The Effect of Music on Reading Comprehension

Yueying Dong

Department of Cognitive Science

University of California, San Diego

2019- 2020

An abstract of a thesis submitted to Partial Fulfillment of the Requirements for the Degree of a B.S. in Cognitive Science with a Specialization in Neuroscience with Honors

Committee

Dr. Andrea Chiba, Dr. Xiao Hu, Gina M Dandreapenna

Advisor

Andrea Chiba, PhD

Abstract

Studies on Mozart Effect - the claim that people perform better on spatial abilities after listening to music composed by Mozart - have been controversial. The arousal mood hypothesis is often used as an explanation for the contradictory findings on Mozart Effect studies: The performance in spatial learning abilities arise from the valence and arousal changes induced by music, but not music itself. The current study aims to investigate whether the mood-induced performance enhancement can be generalized to reading comprehension, utilizing both self-reports and physiological measures. We hypothesize that music would enhance both self-reported and physiological valence and arousal, and leads to better performance on reading comprehension measurements; Moreover, music would help maintain arousal after reading. Self-reported arousal is measured by two questions: arousal and emotional intensity; Self-reported valence is measured by the sum of 8 mood questions assessing the positive and negative emotions; Physiological arousal by EEG alpha power and heart rate mean (beats/ min); Reading comprehension is measured by accuracy on two multiple choice questions. The results show that 1) music enhances self-reported valence and arousal, but such enhancement is not reflected in physiological measures. 2) Music does not help or harm reading comprehension, regardless of subjects' music listening habits. However, with lyrical background music, subjects perform better on inference questions. 3) There is a lesser decrease in self-reported arousal after reading when music is present, but there is no significant change in EEG frequencies of the subjects. Unexpectedly, here is a dissociation between physiological and self-reported measures. To better understand the effect of music on reading comprehension

Introduction

Students around the globe often choose to listen to music while engaging in their studies. It is believed that music can improve their learning skills, and such belief is not without scientific foundation: A 1993 study found that individuals showed improved spatial abilities after listening to Mozart for 10 minutes, relative to those who listened to relaxation instructions or silence [1]. This enhancement is known by the public as the “Mozart effect” and is typically seen as proof of the benefits of music on learning. However, as appealing at it may be, the Mozart effect has not routinely been replicated, and the literature on the beneficial effect of music on learning is mixed: some find music helpful, but others do not.

The arousal mood hypothesis has been offered as an explanation of the Mozart Effect. It claims that spatial learning enhancement arises from valence/arousal changes, due, in this case, to the music [2]: A meta-analysis of 16 studies on the Mozart Effect concluded that enhanced learning -- specifically, spatial learning and abstract reasoning -- results from participants enjoying the music and associated arousal changes [3]. Subsequent studies confirmed this finding. Thompson and Schellenberg compared the effects of happy and sad music on spatial learning and found the different musical selections caused differential enjoyment, arousal, and mood; and when controlling for such enjoyment, arousal and mood differences, the Mozart effect disappeared. Moreover, other studies have shown that these arousal improvements are not exclusively produced by music--other types of manipulations such as a reading passage could influence arousal and lead to similar spatial learning enhancement [4].

In sum, according to the arousal-mood hypothesis, music influences both arousal and mood by affecting one’s levels of energy and of pleasure [5]. These affective qualities, considered achievement emotions in control value theory [6], are tied to the accomplishment of activities -- in this case, the spatial abilities improvements observed in Mozart effect. Thus, as the theory predicts, the arousal and mood change elicited by self-selected background music can regulate learning and performance [7].

