. Therefore, music therapy could potentially be a highly powerful way to improve the moods of patients, especially individuals with MCI and AD whose impaired emotional processing has been linked to agitation, unsettledness, aggression, and other neuropsychiatric symptoms[26].

1.3 Rationale of this Study

While symptomatic treatments exist, we have yet to discover a treatment for MCI or AD that would slow the course of memory decline over the long term (**Figure 3**, line 4). Given this gap and the numerous advantages of music therapy for alleviating memory and cognitive impairment, I think that music may offer us a fresh, accessible, and non-invasive approach to treating these disorders. This study investigates the short and long-term impacts of musical treatment on verbal episodic and working memory as well as the effects of different music genres.

1.4 Innovativeness of this Study

1. Music is an accessible, non-invasive, and sustainable potential supplement to AD treatments that could be used in the long term. It could provide symptom-relieving as well as emotional value. In addition, this investigation quantitatively assesses the effects of different music genres and would give insight into potential personalized music therapy treatment options for memory improvement.
2. This study combines advanced machine learning technology, employing the random forest algorithm to predict long-term treatment outcomes using short-term data. In addition, both portable EEG and machine learning used in this study are prominent subjects in modern neuroscience.
3. This study involves the creation of a mobile application that provides personalized music recommendations, memory change quantification, and future treatment outcome prediction functions for users seeking to improve their memory, which makes treatment portable and easily accessible.

2. Methods and Materials

2.1 Experiment Objectives

1. To investigate the impact of musical stimuli on verbal episodic and working memory, indicated by the ten-word test (TWT) and digit span (DS) test, and observe the effect of different music genres.
2. To utilize machine learning methodologies to explore a potential way of enhancing the efficiency of future treatment through predicting long-term outcomes from initial short-term results.
3. To examine the effect of musical stimuli on brain cortical electrical activity across distinct regions through the analysis of EEG data.

2.2 Experiment Procedure

This experiment was conducted in short-term one-week (**Figure 4**) and long-term three-week cycles (**Figure 5**). The music genre for each participant changes after the end of each cycle. Subjects listen to music at the same time of day in the same surrounding environment throughout the experiment. TWT and DS tests are conducted following music stimuli, and raw EEG data is collected before and after subjects listen to music, which is used to observe the impact of music on regional brain cortical activities (**Figure 6**).

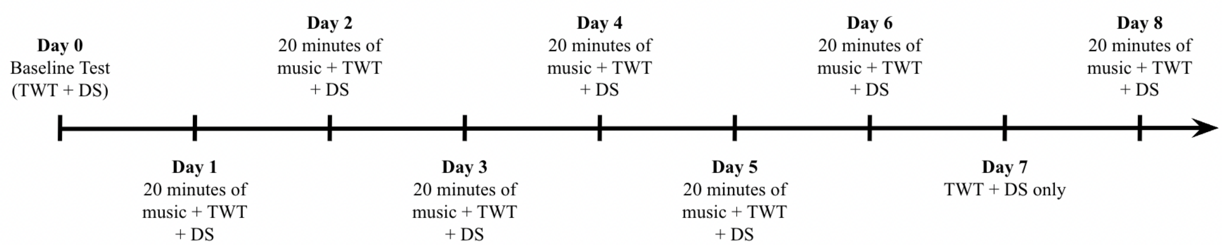


Figure 4. One-week cycle timeline. Day 0 is the baseline test, and participants listen to 20 minutes of music every day, followed by a TWT and DS test. The exception is day 7, during which participants do not listen to music, serving as a control day within the experimental trial.

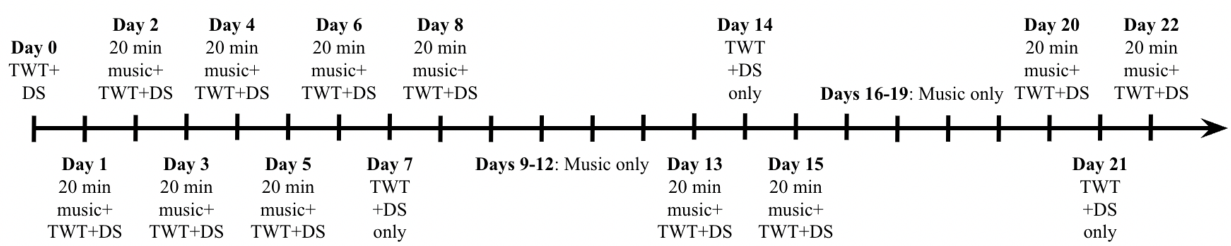


Figure 5. Three-week cycle timeline. Day 0 is the baseline test, and participants listen to 20 minutes of music every day, followed by a TWT and DS test for three weeks. The days during which participants receive no music treatment are days 7, 14, and 21.

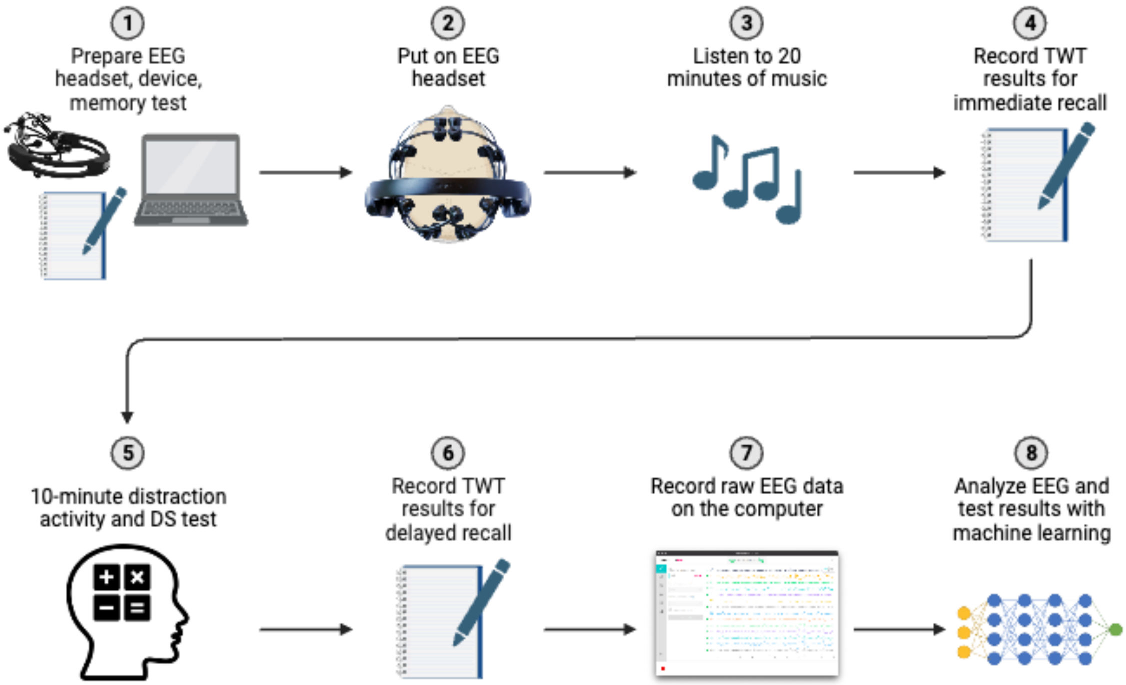


Figure 6. Experiment procedure. Participants listen to music with the EEG headset, followed by the TWT and DS tests. Test results and raw EEG data are then collected and analyzed alongside machine learning methods. Figure prepared by biorender.

2.2.1 Ten-Word Test

The TWT presents a higher level of difficulty relative to the five-word test (FWT), a rapid evaluation of verbal episodic memory[27,28]. The original FWT was an effective tool for the quick screening of AD, assessing the ability of subjects to recall a short list of words belonging to different semantic categories. These semantic categories – building, insect, drink, vehicle, and kitchen utensil – are strategically used to enhance the ability of the subject to apprehend and retrieve the given words.

Differing from the original FWT, the TWT encompasses several distinct features, including a weighting difference between spontaneous and cued recall[28], an attention interference duration of 10 minutes, and an expanded word list containing 10 items.

A step-by-step procedure of the TWT is provided below:

- Show a list of 10 words, every 2 words belonging to a different semantic category, to the subject and confirm that they have understood all the words clearly.
- Ask the subject to repeat the 10 words immediately. For the words not recalled spontaneously, ask "What was the name of the…?", providing the corresponding semantic category as a clue.
- Take note of the spontaneous recalls (2 points) and cued recalls (1 point).
- Show the unremembered words and ask the subject to provide them in response to the corresponding semantic category.

- Divert the attention of the subject for 10 minutes by asking about numerical calculations, and make sure that their attention is no longer focused on the words.
- Ask the subject to repeat the 10 words again, taking note of the delayed recall score. For the words not recalled spontaneously, ask "What was the name of the…?", providing the corresponding semantic category as a clue.
- Take note of the spontaneous recalls (2 points) and cued recalls (1 point).
- Calculate the sum of the immediate recall and delayed recall scores as a total score out of 40 (**Table 1**).

Table 1. Example TWT result tracker

Item	Semantic Group	Immediate recall		Delayed recall	
	"What was the name of the…"	Spontaneous (2pts)	Cued (1pt)	Spontaneous (2pts)	Cued (1pt)
Item 1	Building				
Item 2	Insect				
Item 3	Drink				
Item 4	Vehicle				
Item 5	Kitchen utensil				
Item 6	Building				
Item 7	Insect				
Item 8	Drink				
Item 9	Vehicle				
Item 10	Kitchen utensil				

2.2.2 Digit Span Test

In this study, the DS test is performed by verbally providing a series of numbers and asking the subject to repeat them both in the presented and inverse order[29-31]. This assesses the verbal working memory of participants[31] and the ability to recall numeric information. At the end of each test, the total number of recalled numbers is calculated.

A step-by-step procedure for the DS test is provided below:

- Verbally present a list of 6 double-digits at a rate of 1 double-digit per second. Digits are chosen randomly and do not appear in regular ascending or descending order.
- Immediately ask the subject to recall the series in the presented order, then the inverse order.
- Record the score as the number of correct double-digits recalled in both orders out of 12.

2.2.3 Machine Learning

The random forest supervised machine learning algorithm comprises a collection of random decision trees that are less sensitive to changes in training data. Through the process of bootstrapping and aggregation, the trees are trained with different, randomly selected datasets and features to create diverse trees and decision nodes (**Figure 7**). Predictions in my experiment stem from the mean value from the results of multiple decision trees, and this low inter-tree correlation increases the likelihood of yielding an accurately predicted outcome.

I employed the random forest machine learning algorithm with the goal of creating a model to potentially enhance the efficiency of future music treatment. Instead of undergoing a 3-week investigation to determine the long-term impacts of a certain genre of music on TWT and DS scores, I built a model capable of predicting long-term outcomes from the initial 7-day dataset. This approach enables the possibility of evaluating the effectiveness of a particular treatment genre in advance and reducing the time demands of trial-and-error processes.

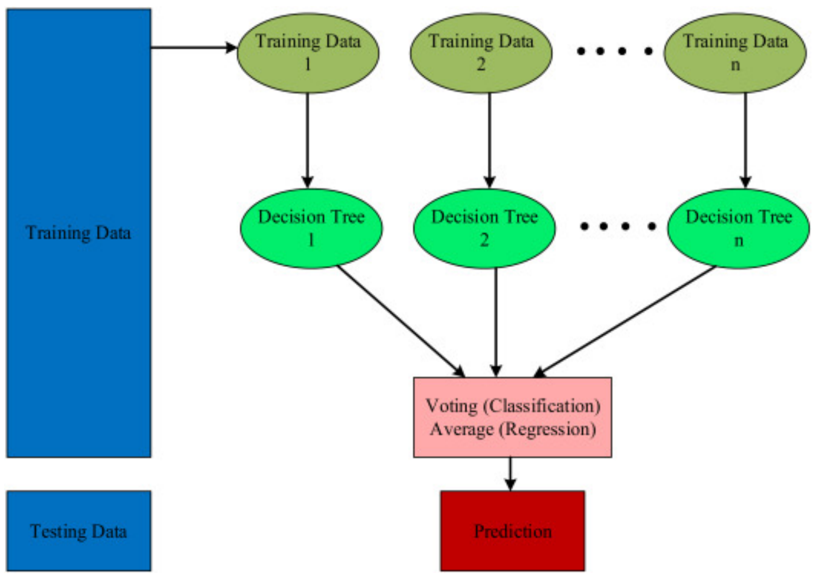


Figure 7. Random forest algorithm step-by-step visualization[32]. My model completes a regression task; therefore, the prediction is the average of decision tree outcomes.

2.2.4 EEG Data Collection

Subjects wear an EEG headset (**Figure 8**), and data on cortical electrical activity is collected before and after musical treatment. The data collected could potentially provide insights into the impact of music on brain activities. I first soak 16 sensor felts in saline solution, then squeeze out excess fluid and secure one sensor felt in each sensor. The EEG device is then connected to the computer using a USB cable, and I ensure the contact and EEG quality are optimal and the participant feels comfortable with the headset position before raw EEG data is collected (**Supplementary Figure 1**).

To test the precision of the EEG, I conducted an additional experiment to distinguish between signals from muscle movement and brain waves. Participants first remained still with the EEG on, then moved their eyebrows, and finally thought about moving their eyebrows without moving them. Next, participants moved their arms and then thought about moving their arms without moving them. By observing the raw EEG data while doing these movements, we could find out whether the EEG captures muscle movement only or both muscle movement and motor neuron electrical potential.

