Can Background Music Facilitate Learning? Preliminary Results on Reading Comprehension

Xiao Hu
University of Hong Kong Shenzhen
Institute of Research and Innovation
P. R. China
xiaoxhu@hku.hk

Fanjie Li
Faculty of Education
University of Hong Kong
Hong Kong S. A. R.
fanjie@hku.hk

Runzhi Kong
Faculty of Education
University of Hong Kong
Hong Kong S. A. R.
runzhi@hku.hk

ABSTRACT

It is a common phenomenon for students to listen to background music while studying. However, there are mixed and inconclusive findings in the literature, leaving it unclear whether and in which circumstances background music can facilitate or hinder learning. This paper reports a study investigating the effects of five different types of background audio (four types of music and one environmental sound) on reading comprehension. An experiment was conducted with 33 graduate students, where a series of cognitive, metacognitive, affective variables and physiological signals were collected and analyzed. Preliminary results show that there were differences on these variables across different music types. This study contributes to the understanding and optimizing of background music for facilitating learning.

CCS CONCEPTS

• **Applied Computing ~Arts and Humanities** ~Sound and music computing • **Applied computing**~Education

KEYWORDS

Background music, reading comprehension, affect, learning performance, meta cognition, physiological signals.

ACM Reference Format:

X. Hu, F. Li, and R. Kong. 2019. Can background music facilitate learning? Preliminary results on reading comprehension. In *The 9th International Learning Analytics and Knowledge Conference (LAK19), March, 2019, Tempe, AZ, USA.* ACM, New York, NY, USA. 5 pages. https://doi.org/10.1145/3303772.3303839

ACM 978-1-4503-6256-6/19/03…$15.00
https://doi.org/10.1145/3303772.3303839

1 INTRODUCTION

Music, often played in the background as an accompaniment to everyday activities, is known to be powerful in modulating emotion, changing behaviors, and affecting task performance and engagement. Music is socially acceptable, inexpensive, ubiquitous and popular among students from primary school to university, and has unique potential for benefiting students' learning. In fact, a considerable number of students use background music for relaxation, to improve focus on difficult tasks, and to refresh their minds during boring tasks [1]. The effects of music on learning, particularly learning performance and engagement, have been studied across disciplines of education, psychology, and physiology, but inconsistent findings call for new evidence at both behavioral and physiological levels.

The rapidly growing area of learning analytics has started paying attention to the affective aspects of learning, including engagement and emotional states of learners. Towards the ultimate goal of optimizing music listening for learning, this study adopts an experimental approach and aims to investigate how background music affects learning performance and engagement based on analytics at both behavioral and physiological levels. Specifically, this study focuses on reading comprehension, one of the most common learning tasks, with the background of four types of music and one enviornmental sound for comparison purposes. Participants' performances, engagement and metacognition in the five different background sound conditions were analyzed and compared, revealing interesting results that are worthy of further investigation, particularly in association with the physiological data collected in the experiment.

2 RELATED WORKS

The effects of music on learning, particularly on performance and engagement, have been studied across the disciplines such as education, psychology, and physiology, but findings are largely inconclusive. Some studies have found that background music helps facilitate memory encoding and retrieval [2], second language learning [3], and increasing task engagement and performance, particularly for individuals with developmental disabilities [4]. Other studies have detected negative or no effects.

To explain the effects of music on learning, the arousal and mood hypothesis [5] states that music affects learning through modulating arousal and mood. Arousal refers to the level of energy while mood refers to the level of pleasure, equivalent to valence in Russel's two-dimensional model [6] widely adopted in affect studies. The effects of music on valence and arousal are well recognized [7], particularly through music tempo (fast vs. slow) and mode (major vs. minor). Specifically, music with a fast tempo is often deemed arousing while slow music is calming; music in the major mode is often related to positive mood (e.g., happy) whereas minor-mode pieces are more related to negative mood (e.g., sad). Arousal and valence in turn influence learning [8]. It is recognized that learning performance increases along with arousal until an optimum point and decreases afterwards [9]. Positive and negative emotional states are generally deemed helpful and detrimental for learning respectively [10].

Most existing studies on the effects of music on learning are based on behavioral measures (e.g., learning performance, self-reported engagement level), while it has been found that physiological signals are indicative of cognitive activities [11]. For instance, heart rate variability is recognized as an indicator of mental workload [12]. Thanks to the advancement of technologies in collecting and analyzing physiological signals, physiological measures have recently attracted attention in the rapidly growing area of learning analytics where heart rate, blood volume pulse, skin conductance, and skin temperature are shown to be able to detect emotional states of learners with reasonable accuracy [13]. As affect is known to be related to learning performance and engagement [8], studies have also attempted to predict, based on physiological signals, various learning-related constructs including cognitive load [14], perceived difficulty [15], and learning performance [12].