Because music can affect spatial abilities via mood induction, it is likely that listening to music during studies may also impact reading comprehension, another aspect of learning that is crucial for students. Consequently, it is important to investigate if the valence and arousal changes potentially elicited by self-selected music can enhance reading comprehension. In a previous study, 100 individuals read either with or without self-selected background music and were then tested on their reading accuracy. Self-selected music was found to improve reading accuracy based on post-reading test question performance, accompanied by an increase in self-reported valence/arousal ratings after listening to music and a decrease after reading a passage [8]. This earlier study primarily utilized behavioral measurements, such as standardized scores, to test performance before and after listening to music, in addition to self-reported valence and arousal ratings. Consequently, the current project aims to explore the impact of students' natural music listening choices on arousal and reading comprehension, using not only behavioral measurements--including reading performance and self-reported valence/arousal scores--but also physiological measures such as EEG and heart rate, measurements commonly used to assess one's arousal state. Usually a decrease in EEG alpha (8-13 Hz) power is associated with an elevated arousal level; however, the role of lower (8-10 Hz) and upper (10-13 Hz) alpha seem to diverge in music-related tasks such as listening and composing music [9]. A decrease in upper alpha is related to language specific processing and long-term memory performance, whereas a decrease in lower alpha is the result of arousal and attention processes [10]. Therefore, in this study, we will examine the alpha bands separately. Heart rate usually increases with arousal and attention [11].

In this within-subject study, 12 subjects read 8 GRE-level passages, 4 with self-selected background music and 4 without (see section 2.2). We examined their self-report of valence and arousal before and after listening to music and before and after reading. Additionally, we examined their heart rate dynamics and their EEG activity (as measured centrally from the scalp). By comparing their EEG and heart rate measures before and after music listening and before and after passage reading, as well as behavioral measurements such as multiple-choice accuracy and self-reported valence/arousal scores, this study aims to answer the following questions.

First, we want to investigate whether music changes valence and arousal, using self-reported and physiological measures. And if so, does this change vary across subjects? Second, we would like to investigate whether self-selected background music improves accuracy on reading comprehension measurements. We will explore one interpersonal variation, the music listening habits (i.e. the frequency of listening to music while studying), and one music characteristic, lyrics, to see if they lead to differential performances on reading comprehension. Third, we want to know if music helps maintain arousal after reading, for both self-reported and physiological measures.

Methodology

Participants

The participants were 12 undergraduate students from University of California, San Diego. (19 to 30 years of age). The research reported in this paper was approved by the UCSD Human Research Protection Program Institutional Review Board (IRB).

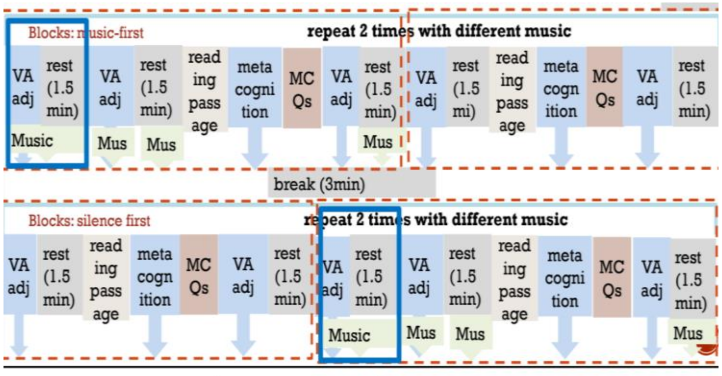
Experiment Design and Procedure

Prior to the experiment, the participants self-assessed their personality traits (Big Five), music training, music preference and habits on studying with music. If they indicate that they usually study with music, they are asked to bring four pieces of music that they listen to; if not, they are asked to bring four pieces of music that they enjoy and with which they are familiar. Each piece of music is then processed using Audacity: Each piece is spliced to 5 minutes in length with standardized volume as well as fade in/fade out effects to ensure smooth transitions between each loop of music (i.e. if participants need more than 5 minutes to read the passage in each trial, then the same piece of music will loop again).

This within-subject experiment consists of one baseline trial and two blocks, each including two control trials and two experimental trials. The order of control and experimental trials varies for the blocks (see figure 1). In each of the four experimental trials, a different piece of self-selected music is played in the background; in the control condition, no music is played.