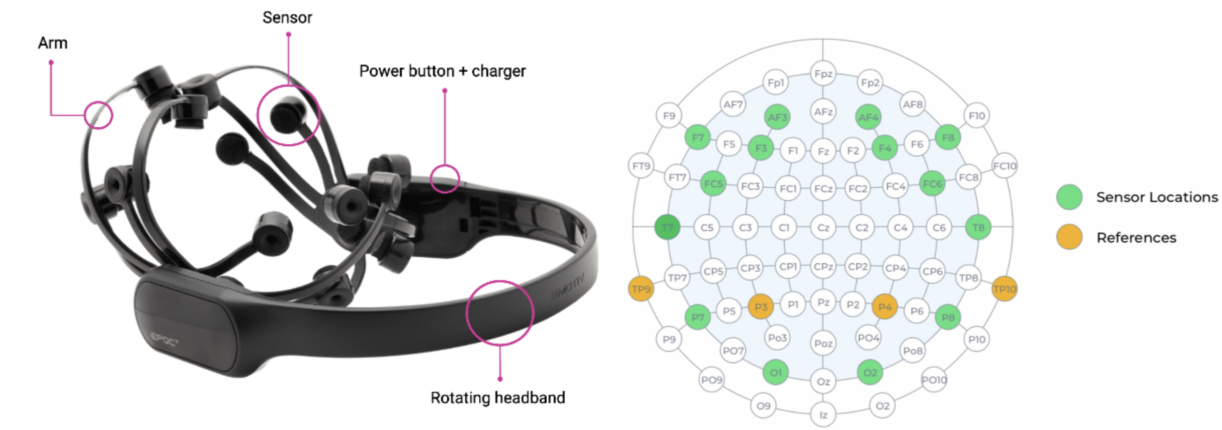


Figure 8. EEG headset (left) and sensor & reference locations (right). The different parts of the EEG headset are labeled. The 14 sensors and 2 references are in contact with the scalp, each collecting data from a different area of the scalp. The headband is rotatable to ensure a comfortable fit for each participant (left). Figure prepared by biorender. The locations of the 14 sensors are labeled in green, and the reference location options, which ensure the headset is in the correct position, are labeled in orange (right)[33].

2.3 Music Genre Classification

After initial research on scientific music classification, I made sure that all favored music pieces of participants fall under the following genres: classical, country, folk, jazz, blues, pop, disco, rock, and metal[34]. To ensure that subjects are familiar with the music genres, I provided them with excerpts from different music genres and asked them to associate one with each genre, and made modifications based on familiarity feedback. The negative control involves no musical or auditory stimuli and 40 Hz music is a potential positive control.

3. Results

3.1 Participant Preference Distribution of Music Genres

I recruited 22 subjects from my local community and hometown, including 14 males and 8 females ranging from the age of 9 to 84 (**Supplementary Table 1**). The first cycle involved 17 participants, and the second cycle involved all 22 participants. After listening to excerpts from each music genre, most subjects had trouble distinguishing between jazz and blues, pop and disco, and rock and metal, while it is easier to distinguish between classical, country, and folk music. Therefore, I decided to include blues in jazz, disco in pop, and metal in rock. 40 Hz music is listed as a potential positive control. The music preference levels of participants are also recorded (**Figure 9**; **Supplementary Table 2**). Most subjects reported experiencing more positive emotions after listening to their preferred music and no clear emotional change after listening to unpreferred music. Some subjects reported experiencing fewer positive emotions after listening to unpreferred music, and others reported feeling no emotional change regardless of the music genre.

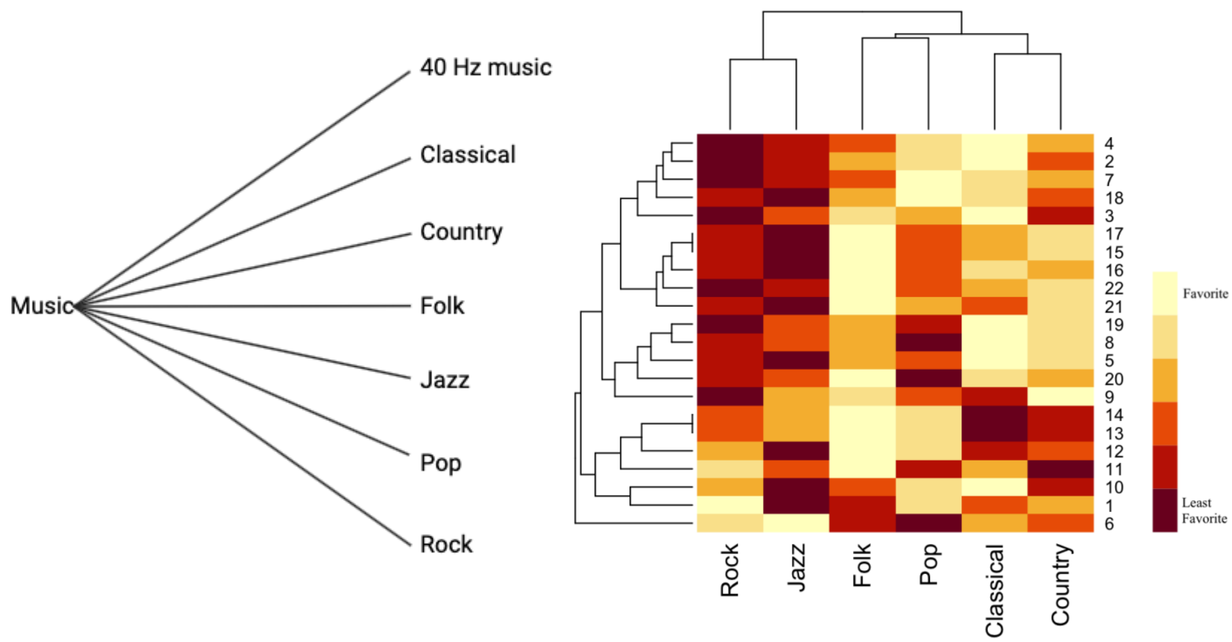


Figure 9. Classification of music (left) and heatmap showing distribution of music genre preference of subjects (right). 40 Hz music is the potential positive control. The lighter the color, the higher the preference for a certain genre of music.

3.2 TWT and DS Short-Term and Long-Term Results

TWT and DS results (section 3.2.1) are presented separately, both with a score vs. time graph showing individual changes, an average score changes vs. time graph showing averaged changes of different music genres relative to the baseline, and a box plot comparing percentage changes across genres. The score vs. time graphs show results from the first cycle of 17 participants, while the average score changes vs. time graphs and box plots show results from both the first and second cycles, including 22 subjects in total. To facilitate the efficiency of future musical treatment evaluation, I employed the random forest machine learning algorithm (section 3.2.2) to create a model that predicts long-term (day 22) TWT and DS results based on short-term treatment data (first 7 days).

3.2.1 Short and Long-Term Changes of TWT and DS Scores

Unsurprisingly, TWT results from the first cycle show high data variability and marginal differences in the one-week short term (**Figure 10**). Nonetheless, long-term results show that 94% (16/17) of all subjects demonstrated higher TWT scores at the end of the cycle compared to their baseline score. In particular, 76% (13/17) of subjects had a 20% or higher score increase, and 24% (4/17) of subjects had a 50% or higher score increase (**Table 2**). Subjects aged 65 and above showed score improvements ranging from 25% to 66.7%. Overall, the baseline and final TWT scores are lower for older participants, but there is a major score increase for participants over the age of 80. The reduction in TWT scores on days 7, 14, and 21 (corresponding to the days without music) indicates the importance of maintaining music therapy consistently. Since there was only one individual who listened to rock, jazz, pop, and country music during the first cycle, we cannot come to definitive conclusions solely based on those singular scores.

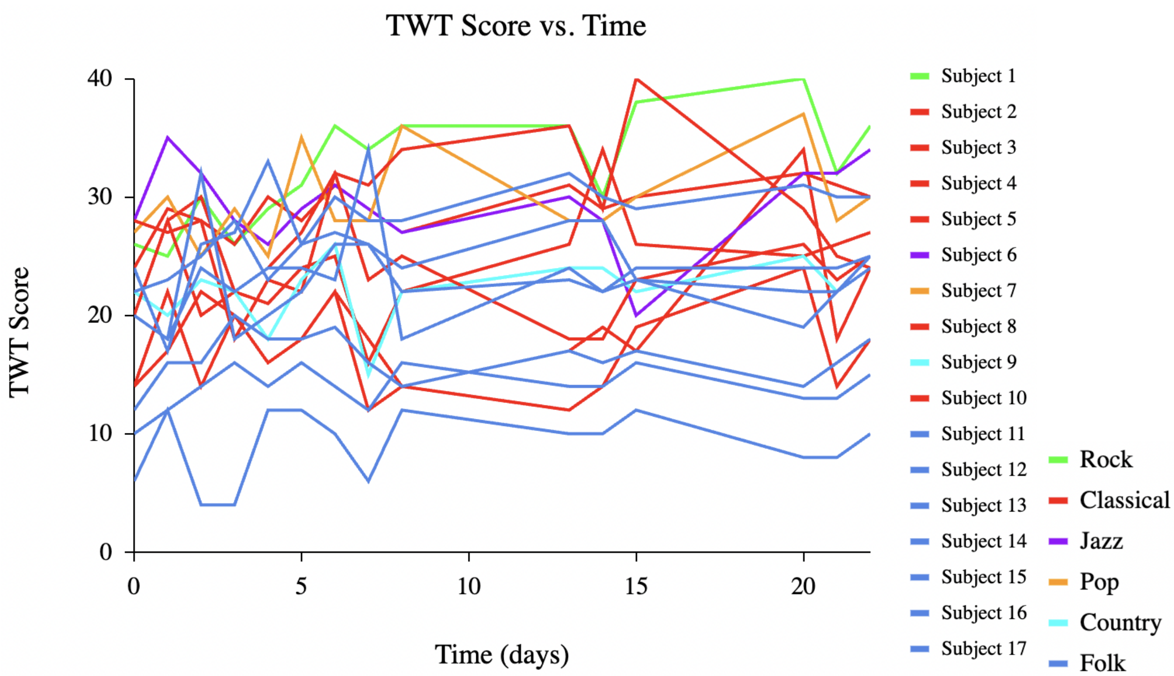


Figure 10. TWT scores over time. There is a general upwards trend for almost all participants and a decrease in the TWT scores on days 7, 14, and 21 (the days without music) for some subjects.

To better visualize the changing TWT score trend of subjects who listened to different genres of music in both cycles, I analyzed the TWT scores over time compared with the baseline score (**Figure 11**).

Results suggest that subjects who listened to classical and folk music exhibited the most consistent averaged TWT scores at or over the baseline (indicated by the x-axis).

I further quantified the percentage score increase (and decrease) for the negative control trial, all six genres of music, and 40 Hz music (**Figure 12**). Folk music resulted in the highest average percentage improvement for the TWT, closely followed by classical music, while some participants who listened to classical music demonstrated the highest individual percentage improvements. Score changes for rock, jazz, pop, country, and 40 Hz music are not statistically significant, according to their p-values. Despite the high variability in individual score changes, results suggest that classical and folk music showed effectiveness in improving TWT scores. To provide a more comprehensive view of the influence of music on memory abilities, I included a result analysis of the numerical DS test that assesses verbal working memory in conjunction with the TWT that assesses verbal episodic memory.

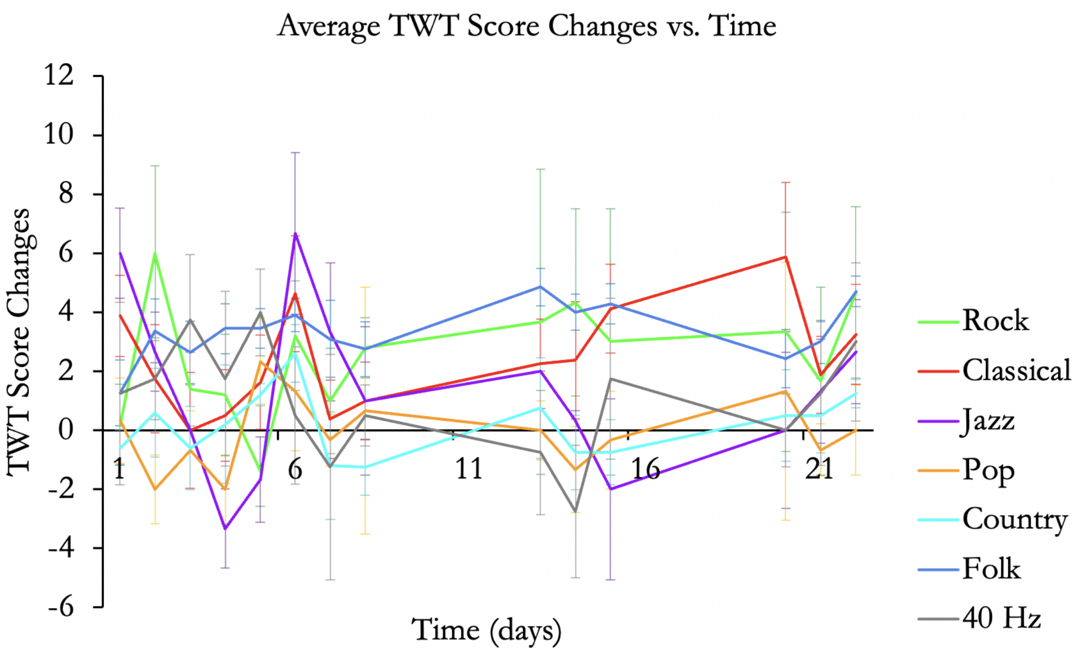


Figure 11. Average TWT score changes over time. Error bars are calculated as standard deviation. The average TWT score changes are calculated by subtracting the baseline score from the scores obtained on all other days of the cycle. Points above the x-axis indicate scores above the baseline, and points below the x-axis indicate scores below the baseline.