Inspired by previous work, this study aims to explore the effects of different types of background sound, with varying combinations of tempo and mode, on the learning task of reading comprehension, from both behavioral and physiological levels.

3 EXPERIMENT DESIGN

3.1 Learning Tasks

As reading comprehension is an essential activity in day-to-day learning, we adapt the task designed in [16] where short texts were presented to participants followed by simple text-based and inference questions on the content of the reading texts. The questions were designed to measure learning performances with varying difficulty levels.

In this study, five text passages were selected from newsela.com, an online platform of instructional content on different topics. We chose passages under the category of Science and Mathematics as they could involve less emotion than those in socio-cultural categories. The readability levels of the passages were assessed and controlled using tools from readable.io and Coh-matrix [17]. The five passages were of similar readability, with average Flesch-Kincaid Grade Level of 10.2 (standard deviation = 0.52). The Flesch-Kincaid Grade level is a widely used readability score corresponding to a U.S grade level. All the passages also had similar lengths, with an average of 136.2 words (standard deviation = 8.48).

As a measure of reading comprehension performances, four true/false questions were designed for each passage, with two text-based and two inference questions. Participants were allowed to read the passages and complete the questions at their own pace so that their thinking process was uninterrupted. Time spent on reading and question answering was taken as additional metrics of learning performance.

3.2 Background Audio

As vocal music is deemed distracting in learning tasks [18], 1000 instrumental pieces under Creative Common license were randomly selected from the Jamendo collection of independent music. The audio signal processing library, Madmom [19], was used to estimate tempo and key (mode) information from the pieces. A music emotion recognition program [20] was applied to estimate the changes of emotion during each piece. Afterwards, only pieces with steady tempo, few changes of keys and consistent emotions remained. Based on previous findings on emotion and music tempo/mode [7], 100 music pieces were selected and balanced in terms of mode (major, minor), and tempo (slow: about 50-60 beats per minute, fast: about 130-150 beats per minute). Finally, a musically trained researcher listened to the pieces and manually selected four pieces with four conditions: a) major, slow, b) major, fast, c) minor, slow, d) minor, fast respectively. For comparison purposes, another piece of environment sound representing sea waves and seagull sounds was selected from naturesoundsfor.me, an online repository of nature sounds.

3.3 Procedure

A pre-survey was filled by participants before starting the experiment, which included demographic information, music preference, English reading ability, and perceived effectiveness of background music on learning. The experiment was individually conducted for each participant. At the beginning of the experiment, the participant was guided through a practice block where another text passage and music piece similar to the ones in formal experiment were presented. The practice block had the same questions and sequence of operations as those in the formal experiment so as to help the participant become familiar with the experiment procedure.

The formal experiment consisted of five blocks, each of which was assigned with a different condition of background audio. Each block had multiple steps. First, before any stimulus was presented, the participant was asked to report his/her emotional states as measured by the arousal and valence dimensions. Based on the Affect Grid [6] widely used to measure emotions, the arousal and valence dimensions were operationalized as two 9-point Likert scale questions, with options ranging from very calm (1) to very energetic (9), and from very unpleasant (1) to very pleasant (9).

Next, the background audio piece assigned to the block was played for one minute while no reading materials were presented. During this one minute, the participant was asked to focus on listening to the audio and look at a white cross displayed on the center of the black background shown on the computer screen. At the end of the minute, the audio continued, and the participant was asked to rate his/her emotional states again with the arousal and valence questions, as well as his/her familiarity and enjoyment with the audio, on 9-point Likert scales.

Afterwards one of the selected text passages (learning material) was presented on the computer screen for the participant to read while the background audio continued playing. This was to simulate the scenarios of studying with background music. After finishing reading the passage, the participant clicked the "Enter" key on the keyboard to enter the section of question answering. At this moment, the background audio stopped so that the participant could answer questions without the influence of background audio. Besides questions on the participant's current levels of valence and arousal, this section also had metacognition questions, including 1) engagement level, 2) perceived difficulty of the passage, 3) familiarity with the topic of passage, 4) perceived degree of comprehension, and 5) confidence in answering questions regarding the passage, all on 9-point scales such as 1 ("I strongly feel I did not engage in the task") to 9 ("I strongly feel I engaged in the task") [16].