At the start of the experiment, we recorded participants' baseline EEG using Cognionics headsets [16] and heart rate using E4 Empatica wristband [17]. Next, to familiarize participants with the experiment procedure,

we present them with a practice block that simulates the reading task. The formal blocks followed completion of the practice block. In an experimental trial (see figure 2), participants were first asked to rate their valence and arousal (VA1). There are 8 valence questions for each trial, and participants are asked to self-assess the directionality of their emotions. (For example, Are you feeling calm? Yes/No) For each valence question that the participants indicated positive emotions we scored it 1, and for negative emotions we scored it 0. Then we took its sum for comparison. There are two arousal-related questions: while the arousal question asks participants to assess their alertness level (i.e. Please indicate your current emotion status: level of arousal: 1=Very calm, 9=Very energetic), the emotional intensity question assesses the intensity of their current feeling (i.e. Please indicate your current emotion status: level of valence: 1=Very unpleasant; 9= Very pleasant). VA1 is then followed by a 90-second resting period (Resting1). During Resting1, participants listened to their self-selected music while EEG and heart rate measurements were recorded. Then they again rated their valence and arousal (VA2), followed by a second resting period (Resting2). The music continued, and the participants were presented with a GRE-level reading passage that was emotionally neutral and with low arousal. After participants finished reading the passage, they were given three metacognition questions concerning engagement, difficulty, and understanding. These were followed by two multiple choice questions on the content of the passage, for the purpose of gauging reading accuracy. To simulate a real-life learning environment, there was no time limit for reading or multiple-choice questions. After reading, participants rated their valence/arousal level once more (VA3), followed by another 90-second resting period with music (Resting3). The control trials were structured in a similar way, except without background music; consequently, there is no self-reported valence arousal measure before music, or Resting1, so the control trials are shorter than the experimental trials.



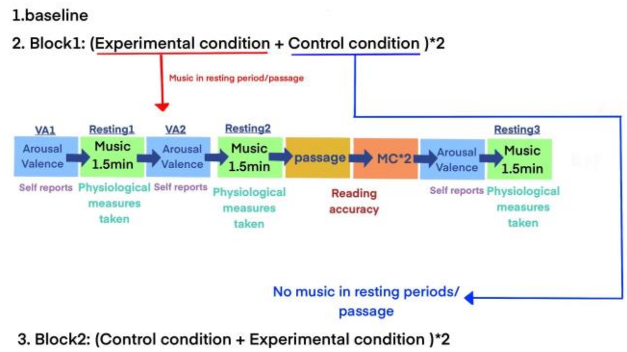


Figure 1(left): Experimental procedure. There is a total of one baseline and two blocks, in block1 and 2, the order of experimental and control trials are reversed. Block1: experimental + control + experimental + control. Block2: control + experimental + control + experimental. Figure 2 (right): Experimental condition in greater detail. For experimental trials, music is played in the background starting at Resting1 to Resting 3. For control trials, there is no background music. In VA1, 2 and 3 the self-reports of valence and arousal scores are taken. In Resting 1, 2 and 3 the physiological (EEG and heart rate) recordings are taken.

Measures

The behavioral measures include

1. Self-reported valence and arousal change, measured by the self-reported valence and arousal before and after music, before and after reading.
2. Reading comprehension measured by multiple choice accuracy.

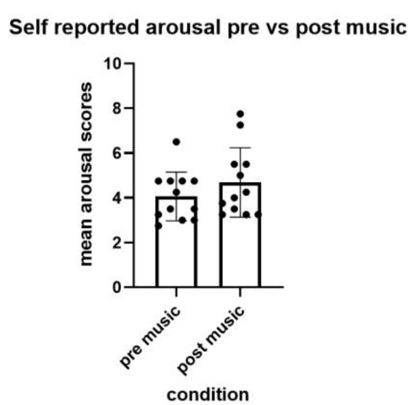
The physiological measures include

1. EEG spectrum: recorded using Lab Streaming Layer (LSL)
2. Heart rate, mean and median (beats/min): recorded using E4 wristband

Results

Does music change valence and arousal, in self-reported and physiological measures?