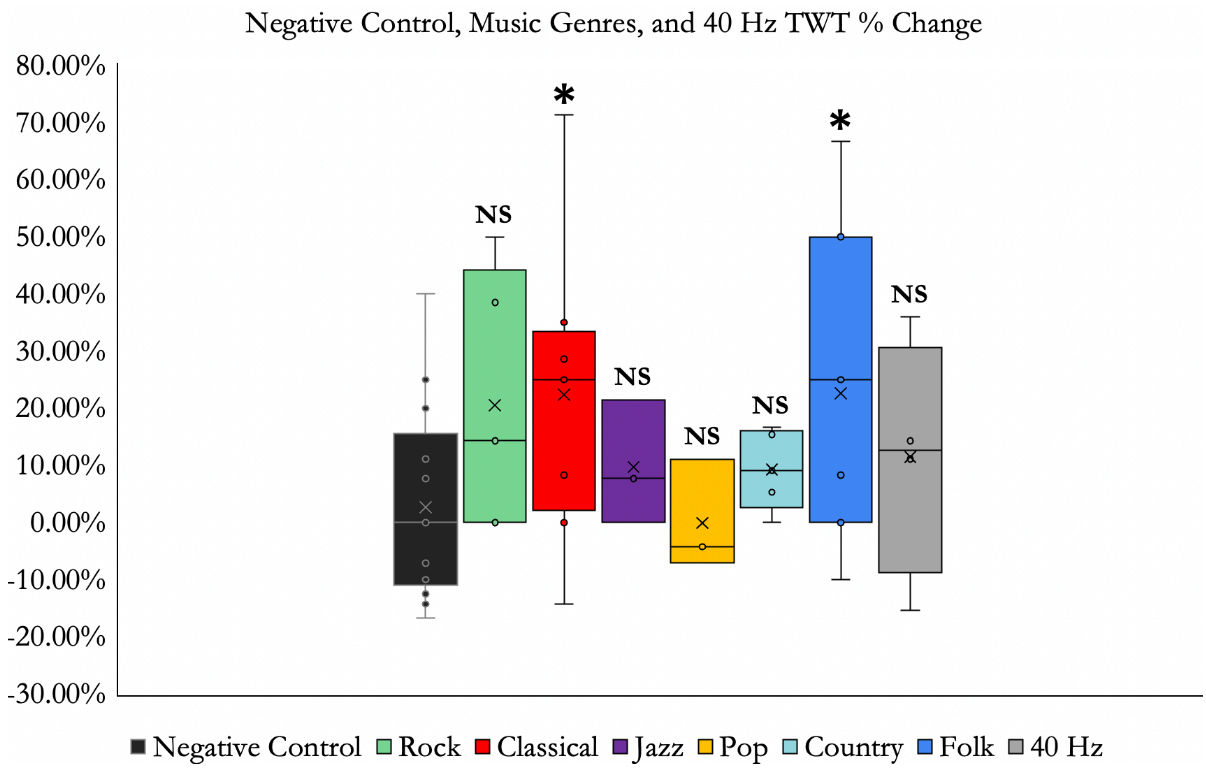


Figure 12. Negative control, music genres, and 40 Hz percentage change of day 22 TWT score compared to baseline score. *P < 0.05 by one-way analysis of variance (ANOVA). “NS” indicates “not significant.” The black crosses indicate the average percentage change.

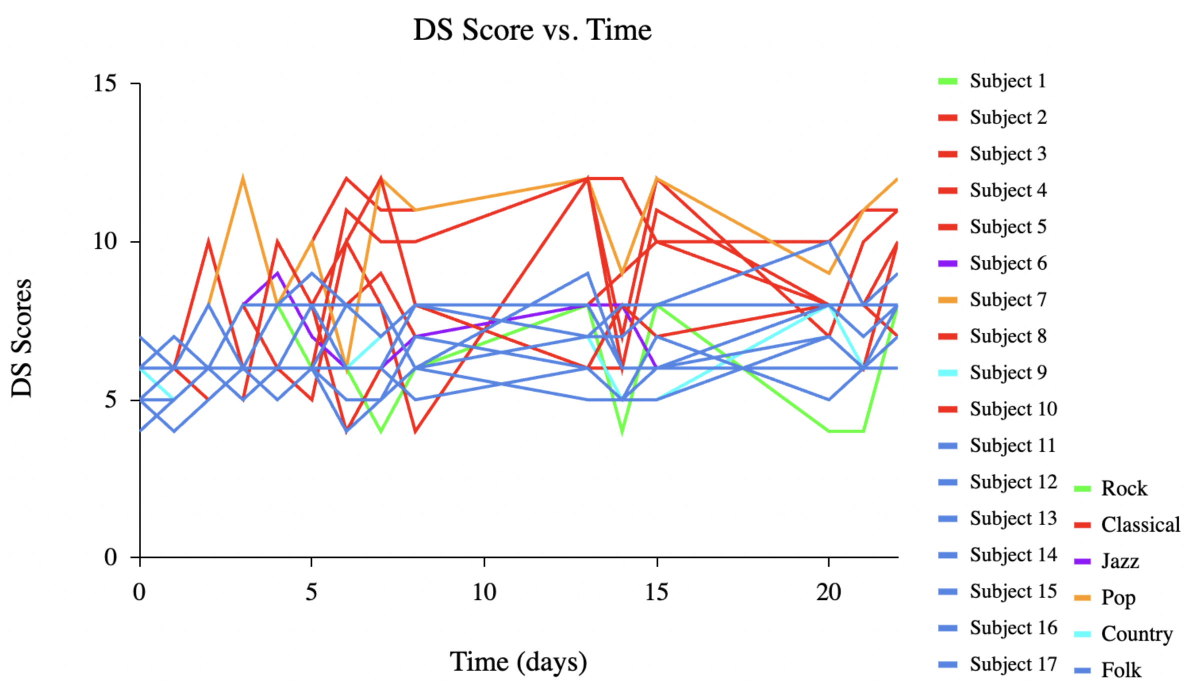


Figure 13. DS scores over time. There is a relatively minor upwards trend for most participants and a decrease in the DS scores on days 7, 14, and 21 (the days without music) for some subjects, particularly obvious on day 14.

DS results from the first cycle also showed high variability. 94% (16/17) of subjects demonstrated higher DS scores at the end of the cycle compared to their baseline score (**Figure 13**), while one subject demonstrated no change. In particular, 76% (15/17) of participants had a 20% or higher score increase, and 47% (8/17) of participants had a 50% or higher score increase (**Table 3**). Subjects aged 65 and above showed score improvements ranging from 20% to 50%. The percentage score increase showed no clear correlation with age. The decreased scores on days 7, 14, and 21 (corresponding to the days without music) indicate the importance of maintaining consistent music therapy.

To better visualize the changing DS score trend of subjects who listened to different genres of music in both cycles, I analyzed the changes in DS scores over time relative to the baseline score (**Figure 14**). The results show that subjects who listened to classical and folk music obtained DS scores consistently above the baseline score (indicated by the x-axis), suggesting the effectiveness of classical and folk music in improving verbal working memory.

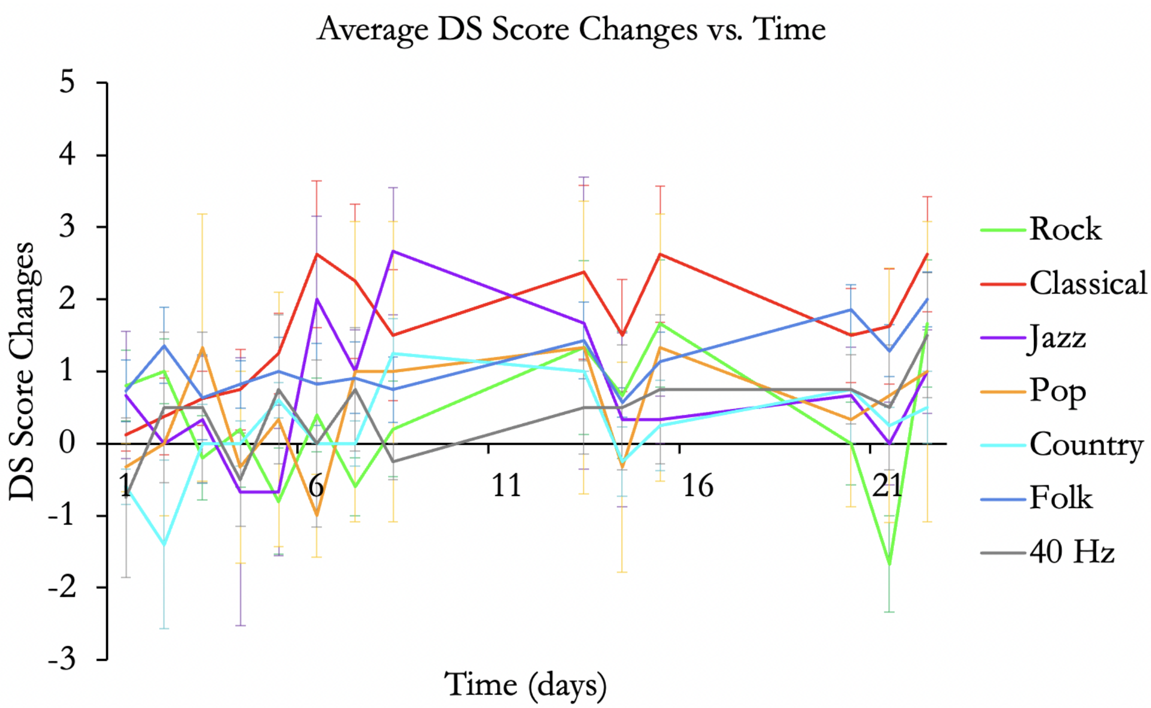


Figure 14. Average DS score changes over time. Error bars are calculated as standard deviation. The average DS score changes are calculated by subtracting the baseline score from the scores obtained on all other days of the cycle. Points above the x-axis indicate scores above the baseline, and points below the x-axis indicate scores below the baseline.

I further quantified the DS percentage increase (and decrease) for the negative control trial, all six genres of music, and 40 Hz music. Results show that classical music resulted in the highest average percentage improvement for the DS, while some subjects who listened to folk music demonstrated the highest individual percentage improvement (**Figure 15**). However, score changes resulting from rock, jazz, pop, country, folk, and 40 Hz music are not statistically significant, according to their p-values. Therefore, results only suggest the effectiveness of classical music in improving long-term DS scores. To summarize, results show that classical music is effective for enhancing both TWT and DS scores in the long term, and folk music is effective for enhancing TWT scores in the long term.

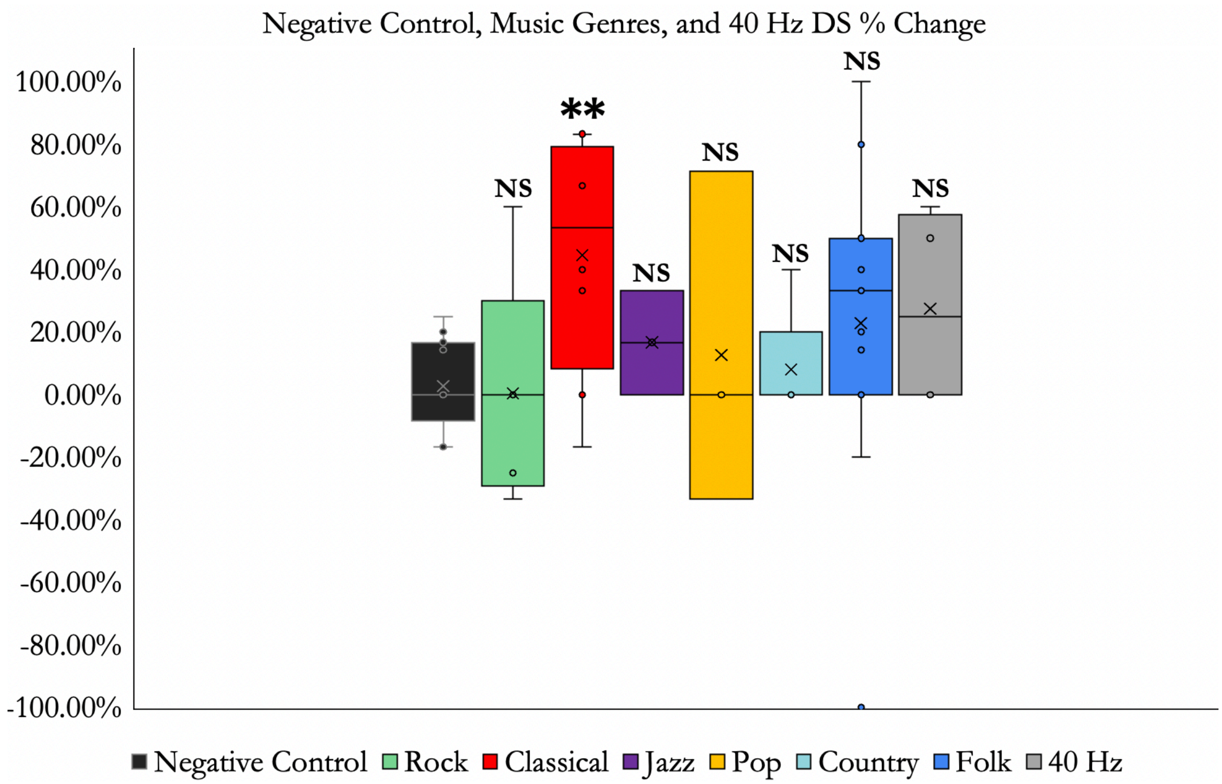


Figure 15. Negative control, music genres, and 40 Hz percentage change of day 22 DS score compared to baseline score. **P < 0.01 by one-way analysis of variance (ANOVA). "NS" indicates "not significant." The black crosses indicate the average percentage change.