After the metacognition questions, the participant was asked to retell the content of the passage in an open-ended manner, followed by answering the four pre-designed true/false questions regarding the content of the passage. As the final step in a block, the participant was asked to sit back and relax for 20 seconds before starting the next block. This multi-step process was repeated for five blocks, each with a different background audio and text passage. After completing all five blocks, an exit interview was conducted to solicit the participant's opinions on the experiment, including but not limited to the background audio, the reading tasks, emotion and perceived engagement.

The order and combinations of background audio and text passages were counter balanced based on a Latin Square arrangement. All answers and interactions of the participants were logged, together with timestamps. During the experiment, participants wore a research-grade wearable wristband, Empatica E4 [21], for collecting physiological signals, including electro-dermal activity (EDA), blood volume pulse (BVP), heart rate (HR) and skin temperature (TEMP), with sampling rates of 4 Hz, 64Hz, 1Hz, and 4Hz respectively. For signal stabilization, the wristband was mounted on a participant's undominant wrist 2 mins before the task started.

The experiment was conducted in a quiet room with all participants having read and signed the informed consent before starting. Each experiment session lasted for about one hour and each participant was paid a nominal renumeration of about 20 U.S. dollars.

4 PRELIMINARY RESULTS

4.1 Participants

33 (14 male) postgraduate students in a major comprehensive university participated in this experiment. Their mean age was 26.5 (standard deviation = 3.9). Eight of them had formal music training. About equal number of participants had positive (12), neutral (11) and negative (11) attitudes towards the influence of background music on learning. When asked how often they listened to background music while studying, they answered with "almost always" (1), "often" (6), "sometimes" (9), "rarely" (9), and "never" (8).

4.2 Effects on Emotion Changes

During each block in the experiment, participants' arousal and valence status were measured three times: at the beginning, after listening to the audio stimulus, and after reading the passage with background audio. The changes of emotion state after focused listening to the audio and after reading text with the audio in the background could then be calculated and compared across the five conditions of background audio. A one-way ANOVA was applied to test whether there were significant differences on emotion changes among the five audio conditions. As multiple comparisons were involved, Bonferroni correction was applied to control Type I error. Table 1 shows the results.

Table 1: Effects on Emotion Changes

Changes	Emotion	F value	Post hoc Tests
Before and after focused listening	Arousal	13.64***	Fast-minor (1.55) vs. Slow-minor (-0.33)***
	Valence	5.06***	Fast-major (0.79) vs. Slow-minor (-0.27)**
Before and after reading	Arousal	0.98	-
	Valence	2.44	-

***: $p < 0.001$; ** $p < 0.01$; values in parentheses are mean changes in that condition

The results show that different audio stimuli had different effects on emotion change after focused listening to the audio. Specifically, the music piece with fast tempo and minor mode on average increased arousal level for 1.55 scale, whereas the music piece with slow tempo and minor mode decreased arousal level for 0.33 scale on average. For valence, the music piece with fast tempo and major mode increased valence level for 0.79 scale on average while the slow piece with minor mode decreased valence level for 0.27 scale on average. However, all five audio pieces had similar effect on emotion change after reading the text with background audio.

4.2 Effects on Metacognition

One-way ANOVA revealed a significant difference among the five audio conditions on participants' self-reported engagement level with the text passages they were reading ($F(4, 160) = 2.70$, $p = .033$). From the post-hoc tests, the largest difference was between the fast-minor music group and the slow-major music group ($p = .027$). The former had an average scale of 5.48 while the latter was 5.03, which implies that the fast-minor music piece helped the participants engage more with the reading text than the slow-major piece in this experiment. This corroborates findings in the literature that fast tempo improves arousal and certain degrees of arousal help learning. There was no significant difference (at $p < 0.05$ level) among the five audio conditions on other metacognition variables including perceived difficulty and understanding levels of the reading passage, as well as confidence in answering questions related to the passages.

4.3 Effects on Learning Performance

Learning performance was preliminarily calculated by the accuracy of answering the true/false questions after reading each passage, time duration of reading each passage, as well as time duration of answering the true/false questions. The one-way ANOVA analysis indicated no significant difference on these variables among the five audio conditions ($F(4, 160) = .488$, $p = .744$ for accuracy, $F(4, 160) = .132$, $p = .970$ for reading time, and $F(4, 160) = .956$, $p = .434$) for question answering time). Looking at the score of question answering, the participants performed quite well across the five audio conditions, with average scores ranging from 75% (environmental sound) to 82% (slow-major piece). The average reading time also evenly distributed across conditions, ranging from 168.0 seconds (slow-major piece) to 181.2 seconds (environmental sound). The range of question answering time was from 36.1 seconds (fast-major piece) to 46.4 seconds (fast-minor piece).