To address the question of whether music changes valence and arousal and whether these effects vary differentially across subjects, we compared the self-reported arousal and valence questions before and after music listening. A two way within-subject ANOVA examining the factors of self-reported arousal (i.e. VA1 and VA2 in Figure 2, corresponding to before and after music listening) and music/no music was applied, across subjects. Then again for emotional intensity, across subjects. For the arousal ratings, after music the arousal ratings increased significantly ($p = 0.0005$). And there are significant differences across subjects ($p = 0.0006$, see figure 3). Emotional intensity ratings were found to significantly increase after music listening ($p < 0.0001$); significant differences were seen across subjects ($p < 0.0001$) and across trials ($p = 0.0025$). Valence exhibited a decreasing trend that did not reach our criterion for significance ($p = 0.066$). Overall, after music, self-reported arousal increased, but the degree of increase varies across subjects; however, valence did not change.



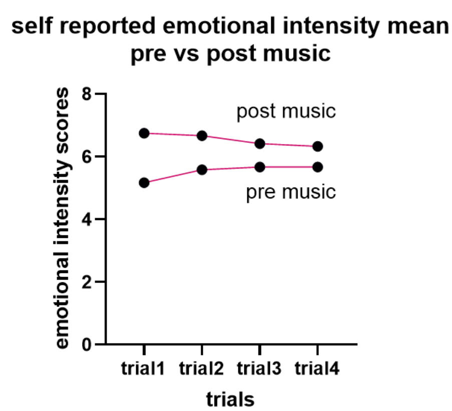
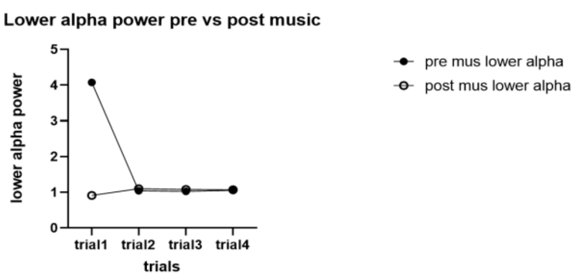


Figure 3 (left): Self-reported arousal pre vs. post music. X axis represents the two conditions: before and after music listening; y axis represents the mean arousal scores. After listening to music, the self-reported arousal scores increased. Figure 4 (right): Self-reported emotional intensity means pre vs. post music. X axis represents the 4 trials, y axis represents emotional intensity scores. After listening to music, the self-reported emotional intensity scores increased.

Next, we looked at the physiological measurements to see how they would correspond to the self-reported arousal increase. A two way within-subject ANOVA, with music and time being two independent measures, was applied to lower and upper alpha power, respectively. For lower alpha, there's a significant music x time interaction ($p = 0.015$), and no significant main effect for music ($p = 0.16$) or across time ($p = 0.097$). For upper alpha, there is a significant interaction between music and time ($p = 0.0065$). No significant difference for pre/post music ($p = 0.19$) or across time ($p = 0.063$) Overall, for both alpha bands, initially alpha decreased, but after the first trial it remained unchanged.



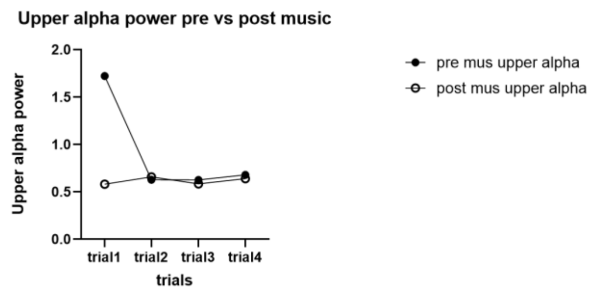


Figure 5 (left): Lower alpha power pre vs. post music. X axis represents the 4 trials, and y axis represents lower alpha power. The lower alpha power decreased only in the first trial. Figure 6 (right): Upper alpha power pre vs. post music. X axis represents the 4 trials, and y axis represents upper alpha power. Similar to lower alpha power, the upper alpha power also decreased only in the first trial.