3.2.2 Machine Learning Predictions of TWT and DS Score Changes

Normally, it would take 3 weeks to assess the impact of a certain musical treatment in the long term. However, I believe that there is potential to shorten the time-consuming process by creating a random forest machine learning model capable of predicting future long-term outcomes based on short-term outcomes. With this approach, it would become possible to discover the long-term influence of a certain musical treatment on the verbal episodic and working memory of individuals in advance and evaluate the effectiveness of the treatment with higher efficiency. Therefore, I created a machine learning model with the objective of using initial short-term data from the first 7 days to predict long-term TWT and DS outcomes on day 22 (**Table 2**; **Table 3**). To better view the accuracy of the model, I summarized the baseline score, predicted long-term score, real long-term score, percentage change, and the difference between predicted and real scores for the TWT in the first cycle.

Table 2. TWT baseline score, day 22 score predicted by machine learning based on day 0-6 scores, real day 22 score, real percentage increase, and the difference between predicted and real scores.

Subject	Age	Music Genre	Baseline TWT Score (day 0)	Predicted TWT Score (day 22)	Real TWT Score (day 22)	Real Score Change	Predicted vs. Real Difference
Subject 1	9	Rock	26	29.3	36	38.5%	6.7
Subject 2	13	Classical	14	21.5	24	71.4%	2.5
Subject 3	14	Classical	14	19.7	18	28.6%	1.7
Subject 4	15	Classical	24	29.4	30	25.0%	0.6
Subject 5	31	Classical	20	25.7	27	35.0%	1.3
Subject 6	31	Jazz	28	30.7	34	21.4%	3.3
Subject 7	43	Pop	27	29.2	30	11.1%	0.8
Subject 8	48	Classical	28	25.4	24	-14.3%	1.4
Subject 9	50	Country	22	24.1	24	9.1%	0.1
Subject 10	53	Classical	20	25.7	25	25.0%	0.7
Subject 11	59	Folk	22	25.9	24	9.1%	1.9
Subject 12	60	Folk	24	27.5	30	25.0%	2.5
Subject 13	64	Folk	20	24.9	25	25.0%	0.1
Subject 14	70	Folk	20	25.3	25	25.0%	0.3
Subject 15	81	Folk	12	18.7	18	50.0%	0.7
Subject 16	83	Folk	10	18.1	15	50.0%	3.1
Subject 17	84	Folk	6	18.1	10	66.7%	8.1

When compared to the real day 22 TWT scores, machine learning predictions demonstrated a difference of less than 2 points for 65% (11/17) of participants, a difference of 2-4 points for 24% (4/17) of participants, with the least accurate predictions observed among the youngest and oldest subjects. To better view the accuracy of the model for DS score predictions, I summarized the baseline score, predicted long-term score, real long-term score, percentage change, and the difference between predicted and real scores for the DS in the first cycle.

When compared to the real day 22 DS scores, machine learning predictions demonstrated a difference of less than 0.6 points for 41% (7/17) of participants, a difference of 0.6-1.2 points for 41% (7/17) of participants, with the remaining 3 predictions differing more than 1.2 points. Additionally, upon assessing the effect of rock music as reflected in both Table 2 and Table 3, it is important to account for the fact that the subject aged 9 years is currently undergoing brain development, which may affect the results obtained.

Table 3. DS baseline score, day 22 score predicted by machine learning based on day 0-6 scores, real day 22 score, real percentage increase, and the difference between predicted and real scores.

Subject	Age	Music Genre	Baseline DS Score (day 0)	Predicted DS Score (day 22)	Real DS Score (day 22)	Real Score Change	Predicted vs. Real Difference
Subject 1	9	Rock	5	7.7	8	60.0%	0.3
Subject 2	13	Classical	6	9.8	11	83.3%	1.2
Subject 3	14	Classical	6	8.7	8	33.3%	0.7
Subject 4	15	Classical	6	9.8	11	83.3%	1.2
Subject 5	31	Classical	6	9.0	10	66.7%	1.0
Subject 6	31	Jazz	6	8.2	8	33.3%	0.2
Subject 7	43	Pop	7	8.7	12	71.4%	3.3
Subject 8	48	Classical	5	8.2	7	40.0%	1.2
Subject 9	50	Country	6	7.9	6	0%	1.9
Subject 10	53	Classical	6	9.7	10	66.7%	0.3
Subject 11	59	Folk	6	8.4	8	33.3%	0.4
Subject 12	60	Folk	7	8.3	8	14.3%	0.3
Subject 13	64	Folk	5	8.1	9	80.0%	0.9
Subject 14	70	Folk	6	8.4	8	33.3%	0.4
Subject 15	81	Folk	5	7.1	7	40.0%	0.1
Subject 16	83	Folk	5	7.2	6	20.0%	1.2
Subject 17	84	Folk	4	7.5	6	50.0%	1.5

3.3 EEG Data Collected Before and After Musical Treatment

EEG data were collected for subjects to find the potential relationship between TWT and DS score changes with corresponding brain wave changes as well as what the brain waves might imply. I observed that participants who listened to classical music exhibited the largest difference between EEG data before and after musical treatment (**Figure 16**). Specifically, there is a reduction in EEG amplitude after listening to classical music. Similarly, the same cohort of subjects who listened to classical music saw large average increases in their TWT and DS scores, which may suggest a potential correlation between the decrease in EEG amplitude and the amplification of TWT and DS scores. On the other hand, I observed that participants who listened to country music exhibited minimal differences between EEG data before and after musical treatment (**Figure 17**), and similarly, the participant exhibited one of the lowest TWT score increases (**Table 2**) and no DS score changes (**Table 3**) during the first cycle. The EEG data before and after folk music (**Supplementary Figure 2**), jazz music (**Supplementary Figure 3**), pop music (**Supplementary Figure 4**), and rock music (**Supplementary Figure 5**) treatment all showed smaller differences when compared with data from classical music. It is worth noting that there is high individual variability and the results only compare data collected for the same individual at different times.

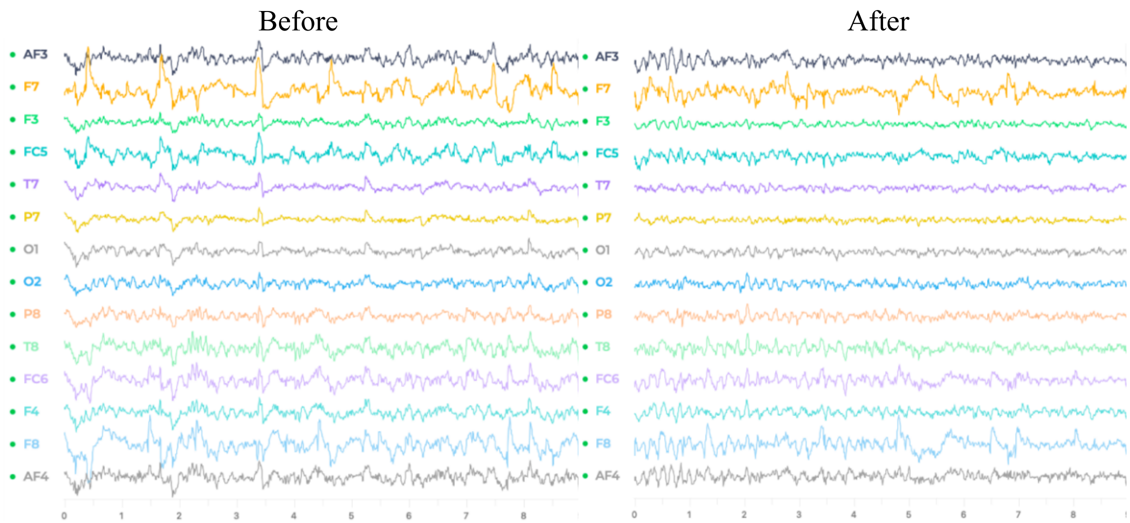


Figure 16. Raw EEG data before classical music treatment (left) and after (right). There is a more obvious difference between the data before and after classical music treatment. This data was collected from a participant who experienced score improvements for both the TWT and DS.

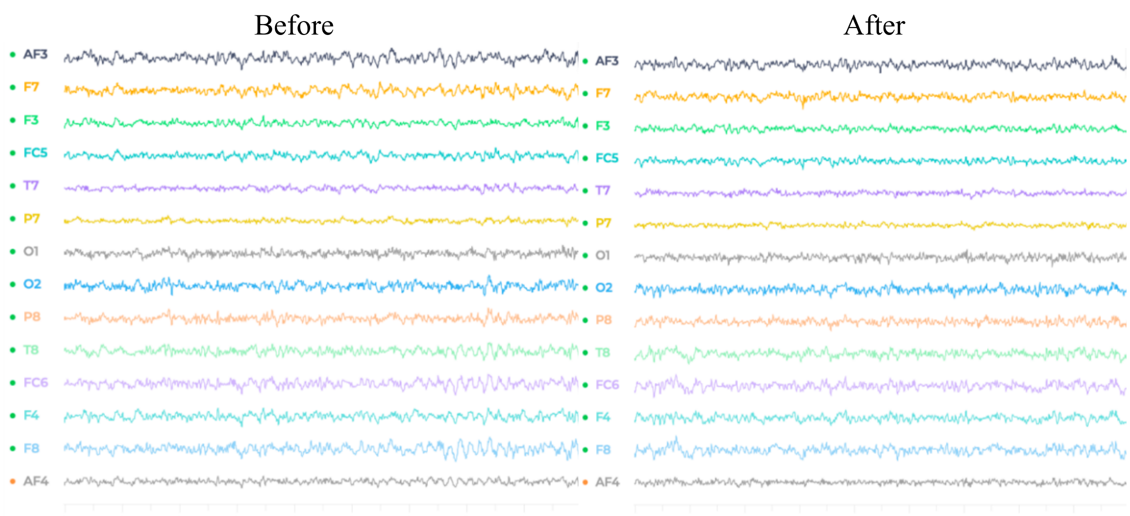


Figure 17. Raw EEG data before country music treatment (left) and after (right). There is a less obvious difference between the data before and after country music treatment. This data was collected from a participant who experienced score improvements for the TWT and no improvement for the DS.

A general trend observed in EEG data before and after musical treatment was that channels around the frontal cortex (AF3, AF4, F7, F8, F3, F4, FC5, FC6) often exhibited notable amplitude differences. This suggests that the frontal cortex region may have experienced the largest effect from music. However, we cannot solely use classical music data to conclude that the effectiveness of music is related to a decrease in amplitude; rather, it is possible that both an increase or decrease in amplitude could suggest influences on memory. Therefore, more data from different individuals in each genre is needed to view a clearer potential correlation between TWT and DS score changes and brain wave changes. Since classical music resulted in the largest EEG amplitude difference before and after treatment, I plotted the amplitude over time and calculated the cumulative sum of amplitudes to further visualize the difference (**Figure 18**).