4.4 Effects on Physiological Signals

At the physiological level, signals collected by the wristband were normalized by z-score normalization. As physiological signals vary across individuals, the normalization was conducted within each individual participant. Time-series of each kind of physiological signals (i.e., HR, BVP, EDA, and TEMP) were then aligned with the starting and ending time of participants' actions including 1) focused music/audio listening and 2) passage reading with music/audio as background. Physiological signals were then split into chunks corresponding to each action of each participant. Indicators of physiological signals during each action period were extracted according to methods in time series and spectrum analysis. The indicators are summarized in Table 2. Results of one-way ANOVA revealed that during the one-minute focused audio listening period, the indicator of LF/HF extracted from EDA signals had a significant difference among the five audio conditions ($F(4, 160) = 2.59$, $p = 0.039$). Post- hoc tests indicated that the fast-major piece and the fast-minor one accounted for the difference ($p = 0.038$). On average, the LF/HF values of EDA while listening to the fast-major piece were significantly higher than those while listening to the fast-minor piece (mean difference = 0.28), meaning the signals fluctuated more equally (less difference between LF and HF) in the former scenario than in the latter. Physiological signals during the periods of passage reading had no significant differences (at $p < 0.05$ level) across the five audio conditions. Together with the result that no significant difference was found in the change of emotions during reading (Table 1), these observations seem to indicate that, when music of different types was used as background for reading, its effect on emotion modulation did not differ as much as when it was used for focused listening.

Table 2: Indicators Extracted from Physiological Data

Categories	Indicators
Descriptive statistics	Mean, Standard deviation, Median, Range
Time series	Means of the values of the 1st (MFD)/ 2nd (MSD) differences of the signals
Frequency domain	High frequency (HF), low frequency (LF), LF / HF

4.5 Effects on Groups of Participants

As participants had different music listening preferences, we grouped the participants based on the following three criteria and examined whether the audio conditions had different effects across the groups: 1) frequency of listening background music while studying, 2) enjoyment with (like or dislike) the audio; and 3) familiarity with the audio. For each criterion, the participants were split into two groups using the median value of the corresponding variable, such that the size of groups was roughly equal. For the variables under comparison, we excluded emotion changes before and after focused audio listening, as they have been shown as significantly different on the sample of all participants (Table 1). Significant results of this analysis are shown in Table 3. Participants who sometimes or more often listened to background music while studying tended to perceive slightly higher valence after reading passages with the slow-major pieces, while perceived lower valence after reading with the environment sound. When the participants disliked the audio, their EDA signals while reading with background audio tended to be different across the audio conditions. The indicators EDA_MFD and EDA_MSD capture the change of the EDA signals across time. The results in Table 3 show that there were more changes on EDA when participants read passages with slow pieces they did not like than with fast pieces or the environment audio they did not like.

Table 3: Effects on Groups of Participants

Group	Variable	F value	Post hoc Tests
Listen to backgroud music	Valence before & after reading	3.07*	Slow-minor (0.25) vs. Environment (-0.75) *
Dislike the audio	EDA_MFD during reading	2.62*	Slow-major (0.10) vs. Environment (0.02)*
	EDA_MSD during reading	3.32*	Environment (0.02) vs. Slow-major (0.13)** Fast-major (0.03) vs. Slow-major (0.13)* Fast-minor (0.03) vs. Slow-major (0.13)* Slow-major (0.13) vs. Slow-minor (0.03)*
Familiar with the audio	Arousal before and after reading	3.68**	Fast-minor (-1.08) vs. Slow-major (0.60)* Fast minor (-1.08) vs. slow minor (1.00)*

**: $p < 0.01$; *: $p < 0.05$; values in parentheses are mean changes in that condition

6 SUMMARY AND FUTURE WORK

This study sought to explore the effect of different types of background music on learning, from the aspects of emotion change, metacognition and learning performances. An experiment was conducted with four types of background music varying in tempo and mode, as well as an environmental sound. Both behavioral and physiological data were collected from 33 graduate student participants for a more complete analysis. Preliminary results show that the different audio pieces had different effects on emotion change after focused listening and self-perceived engagement with the reading passages. However, they did not show significant differences on learning performances. A series of indicators were extracted from physiological signals during focused audio listening and text reading with background audio. Significant difference was observed on one indicator during focused listening but not during text reading when the audio served as background. This seems to imply that the emotion changing function of music is not as prominent when it is used for learning background as when it is listened with focused attention. More differences appeared when comparing the effects of the five audio conditions within sub groups of participants, suggesting the effect of background music may vary across individuals.