Moreover, self-reported arousal scores and lower alpha were only moderately correlated ($r = -0.449$, $p = 0.07$): self-reported arousal scores increase after music listening, while the lower alpha power decreases after music. The observed correlation is far weaker than predicted by literature and does not meet our criterion of significance. Among all the 48 cases (12 subjects * 4 trials for each subject) Only in 37.5% cases (18/48) did subjects' self-reported arousal change match their lower alpha power change. As with lower alpha, there was a weak negative correlation between self-reported arousal and upper alpha, but that did not reach significance ($r = -0.076$, $p = 0.41$). In 41% of cases (20/48), subjects reported arousal change matched their upper alpha power change, but the magnitude of the decrease was small. Thus, neither lower nor upper alpha power change corresponds to self-reported arousal.

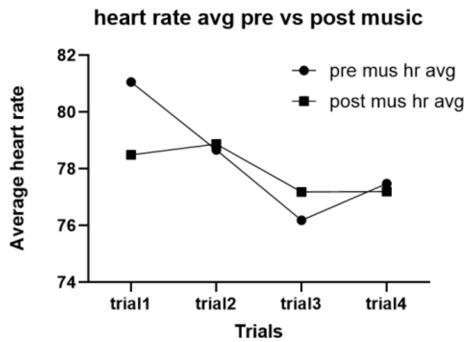


Figure 7 (left): Heart rate average pre v.s. post music. The x axis represents trials, and the y axis represents average heart rate (beats/min). There's a significant decrease in heart rate throughout the trials, but there's no significant change in mean heart rate before and after music.

To investigate whether the other physiological measure, heart rate, would correspond to self-reported arousal after music, we used a two way within-subject ANOVA, with music and trial as the two independent variables. There is no significant difference in heart rate after music listening ($p = 0.44$), but there is a significant decrease across trials ($p < 0.0001$). So, contrary to our prediction, music does not increase heart rate. More surprisingly, though we expected a negative correlation between alpha power and heart rate, at it is suggested by the literature that alpha power decrease and heart rate increase are indicators for arousal; however, we didn't find any correlation (Table 1)

Bands/ p values	Trial1	Trial2	Trial3	Trial4
Lower alpha	- 0.11/0.74	0.09/0.39	-0.03/0.45	0.005/0.49
Upper alpha	-0.1/0.73	0.18/0.58	0.5/0.10	0.17/0.60

Table1: Correlation between upper/ lower alpha power changes and heart rate changes, trial 1 - 4. There is no correlation between alpha power and heart rate changes.

Upon closer examination, we found that 4 out of 12 subjects have a high heart rate baseline: while the usual baseline heart rate ranges from 55 to 85 bpm, these 4 subjects' heart rate baseline ranges from 85 to 120. It seems that these subjects must have been experiencing some stress or energy overload. For these subjects, sometimes an increase in arousal score is accompanied by a decrease in their mean heart rate, for about 10 beats/minutes. Even though this decrease is not consistent across subjects, it's a point of interest that will be statistically testable when we increase our subject population size.

Overall, in only 22.92% of the cases (11/48), the subjects' lower alpha power and heart rate directional change match their self-reported arousal scores in VA questions. And for upper alpha, the matching cases were only 16.67% (8/48). Thus, the physiological data doesn't reflect self-evaluation.

Does self-selected background music improve accuracy on reading comprehension measurements?

In congruence with the previous finding, using a one-way repeated measure ANOVA on reading accuracy, we found that background music does not affect reading accuracy ($p = 0.42$). However, numerically, without music, the overall reading accuracy score is 64% (same for both text and inference questions), higher than the 56% accuracy score (same for both text based and inference based questions) of reading with music.