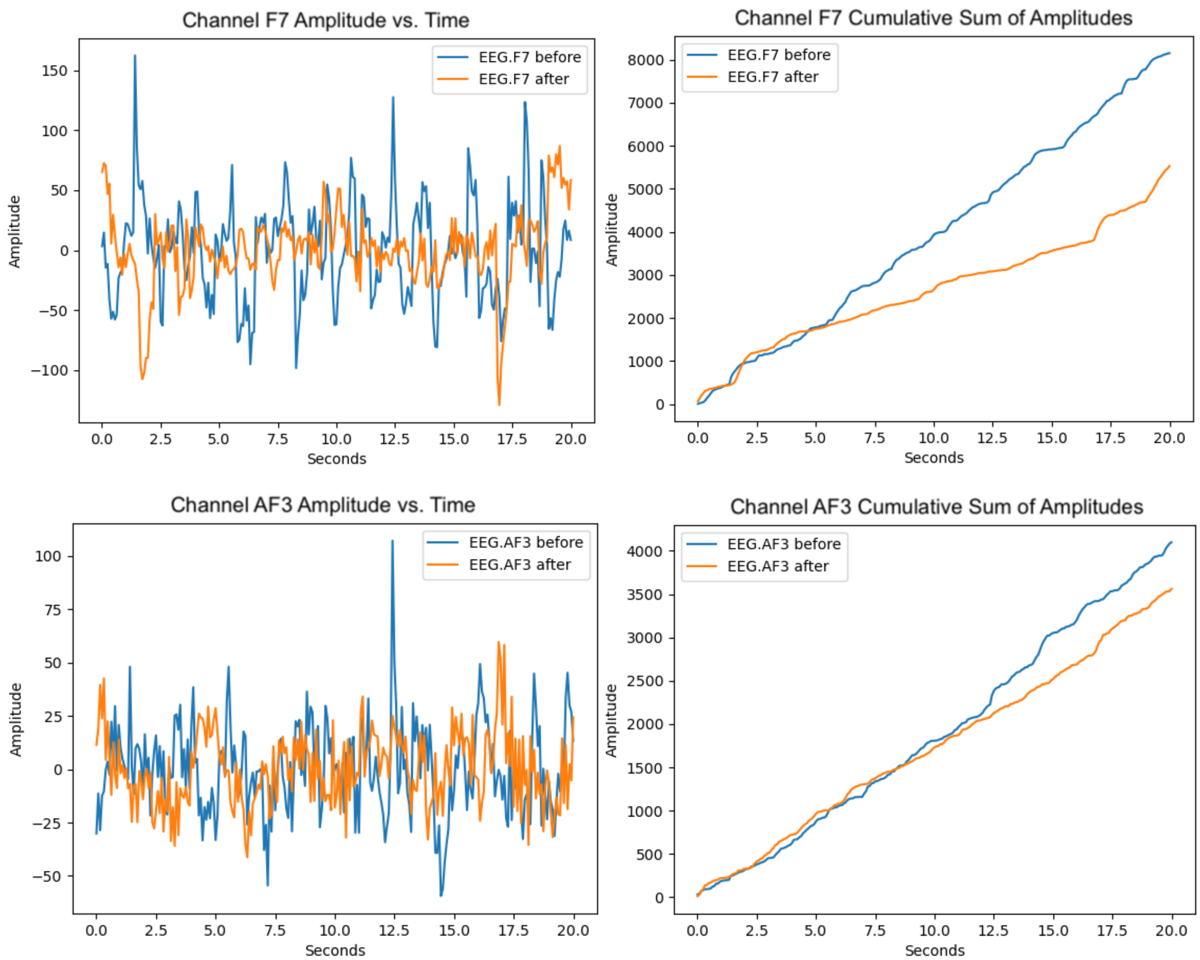


Figure 18. EEG amplitude data (left) and cumulative sum of amplitudes over time by sensor (right), before and after classical music treatment. After 20 minutes of classical music treatment, the amplitude of all sensors decreased. Channel F7 (frontal cortex) exhibited the largest decrease and channel AF3 (frontal cortex) exhibited the smallest decrease in amplitude.

All channels experienced a decrease in amplitude, although different channels displayed varying degrees of amplitude reduction (**Figure 18**). Notably, channel F7 exhibited the most substantial reduction in amplitude subsequent to classical musical treatment in comparison to its state before treatment, whereas channel AF3 showed the most modest amplitude decrease. The results suggest that 20 minutes of classical music treatment may be associated with a decrease in the intensity of the cortical electrical signals of the brain. However, we have not yet arrived at a definite conclusion. Interestingly, while classical music treatment resulted in a decrease in amplitude, we found that folk music treatment resulted in an increase in amplitude (**Supplementary Figure 6**). We acknowledge the limitations of this experiment, and it is possible that there may not be a direct association between EEG alterations and memory test performance. To further investigate the potential implications of changes in brain waves, I analyzed the variations in the band power for different wave types prior to and following classical music treatment (**Figure 19**, which shows the amplitude of five types of brain waves over time; **Figure 20**).

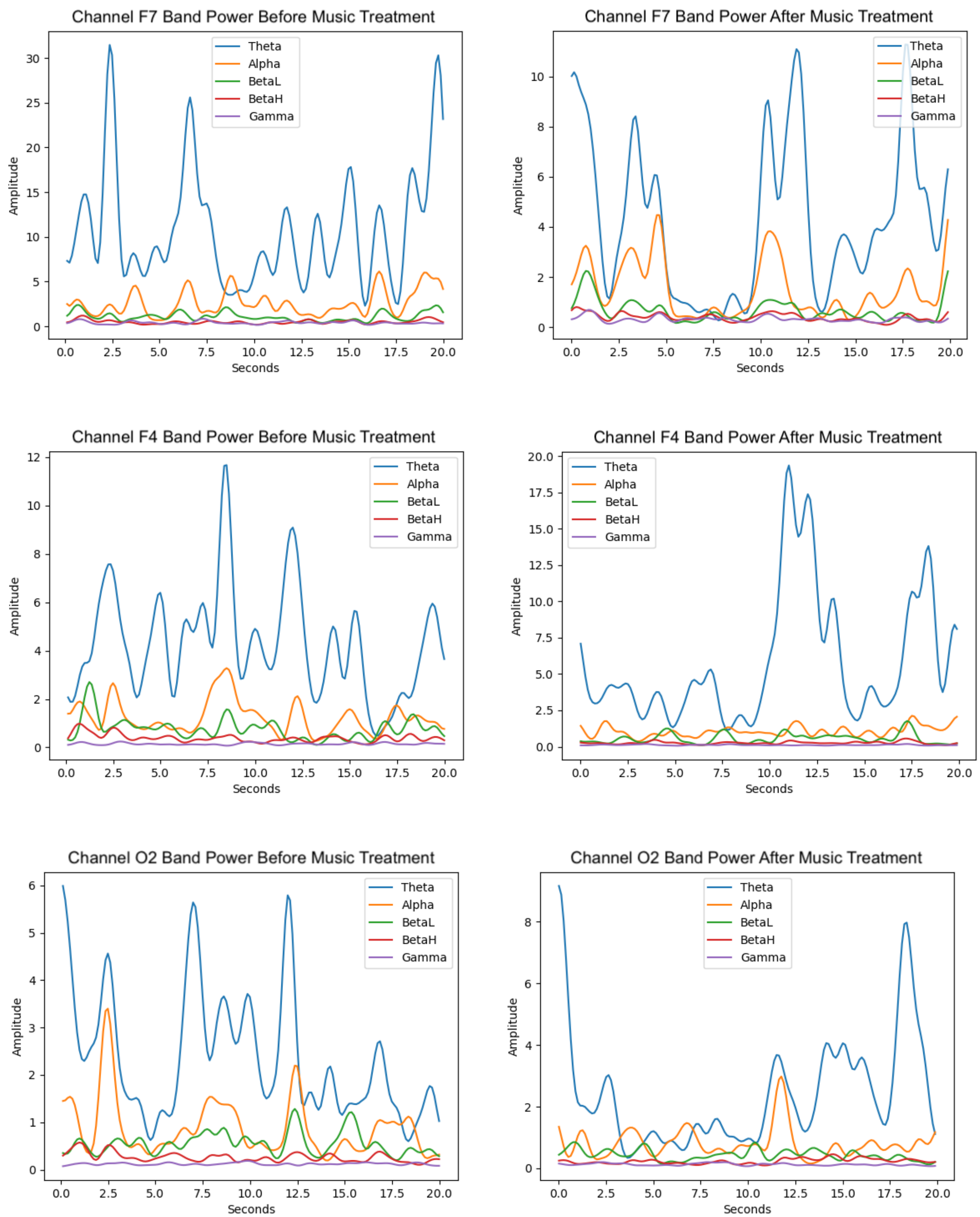


Figure 19. Amplitude of channel band power (theta, alpha, low beta, high beta, gamma) over time. Channels F7 and F4 displayed notable band power differences before and after classical music treatment, while channel O2 displayed the least amount of difference. Note that the y-axes are adjusted to different scales for better visualization.

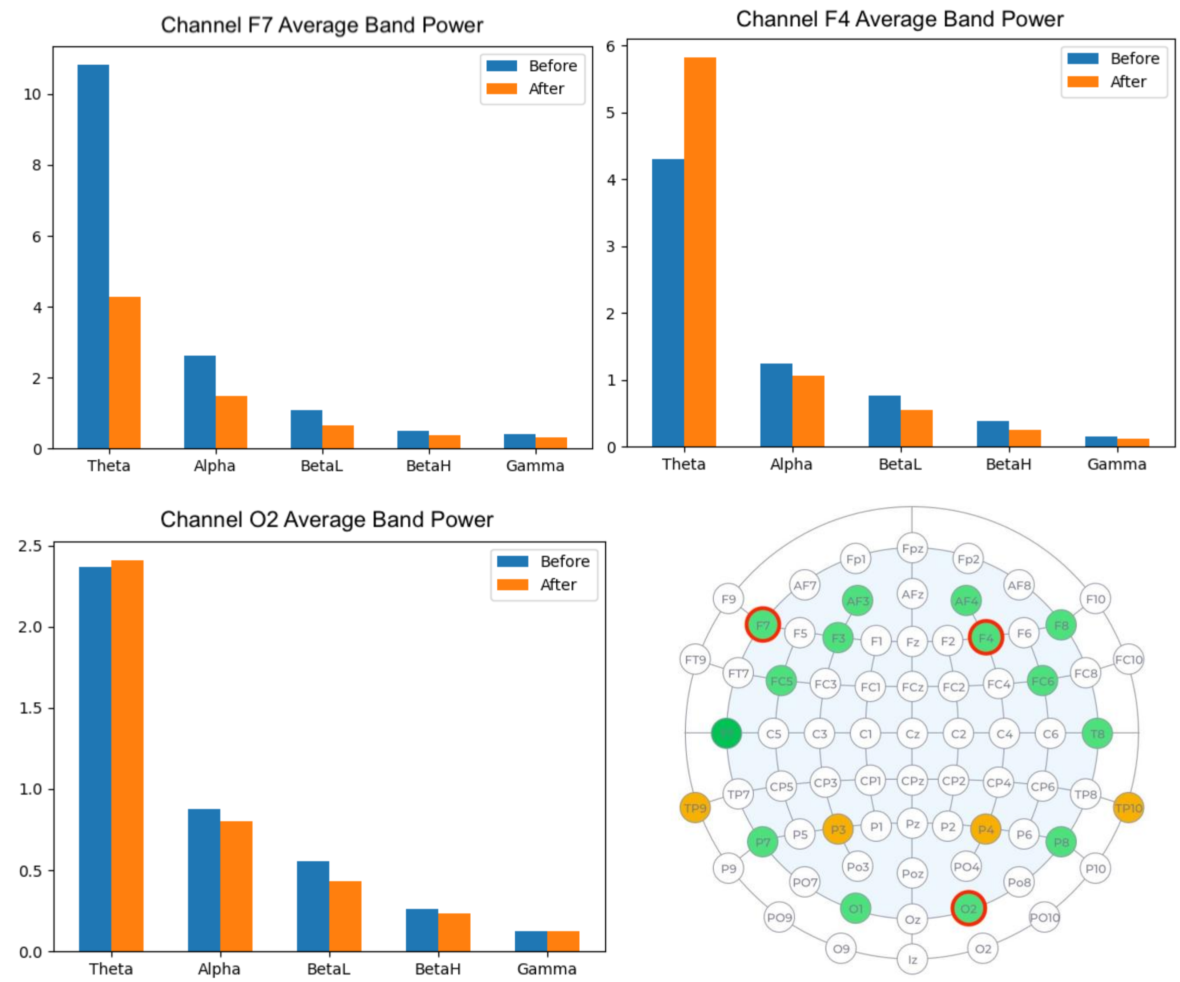


Figure 20. Average band power of channels F7, F4, and O2 (theta, alpha, low beta, high beta, gamma). After treatment, channel F7 showed lower theta (θ) waves, channel F4 showed higher θ waves, and channel O2 displayed the least change. All three channels showed lower alpha (α), beta (β), and gamma (γ) band powers after treatment.

Analysis of different band powers may be informative of the mental state of the subject. Theta (θ) waves have been associated with the state of being deeply relaxed and inwardly focused; alpha (α) waves indicate relaxation; low beta (β) waves signify increases in focus and attention; high beta waves are often associated with stress and anxiety; and gamma (γ) waves indicate intensive thought and concentration[35,36]. To enhance the clarity of the alterations, I calculated the average band power for θ, α, low and high β, and γ waves prior to and following classical music treatment (**Figure 20**). A decrease in θ band powers across 6/14 channels suggests an increase in inward focus; conversely, an increase in θ band powers across 8/14 channels suggests an increase in states of relaxation[6]. A decline in α, β, and γ waves in nearly all channels may indicate higher engagement, wakefulness, and less intensive cognitive activity[6]. A comprehensive investigation is required to establish conclusive connections between EEG band power fluctuations and memory performance. Intriguingly, folk music treatment frequently resulted in increased band powers across channels (**Supplementary Figure 7**; **Supplementary Figure 8**), suggesting that there may not be a clear association between band power variations and memory abilities at this stage.

4. Discussion and Conclusion

4.1 Methodological Variables in the Experiment

4.1.1 White Noise and Music Therapy Placebo Effect

In this study, data from the negative control group was collected without musical or auditory stimulation. One alternative was to include auditory stimulation such as white noise, a sound that contains all audible frequencies. White noise can be used to improve sleep and relieve symptoms for patients with attention-deficit/hyperactivity disorder[37]. Using white noise as the control may also help us determine whether the results were due to the placebo effect or the effect of music. In this study, however, participants were exposed to no auditory stimulation in the negative control trials to imitate their most realistic daily surrounding environment.

4.1.2 Effect of Lyrics

In this study, subjects listened to instrumental music (without lyrics) regardless of the genre. The purpose of listening to instrumental music is to exclude the possible influence of lyrics on the following TWT and DS tests. It has been found that music with lyrics could be distracting and harms performance in cognitive tasks, including verbal memory[38]. Although participants do not listen to music while completing the TWT and DS, it is important to make sure that we avoid possible continued interference with cognitive tasks.