Further analysis will be conducted by comparing participant groups split with more fine-grained criteria such as music training background, attitude to the effect of background music on learning, etc. The analyses conducted so far have separately examined individual variables. Prediction models can be constructed in the future to explore the relationships and dynamics among the learning-related variables in behavioral and physiological levels.

7 ACKNOWLEDGEMENTS

This study is partially supported by National Natural Science Foundation of China: 61703357 and an Early Career Scheme grant from the Research Grants Council of the Hong Kong Special Administrative Region, China: HKU 27401114.

References

[1] S. Hallam, 2010. "The power of music: Its impact on the intellectual, social and personal development of children and young people," International Journal of Music Education, vol. 28, no. 3, pp. 269-289.

[2] 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.

[3] H. J. Kang and V. J. Williamson, 2014. "Background music can aid second language learning," Psychology of Music, vol. 42, no. 5, pp. 728-747.

[4] Schwartz, K. M. Ayres, and K. H. Douglas, 2017. "Effects of music on task performance, engagement, and behavior: A literature review," Psychology of Music, vol. 45, no. 5, pp. 611-627.

[5] W. F. Thompson, E. G. Schellenberg, and A. K. Letnic, 2012. "Fast and loud background music disrupts reading comprehension," Psychology of Music, vol. 40, no. 6, pp. 700-708.

[6] 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.

[7] T. Eerola and J. K. Vuoskoski, 2013. "A review of music and emotion studies: approaches, emotion models, and stimuli," Music Perception: An Interdisciplinary Journal, vol. 30, no. 3, pp. 307-340.

[8] R. Pekrun and R. P. Perry, 2014. "Control-value theory of achievement emotions," International Handbook of Emotions in Education, 120-141.

[9] R. M. Yerkes and J. Dodson, 1968. "The relation of strength of stimulus to rapidity of habit-formation," Punishment: Issues and experiments, pp. 27-41.

[10] E. G. Schellenberg, , 2012. "Cognitive performance after listening to music: A review of the Mozart effect," Music, Health, and Wellbeing, pp. 324-338.

[11] C. François and D. Schön, 2014."Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: the role of musical practice," Hearing Research, vol. 308, pp. 122-128.

[12] B. Cowley, N. Ravaja, and T. Heikura, 2013."Cardiovascular physiology predicts learning effects in a serious game activity," Computers & Education, vol. 60, no. 1, pp. 299-309.

[13] S. D’Mello, A. Kappas, and J. Gratch, 2018. "The affective computing approach to affect measurement," Emotion Review, vol. 10, no. 2, pp. 174-183.

[14] M. S. Hussain, R. A. Calvo, and F. Chen, 2014. "Automatic cognitive load detection from face, physiology, task performance and fusion during affective interference," Interacting with Computers, vol. 26, no. 3, pp. 256-268.

[15] P. Pham and J. Wang, 2016. "Adaptive review for mobile MOOC learning via implicit physiological signal sensing," in Proceedings of the 18th ACM International Conference on Multimodal Interaction, pp. 37-44.

[16] A. C. Strain, R. Azevedo, and S. K. D’Mello, 2013. "Using a false biofeedback methodology to explore relationships between learners’ affect, metacognition, and performance," Contemporary Educational Psychology, vol. 38, no. 1, pp. 22-39.

[17] A. Graesser, Y. Ozuru, and J. Sullins, 2010. "What is a good question?," Bringing Reading Research to Life. pp. 112-141.

[18] A. Furnham and K. Allass, 1999."The influence of musical distraction of varying complexity on the cognitive performance of extroverts and introverts," European Journal of Personality, vol. 13, no. 1, pp. 27-38.

[19] S. Böck, F. Korzeniowski, J. Schlüter, F. Krebs, and G. Widmer, 2016. "Madmom: A new Python audio and music signal processing library," in Proceedings of the 2016 ACM on Multimedia Conference, pp. 1174-1178: ACM.

[20] Y.-S. Huang, S.-Y. Chou, and Y.-H. Yang, 2018. "Pop Music Highlighter: Marking the Emotion Keypoints," arXiv preprint arXiv:1802.10495.
S. Losorelli, D. T. Nguyen, J. P. Dmochowski, and B. Kaneshiro, 2017.

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