We also compared the effect of music listening habits on reading comprehension. Using a two-way ANOVA, with music and music listening habit (often listen to music/sometimes listen to music/rarely listen to music) as the independent variables. Despite variability in reading habits across subjects, there was no significant difference in reading accuracy, when reading with or without music, and this is true for both text-based (listening habit: $p = 0.51$; music: $p = 0.89$; habit x music: $p = 0.36$) , and inference-based questions (listening habit: $p = 0.69$; music: $p = 0.98$, habit x music: $p = 0.80$).

A one way between-subject ANOVA was employed to examine the effects of lyrics on reading accuracy. For text-based questions, there's no significant difference in reading accuracy between lyrics and no lyrics groups (p = 0.883). However, for inference-based questions, there's a significant difference between groups (p = 0.017), suggesting that exposure to lyrical music leads to better performance on inference-based questions. However, only 5 out of 12 subjects had chosen lyrical music; thus, running the experiment again with a larger sample size will indicate whether this distinction between text and inference-based questions is reliable.

Does music help maintain arousal after reading, in self-reported and physiological measures?

To investigate whether music helps maintain arousal after reading we first looked at the self-reported arousal and emotional intensity scores.

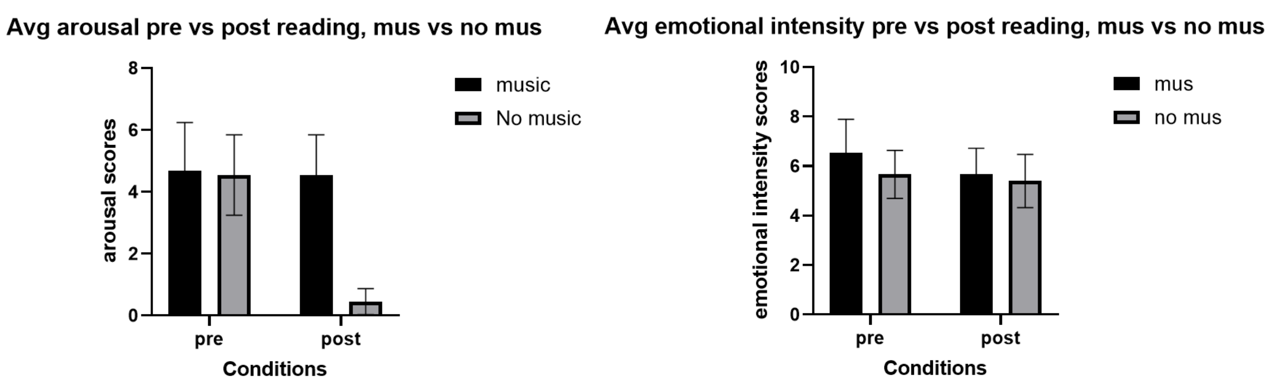


Figure 8 (left): Average self-reported arousal scores pre vs. post reading, with music vs. without music. The x axis represents the conditions, pre and post reading. The y axis represents average self-reported arousal scores. With music, there is a significant decrease in arousal scores after reading. Figure 9 (right): Average self-reported emotional intensity scores pre vs. post reading, with music vs. without music. The x axis represents the conditions, pre and post reading. The y axis represents average self-reported emotional intensity scores. Without music, there is a significant decrease in emotional intensity scores.

A within-subject two-way ANOVA was applied to the arousal and emotional intensity scores respectively, using music and reading as two independent variables. Arousal is significantly lower for no music condition ($p < 0.0001$), and after reading there is a significant decrease in arousal scores ($p < 0.0001$). However, without music, the decrease is larger. Emotional intensity is also significantly lower for the no music condition ($p = 0.013$), and the decrease after reading is significant ($p = 0.018$). In contrast to arousal scores, with music the decrease in emotional intensity is larger ($p = 0.011$).