4.1.3 Memory Methods

Feedback from a few participants indicated that they may choose to use certain memory methods to assist them in the TWT, such as associating every word with an image. Therefore, there is a likelihood of overestimating the true memory ability of certain subjects. To improve result reliability, I informed all participants to complete the TWT and DS without using memory methods. However, although subjects restrained themselves from using memory methods, they could have developed subconscious ways of remembering after a few days of testing or acquired familiarity with the semantic category cues over time.

4.2 EEG Precision

I conducted an additional EEG precision experiment (section 2.2.4) because I observed that the EEG detects muscle movement, especially from the eyebrows. To rule out the possibility that the EEG detected only facial or body movement, I collected data under three different circumstances: not moving, thinking about moving the eyebrows and arms, and moving the eyebrows and arms. Results showed that the headset is capable of capturing cortical electrical activity without actual muscle movement, but the amplitude change is small (**Figure 21**). There were minor waves detected when the participant wanted to move their arms and when they actually move their arms (**Figure 22**), indicating that facial movement has the largest effect on brain wave detection.

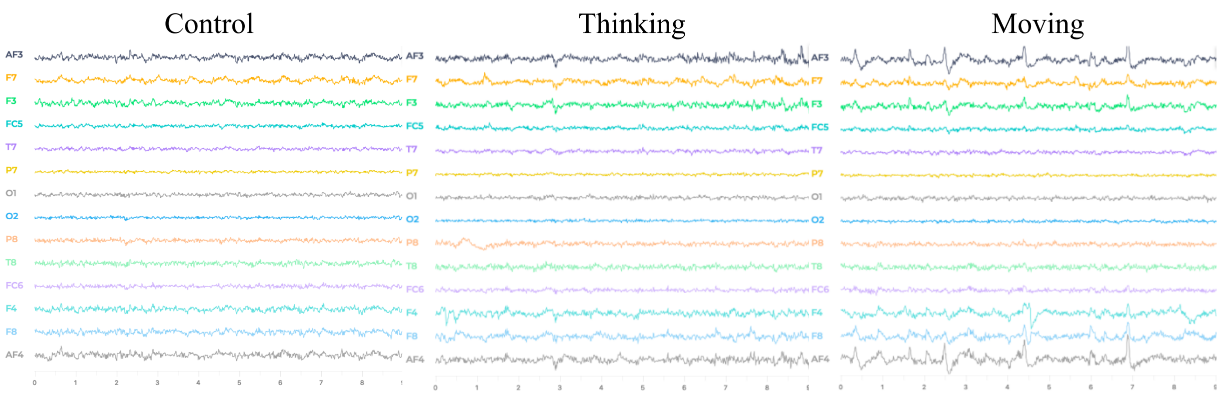


Figure 21. Raw EEG data without facial movement (left), wanting to move (middle), and with facial movement (right).

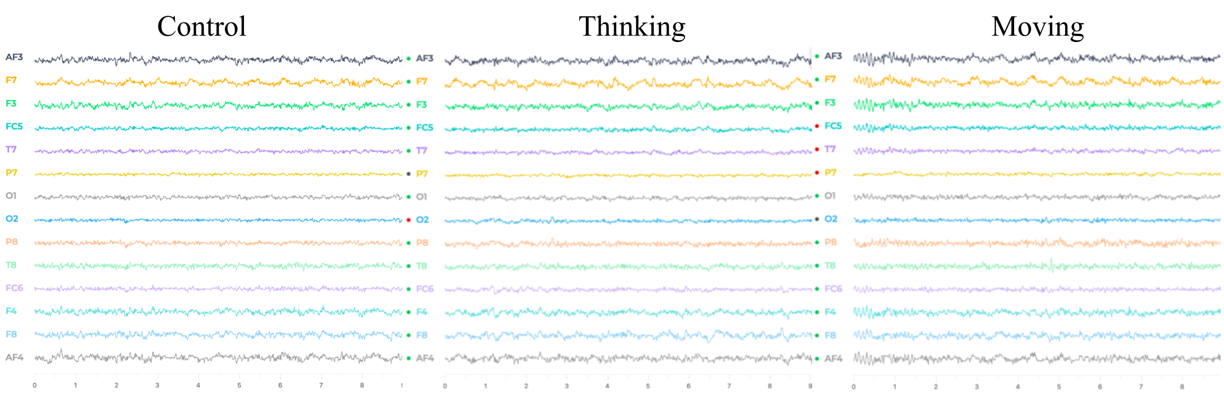


Figure 22. Raw EEG data without arm movement (left), wanting to move (middle), and with arm movement (right).

4.3 Music Preference and Effectiveness

Current results from the first cycle show that classical and folk music have been effective at improving TWT and DS performance. However, we are uncertain whether memory improvement is attributed to participants listening to their favorite music or to specific music genres. I plan to continue this study by collecting more data encompassing various genres and preferences and continue utilizing machine learning methodologies to analyze the factors that determine the effectiveness of music. Next, the third cycle will involve subjects listening to music they like and dislike alongside white noise. As I continue this study, I intend to substitute the music genres with different preference level indicators and employ machine learning techniques to investigate the potential correlations between preference levels and the effectiveness of music. As well, I plan to quantitatively assess the emotional impact of different music genres and use machine learning to predict subject preference levels for different music genres given the treatment outcome.

4.4 Mobile Application Prototype Design

Many participants experienced memory improvements from musical treatment in my study, and I hope to expand this benefit to more people in an accessible way. Therefore, I designed the prototype of a mobile application that provides users with personalized music recommendations, quantifies verbal episodic and working memory abilities, and predicts future outcomes of musical treatment. The prototype displays specific pages asking for basic information, providing users with music recommendations, facilitating the TWT and DS, and providing predictions of long-term TWT and DS scores based on short-term testing data (**Figure 23**). Next, I plan to introduce this mobile application to local senior care centers to make these benefits more accessible.

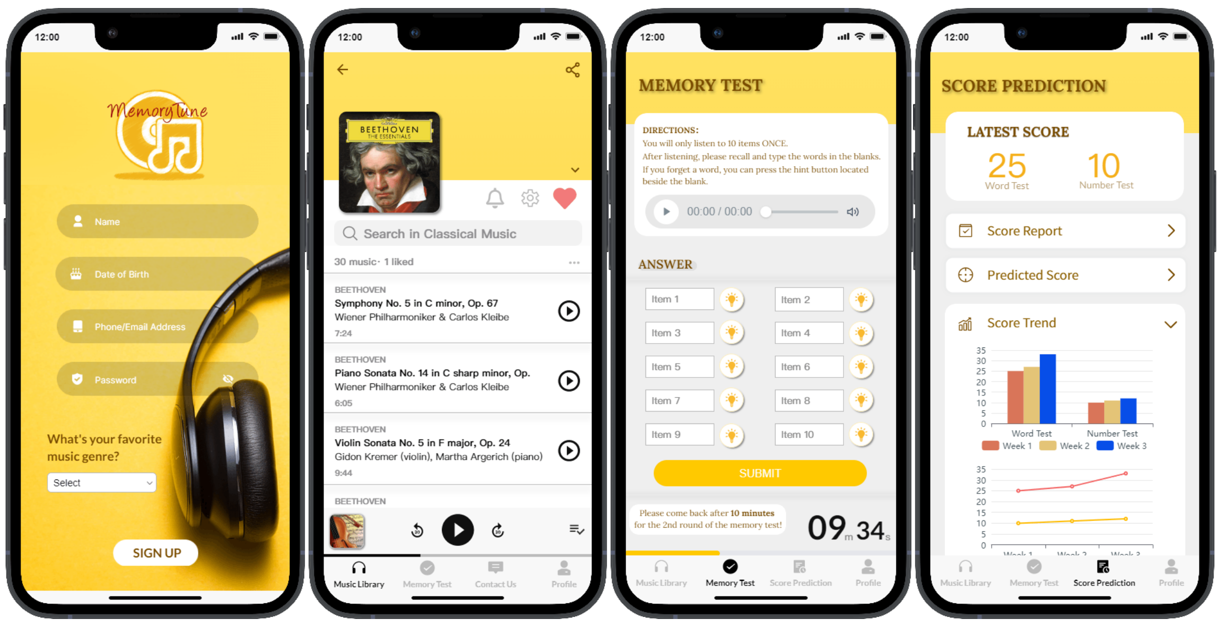


Figure 23. Mobile application prototype design.

4.5 Limitations and Future Prospects

First, I plan to expand the subject pool and conduct more experiment cycles. Additionally, I will continue utilizing machine learning to create models capable of predicting outcomes at various future time points. Subsequent experiments include gathering more data across each music genre and involving patients diagnosed with MCI and AD to provide further insight into future musical treatments.

Second, I aim to introduce this mobile application to senior care centers to help further facilitate musical treatment, quantify verbal episodic and working memory abilities, and predict future outcomes of music therapy. This application will provide a pleasant therapeutic experience alongside its functions focused on assisting memory improvement. With the continuous refinement and optimization of the app, I hope to provide these benefits to more elders, patients, and other individuals in need.

4.6 Conclusion

I found that, on average, classical music stimuli yielded enhancements in long-term TWT and DS scores, while folk music effectively improved long-term TWT scores. Nonetheless, the association between EEG alterations and TWT and DS performance remains inconclusive. In addition, it is important to maintain consistent music therapy. For participants aged 65 and above, folk music treatment resulted in TWT score improvements ranging from 25% to 66.7% and DS score improvements ranging from 20% to 50%. I created a random forest machine learning model that exhibited effectiveness in predicting long-term TWT and DS outcomes based on short-term data, potentially enhancing the efficiency of future treatment evaluation. The discoveries about the effectiveness of different music genres on verbal episodic and working memory could provide valuable insights for future music treatment and prevention of MCI and AD.

Acknowledgment

I would like to thank my instructor, Mr. Andrew McLeod, who assisted me throughout the process by providing valuable suggestions, which helped me improve a lot. I would also like to thank my parents for their support, especially for driving me to other places when I needed to collect EEG data. In addition, I am thankful for everyone who participated in my study. Everyone was kind and supportive, and it gave me even more motivation to work hard throughout this project to achieve my final goals. Finally, I am very thankful for my piano teacher, who cultivated my passion for music and introduced me to the connection between music and neuroscience. When my grandpa started experiencing memory and cognitive issues due to old age, I developed a deep drive to help him and other elders like him. He inspired me to later come up with this project and help those around me.

References

[1] Blinkouskaya Y, Caçoilo A, Gollamudi T, Jalalian S, Weickenmeier J. Brain aging mechanisms with mechanical manifestations. Mechanisms of ageing and development. 2021;200:111575.

[2] Small GW. What we need to know about age related memory loss. Bmj. 2002;324(7352):1502-5.

[3] Etgen T, Sander D, Bickel H, Förstl H. Mild cognitive impairment and dementia: the importance of modifiable risk factors. Deutsches Ärzteblatt International. 2011;108(44):743.

[4] Lane CA, Hardy J, Schott JM. Alzheimer's disease. European journal of neurology. 2018;25(1):59-70.

[5] Jorm AF, Jolley D. The incidence of dementia: a meta-analysis. Neurology. 1998;51(3):728-33.

[6] Nayak CS, Anilkumar AC. EEG Normal Waveforms. StatPearls. Treasure Island, FL: StatPearls Publishing; 2020.

[7] Long JM, Holtzman DM. Alzheimer disease: an update on pathobiology and treatment strategies. Cell. 2019;179(2):312-39.

[8] Weller J, Budson A. Current understanding of Alzheimer's disease diagnosis and treatment. F1000Research. 2018;7.

[9] Skaria AP. The economic and societal burden of Alzheimer disease: managed care considerations. The American Journal of Managed Care. 2022;28(10 Suppl):S188-S96.

[10] Gustavsson A, Norton N, Fast T, Frölich L, Georges J, Holzapfel D, et al. Global estimates on the number of persons across the Alzheimer's disease continuum. Alzheimer's & Dementia. 2023;19(2):658-70.

[11] Shi M, Chu F, Zhu F, Zhu J. Impact of anti-amyloid-β monoclonal antibodies on the pathology and clinical profile of Alzheimer's disease: a focus on aducanumab and lecanemab. Frontiers in aging neuroscience. 2022;14:870517.

[12] Van Dyck CH, Swanson CJ, Aisen P, Bateman RJ, Chen C, Gee M, et al. Lecanemab in early Alzheimer's disease. New England Journal of Medicine. 2023;388(1):9-21.

[13] Visser H. Gait and balance in senile dementia of Alzheimer's type. Age and ageing. 1983;12(4):296-301.

[14] van Iersel MB, Hoefsloot W, Munneke M, Bloem BR, Olde Rikkert M. Systematic review of quantitative clinical gait analysis in patients with dementia. Zeitschrift fur Gerontologie und Geriatrie. 2004;37(1):27-32.

[15] Longhurst J, Phan J, Chen E, Jackson S, Landers MR. Physical therapy for gait, balance, and cognition in individuals with cognitive impairment: a retrospective analysis. Rehabilitation research and practice. 2020;2020.

[16] Raglio A, Attardo L, Gontero G, Rollino S, Groppo E, Granieri E. Effects of music and music therapy on mood in neurological patients. World journal of psychiatry. 2015;5(1):68.

[17] Zatorre RJ, Chen JL, Penhune VB. When the brain plays music: auditory–motor interactions in music perception and production. Nature reviews neuroscience. 2007;8(7):547-58.

[18] Trimble M, Hesdorffer D. Music and the brain: the neuroscience of music and musical appreciation. BJPsych international. 2017;14(2):28-31.

[19] Vanstone AD, Cuddy LL. Musical memory in Alzheimer disease. Aging, Neuropsychology, and Cognition. 2009;17(1):108-28.

[20] Matziorinis AM, Koelsch S. The promise of music therapy for Alzheimer's disease: A review. Annals of the New York Academy of Sciences. 2022;1516(1):11-7.

[21] King J, Jones K, Goldberg E, Rollins M, MacNamee K, Moffit C, et al. Increased functional connectivity after listening to favored music in adults with Alzheimer dementia. The journal of prevention of Alzheimer's disease. 2019;6:56-62.

[22] Iaccarino HF, Singer AC, Martorell AJ, Rudenko A, Gao F, Gillingham TZ, et al. Gamma frequency entrainment attenuates amyloid load and modifies microglia. Nature. 2016;540(7632):230-5.

[23] Adaikkan C, Middleton SJ, Marco A, Pao P-C, Mathys H, Kim DN-W, et al. Gamma entrainment binds higher-order brain regions and offers neuroprotection. Neuron. 2019;102(5):929-43. e8.

[24] Hansen DV, Hanson JE, Sheng M. Microglia in Alzheimer's disease. Journal of Cell Biology. 2018;217(2):459-72.

[25] Martorell AJ, Paulson AL, Suk H-J, Abdurrob F, Drummond GT, Guan W, et al. Multi-sensory gamma stimulation ameliorates Alzheimer's-associated pathology and improves cognition. Cell. 2019;177(2):256-71. e22.

[26] Cárdenas J, Blanca MJ, Carvajal F, Rubio S, Pedraza C. Emotional processing in healthy ageing, mild cognitive impairment, and Alzheimer's disease. International Journal of Environmental Research and Public Health. 2021;18(5):2770.

[27] Dubois B, Touchon J, Portet F, Ousset P, Vellas B, Michel B. " The 5 words": a simple and sensitive test for the diagnosis of Alzheimer's disease. Presse Médicale (Paris, France: 1983). 2002;31(36):1696-9.

[28] Mormont E, Jamart J, Robaye L. Validity of the five-word test for the evaluation of verbal episodic memory and dementia in a memory clinic setting. Journal of Geriatric Psychiatry and Neurology. 2012;25(2):78-84.

[29] Langeluddecke PM, Lucas SK. Wechsler Adult Intelligence Scale–Third Edition findings in relation to severity of brain injury in litigants. The Clinical Neuropsychologist. 2003;17(2):273-84.

[30] Woods DL, Kishiyama MM, Yund EW, Herron TJ, Edwards B, Poliva O, et al. Improving digit span assessment of short-term verbal memory. Journal of clinical and experimental neuropsychology. 2011;33(1):101-11.

[31] Jahanshahi M, Saleem T, Ho AK, Fuller R, Dirnberger G. A preliminary investigation of the running digit span as a test of working memory. Behavioural neurology. 2008;20:17-25.

[32] Srivastava R, Kumar S, Kumar B. Classification model of machine learning for medical data analysis. Statistical Modeling in Machine Learning: Elsevier; 2023. p. 111-32.

[33] EMOTIV. EPOC X User Manual2023 August 18, 2023.

[34] Tzanetakis G, Cook P. Musical genre classification of audio signals. IEEE Transactions on speech and audio processing. 2002;10(5):293-302.

[35] Abhang PA, Gawali BW, Mehrotra SC. Technological Basics of EEG Recording and Operation of Apparatus: Academic Press; 2016.

[36] Abhang PA, Gawali BW, Mehrotra SC. Technical aspects of brain rhythms and speech parameters. Introduction to EEG-and speech-based emotion recognition. 2016:51-79.

[37] Pickens TA, Khan SP, Berlau DJ. White noise as a possible therapeutic option for children with ADHD. Complementary therapies in medicine. 2019;42:151-5.

[38] Souza AS, Barbosa LCL. Should we turn off the music? Music with lyrics interferes with cognitive tasks. Journal of cognition. 2023;6(1).

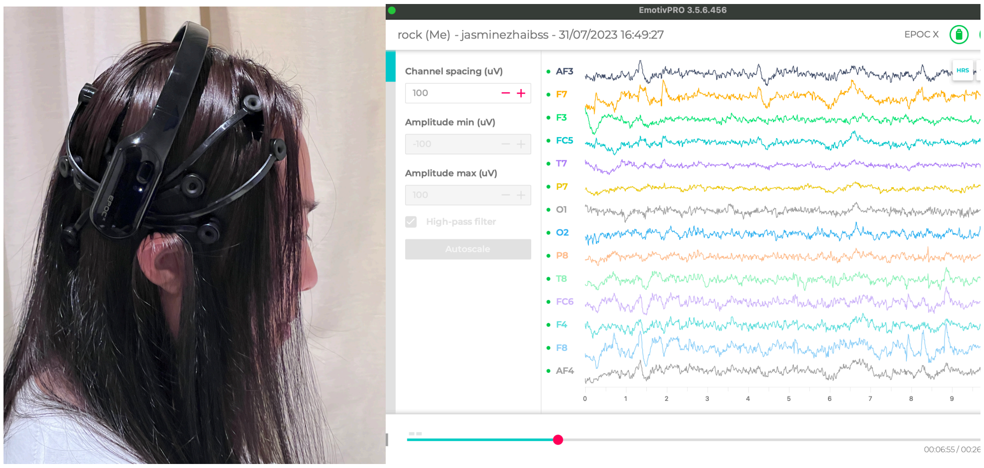
Supplementary information

Supplementary Table 1. Subject information. The first cycle involved subjects 1-17; the second cycle involved all 22 subjects.

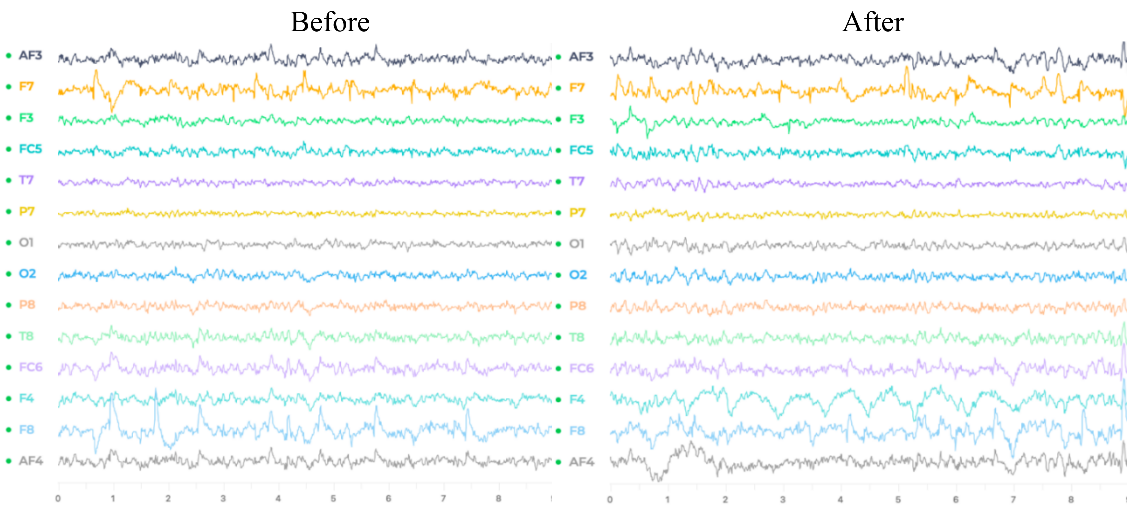
Subject	Age	Gender	Health	Smoke/drink	Medications	Highest education background
Subject 1	9	M	√	No	No	Primary school
Subject 2	13	M	√	No	No	Middle school
Subject 3	14	M	√	No	No	High school
Subject 4	15	F	√	No	No	High school
Subject 5	31	F	√	No	No	Bachelor
Subject 6	31	M	√	No	No	Bachelor
Subject 7	43	F	√	No	No	Master
Subject 8	48	F	√	No	No	Bachelor
Subject 9	53	M	√	No	No	Bachelor
Subject 10	50	M	√	No	No	Bachelor
Subject 11	60	F	√	No	No	College
Subject 12	59	M	√	Drink	No	College
Subject 13	64	F	√	No	No	Middle school
Subject 14	70	M	√	Smoke+drink	No	Middle school
Subject 15	81	M	√	No	No	Primary School
Subject 16	83	M	√	No	No	Primary school
Subject 17	84	M	√	No	No	Primary school
Subject 18	36	F	√	No	No	Bachelor
Subject 19	39	M	√	No	No	Bachelor
Subject 20	59	M	√	Smoke	No	Middle school
Subject 21	80	F	√	Smoke	No	Primary school
Subject 22	81	M	√	Smoke	No	Primary school

Supplementary Table 2. Music preference of subjects, each genre ranked from 1 to 6 (1 being their favorite, 6 being their least favorite)

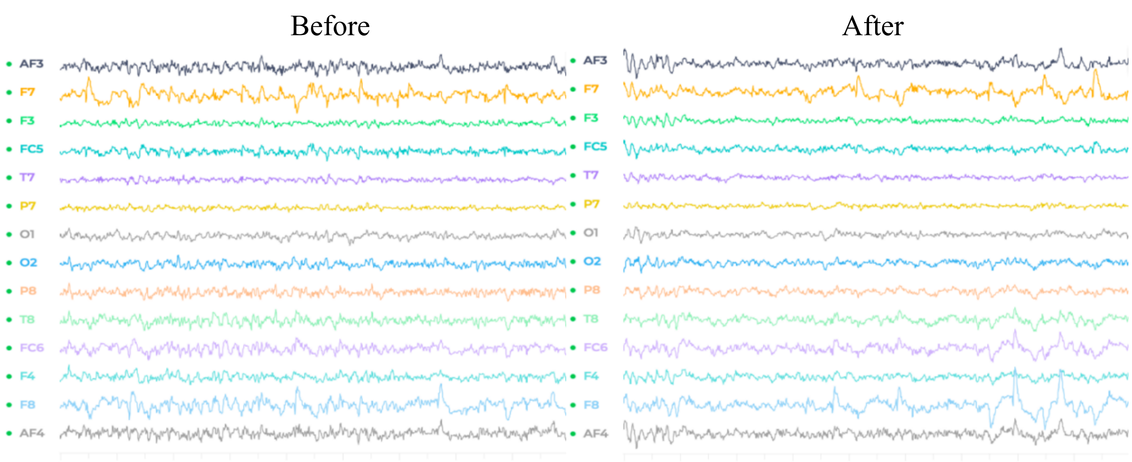
Subject	Classical	Country	Folk	Jazz/blues	Pop/disco	Rock/metal
Subject 1	4	3	5	6	2	1
Subject 2	1	4	3	5	2	6
Subject 3	1	5	2	4	3	6
Subject 4	1	3	4	5	2	6
Subject 5	1	2	3	6	4	5
Subject 6	3	4	5	1	6	2
Subject 7	2	3	4	5	1	6
Subject 8	1	2	3	4	6	5
Subject 9	5	1	2	3	4	6
Subject 10	1	5	4	6	2	3
Subject 11	3	6	1	4	5	2
Subject 12	5	4	1	6	2	3
Subject 13	6	5	1	3	2	4
Subject 14	6	5	1	3	2	4
Subject 15	3	2	1	6	4	5
Subject 16	2	3	1	6	4	5
Subject 17	3	2	1	6	4	5
Subject 18	2	4	3	6	1	5
Subject 19	1	2	3	4	5	6
Subject 20	2	3	1	4	6	5
Subject 21	4	2	1	6	3	5
Subject 22	3	2	1	5	4	6



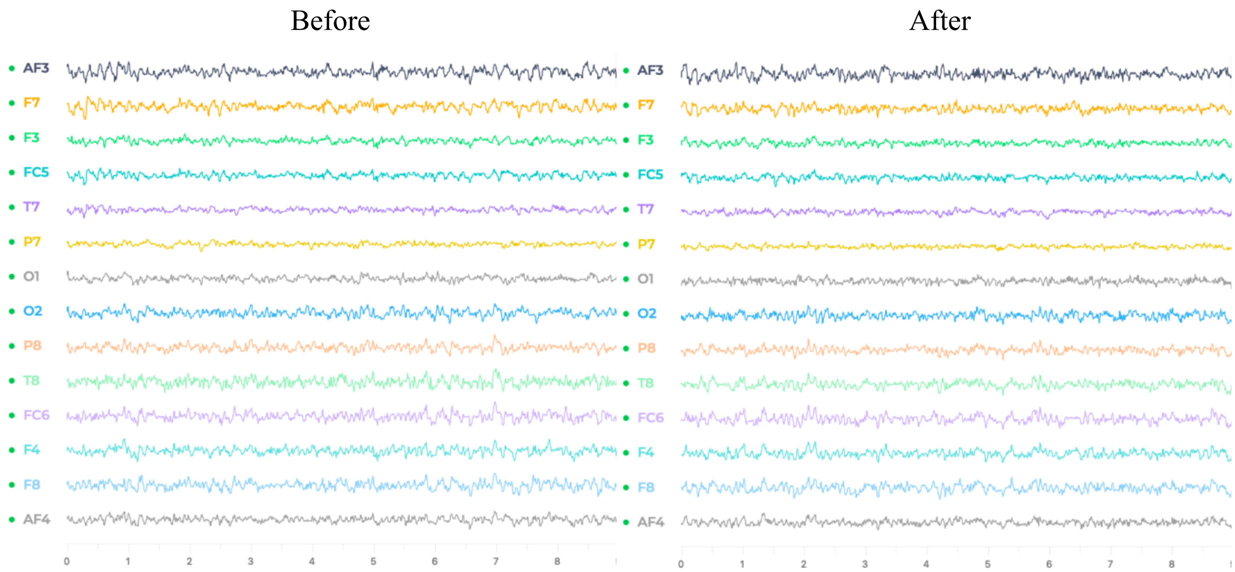
Supplementary Figure 1. Position of EEG headset and sample raw data. The headset is placed in this position, with the headband at the back of the head and all sensors in close contact with the scalp (left). Raw EEG data collected for a participant before listening to music (right).