Next, we investigated arousal after reading using EEG measures. A two-way repeated measure ANOVA was applied to upper and lower alpha bands separately, using music and reading (i.e. before and after reading) as two independent variables. There's no significant difference in EEG lower or upper alpha power after reading for both reading with and without music (Lower alpha - Music: $p = 0.98$, Reading: $p = 0.31$; Upper alpha - Music: $p = 0.56$, Reading: $p = 0.37$). Again, the physiological data does not reflect self-evaluation.

Discussion

After listening to music, there was an increase in self-reported arousal and emotional intensity scores. However, this result is not well-reflected in their physiological measures. Alpha power only decreased in the first trial after participants listened to music and remained the same throughout the rest of the three trials, for both lower and upper alpha. However, the correlation between lower alpha change and self-reported arousal score change is moderate (r = -0.449) and almost reached significance ($p = 0.07$). This correlation is stronger compared to the correlation between upper alpha change and arousal score change, where the correlation is almost non-exist (r = -0.076). A decrease in lower alpha may reflect attentional functions, and a decrease in upper alpha is correlated with semantic, long-term memory performance [10]. Interestingly, in mean heart rate, we also observe the largest difference before and after music in the first trial, even though the degree of decrease is not significant, but it's comparatively larger than the rest of the trials. One plausible explanation is that habitation to music occurred during the course of this experiment. Habituation is defined as a decrease in responsiveness when stimuli are repeatedly presented [12]. Initially, music modulated participants' arousal level, but its effect ceased as participants habituated to the background music.

We also noticed the dissociation between self-reports and physiological measures of arousal. One reason that may account for this dissociation could be interpersonal variations. Studies have shown that self-reports of arousal are affected by a differential level of emotional awareness that varies between individuals [13]. Yet another explanation would be that our subjects have differential levels of baseline arousal. Studies have found that music has calming effects on subjects with high arousal baseline and has an arousing effect on subjects with low arousal [12]. Therefore, for the subjects whose arousal levels were initially high, when they indicated that they are more aroused and alert after listening to music listening, music actually calms them down. In accordance with this hypothesis, we observed that for the 4 subjects with high heart rate baseline (85-120 beats/minutes), sometimes an increase in self-reported arousal is accompanied by a decrease in heart rate. Thus, the lack of correlation between self-reports and physiological reports of arousal may be partially explained by the variability in our subjects' baseline arousal level.

Reading accuracy remained the same regardless of whether music is played in the background, and it is not affected by subjects' music listening habits. Additionally, in contrast to previous studies finding that lyrics resulted in no difference in reading accuracy [8], we found that there's a significant increase in inference question accuracy with lyrics. This result, however, is not consistent with the literature that lyrics can interfere with working memory storage and processing, because lyrics reduce the limited cognitive resources for language processing via visual stream[15] and thus can result in worse reading comprehension. On one hand, it is possible that our result is not replicable, given the small sample size. On the other hand, there are studies that demonstrated that lyrical music only leads to decreased cognitive performance in individuals with lower working memory capacity, but not in individuals with higher working memory capacity [14]. It could be possible that the 12 subjects in this study had higher working memory capacity and thus we did not see a decrease in reading performance with lyrics. However, to further understand the interaction between working memory capacity, lyric music and reading comprehension, a more comprehensive future study will be needed.

Consistent with previous research, with music there was a smaller self-reported arousal decrease after reading. However, with music, the decrease in emotional intensity is larger. Thus, according to the self-reports alone, music partially maintains arousal after reading. But again, the physiological reports did not match self-reports, as there's no significant change in either lower or upper alpha frequency bands in the EEG signals. Again, this could be due to interpersonal variation in emotional awareness.