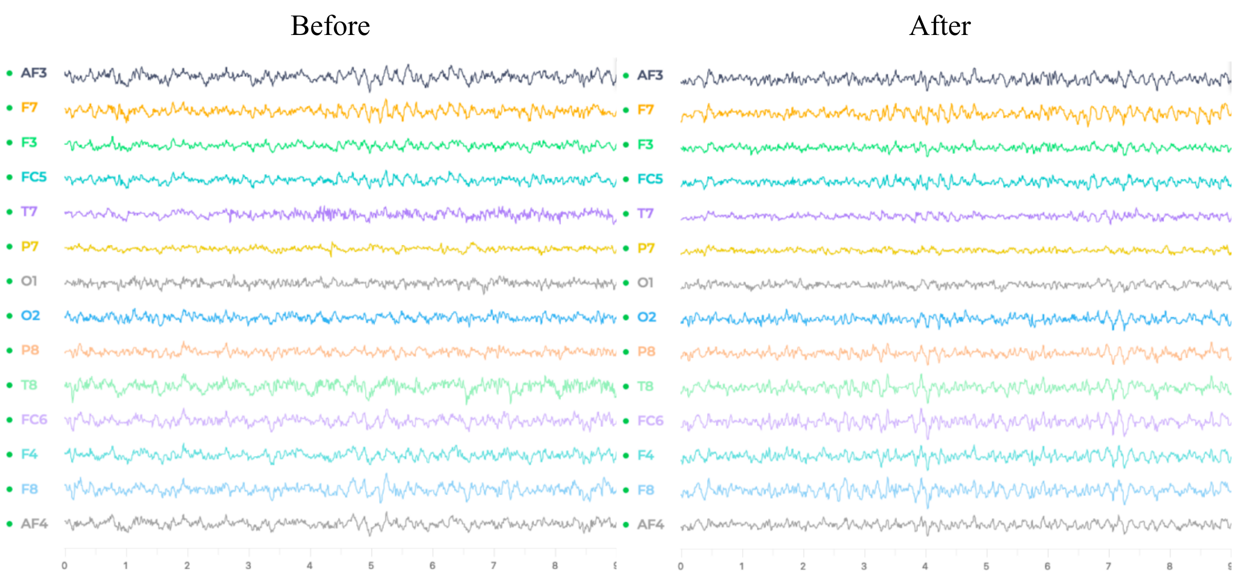
Supplementary Figure 2. Raw EEG data before folk music treatment (left) and after (right)



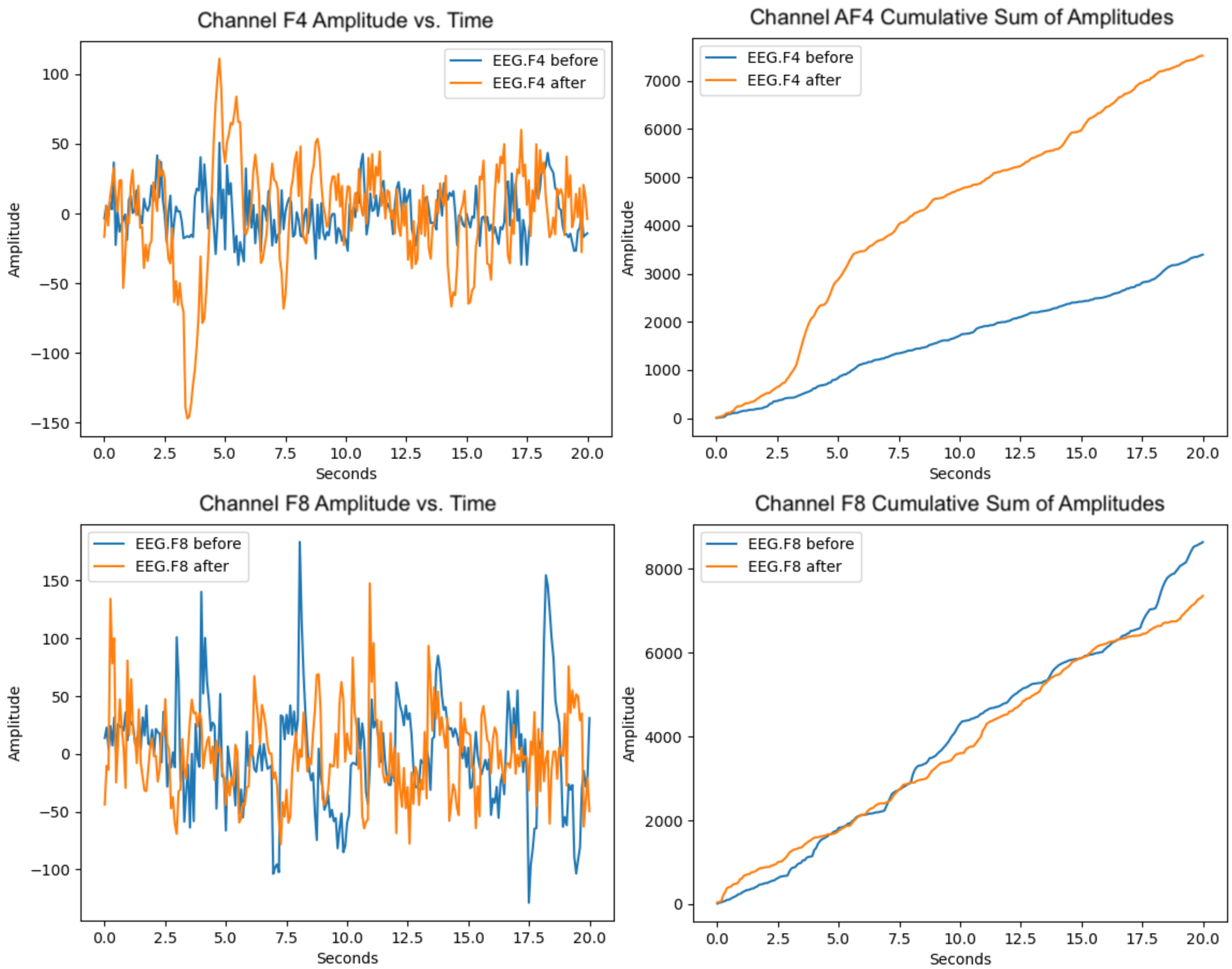
Supplementary Figure 3. Raw EEG data before jazz music treatment (left) and after (right)



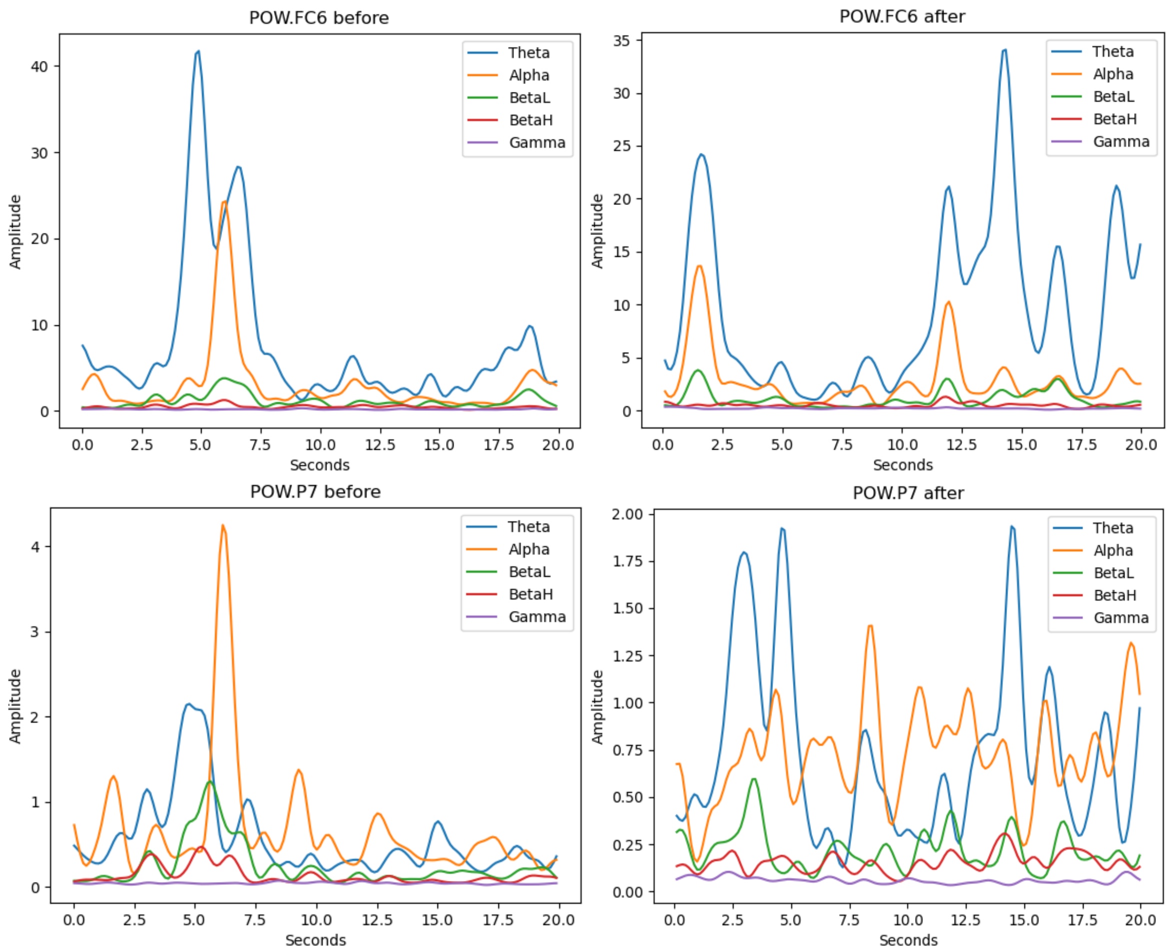
Supplementary Figure 4. Raw EEG data before pop music treatment (left) and after (right).



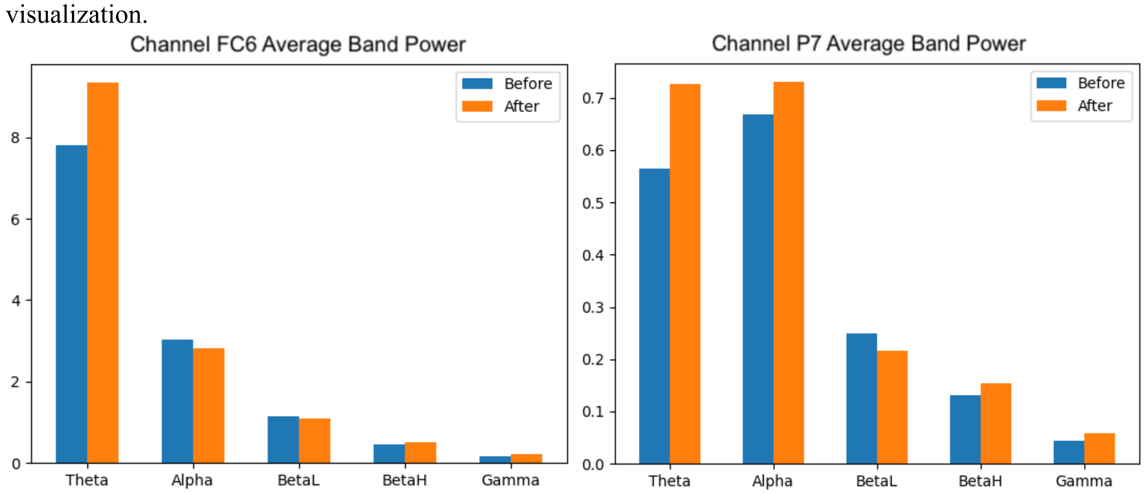
Supplementary Figure 5. Raw EEG data before rock music treatment (left) and after (right)



Supplementary Figure 6. EEG amplitude data and cumulative sum of amplitudes over time by sensor, before and after folk music treatment. After listening to 20 minutes of folk music, the amplitude of all sensors increased. Channel F4 (frontal cortex) exhibited the largest increase and channel F8 (frontal cortex) exhibited the least increase in amplitude.



Supplementary Figure 7. Amplitude of channel band power (theta, alpha, low beta, high beta, gamma) over time. Channel FC6 displayed notable band power differences before and after folk music treatment; while channel P7 displayed the least amount of difference. Note that the y-axes are adjusted to different scales for better visualization.



Supplementary Figure 8. Average band power of channels FC6 and P7 (theta, alpha, low beta, high beta, gamma). After treatment, channel FC6 showed higher θ, high β, and γ waves, and channel P7 displayed the least amount of change.