This discovery underscores the importance of understanding what self-report means, and of developing metrics for physiological scores. For example, if we have heart rate in time series rather than averages, we could look at heart rate variability, which may have been informative. The next step is to have a large number of subjects to determine if the results in this experiment are replicable. If it is, then we need to revisit the literature demonstrating the Mozart effect. If not, then we have to determine what the relationship is between music, reading comprehension and arousal both by self-report and by physiological measures. Moreover, we want to conduct additional studies collecting physiological data and self-reports to understand why they did not match, especially our two physiological measures.

Conclusion

Self-selected music neither helps nor harms reading comprehension, regardless of the listening habit or presence/absence of lyrics. Music leads to increased self-reported arousal, but not in physiological arousal. The results do not align well with existing findings in the literature in that arousal is expected to have increased according to all the measures. We discussed possible reasons for this discrepancy and suggested future studies in which a larger sample size is needed.

References

[1]Rauscher FH, Shaw GL, Ky KN., 1999, "Music and spatial task performance", Nature. 1993;365(6447):611. doi:10.1038/365611a0

[2] W.Thompson, E. Schellenberg, and Gabriela Husain, 2001, "Arousal, Mood, And the Mozart Effect," Psychological Science.

[3]Chabris CF. , 1999, "Prelude or requiem for the 'Mozart effect'?" Nature. 1999;400(6747):826-828. doi:10.1038/23608

[4]Nantais, K. M., & Schellenberg, E. G. (1999). The Mozart effect: An artifact of preference. Psychological Science, 10(4), 370–373. https://doi.org/10.1111/1467-9280.00170

[5]J. A. Russell, A. Weiss, and G. A. Mendelsohn, 1989. "Affect grid: a single- item scale of pleasure and arousal," Journal of Personality and Social Psychology, vol. 57, no. 3, p. 493.

[6]R. Perkrun, A. Frenzel, R. Perry and T. Goetz, 2007. "The Control-Value Theory of Achievement Emotions: An Integrative Approach to Emotions in Education", Emotion in Education, pp.13-16

[7] L. Ferreri and L. Verga, 2016."Benefits of Music on Verbal Learning and Memory: How and When Does It Work?," Music Perception: An Interdisciplinary Journal, vol. 34, no. 2, pp. 167-182.

[8]Que, Y., Zheng, Y., Hsiao, J., & Hu, X. ,2020, "Exploring the effect of personalized background music on reading comprehension." Accepted to Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, Virtual Event, Aug. 1-5, 2020.

[9]Petsche H, Kaplan S, von Stein A, Filz O., 1997 "The possible meaning of the upper and lower alpha frequency ranges for cognitive and creative tasks. " *Int J Psychophysiol*;26(1-3):77-97.

[10]W Klimesch, 1999, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis", Brain research reviews

[11] DC Fowles, 1980, "The three arousal model: Implications of Gray's two-factor learning theory for heart rate, electrodermal activity, and psychopathy", Psychophysiology

[12]RF Thompson, WA Spencer, 1966 "Habituation: a model phenomenon for the study of neuronal substrates of behavior", Psychological review

[13]Salgado, S., & Kingo, O. S., 2019, "Data related to measures of physiological arousal during everyday life experiences and their relation to self-reports of subjective experience of both the event and its memory." *Data in brief, 28*, 104823

[14]Esther Kang, Arun Lakshmanan, 2017, “Role of executive attention in consumer learning with background music”, Journal of Consumer Psychology,Volume 27, Issue 1,Pages 35-48,ISSN 1057-7408,

[15]Alan D. Baddeley, Graham Hitch, “Working Memory”,Psychology of Learning and Motivation, Academic Press,Volume 8,1974,Pages 47-89,

[16]Marini, F., Lee, C., Wagner, J., Makeig, S., & Gola, M. (2019). A comparative evaluation of signal quality between a research-grade and a wireless dry-electrode mobile EEG system. Journal of neural engineering, 16(5), 054001.

[17]Garbarino, M., Lai, M., Bender, D., Picard, R. W. & Tognetti, S. (2014). Empatica E3—A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. The 4th International Conference on Mobihealth, pp. 39-42: IEEE.