The Effect of Music on the Developers' Performance During Small Programming Tasks

Kathrin Wardatzky
University of Zurich
Zurich, Switzerland
kathrin.wardatzky@uzh.ch

Luka Popovic
University of Zurich
Zurich, Switzerland
luka.popovic@uzh.ch

Vladimir Masarik
University of Zurich
Zurich, Switzerland
vladimir.masarik@uzh.ch

ABSTRACT

Listening to music while working is a popular method for software developers to avoid background noises and control their mood. There has been plenty of literature that examines the effect of music in the work context. These, however, mainly rely on self-perceived metrics. But how does the music actually affect the developers' cognitive state? In this work, we propose a study design that investigates the effect of different music styles using electroencephalographic (EEG) sensing. We furthermore include music that the participants selected themselves to account for the subjectivity of a person's music preferences. The study was conducted with 15 participants who performed programming tasks while listening to four different music styles. The results show a high variance between participants which indicates that not only the own taste but also the personality could influence the participants' performance while being exposed to music.

Author Keywords

Developer Productivity, Music, Biometric Sensing

CCS Concepts

•**Human-centered computing** → User studies;

INTRODUCTION

Listening to music during programming activities is popular among developers, particularly in shared office spaces [1]. However, how does this affect the productivity of the developers? Developers who listen to music while working report that music helps them to focus and to avoid distractions. Developers who do not listen to music, on the other hand, state that it would distract them and decreases their productivity [1]. In this study we plan to further investigate the influence of music on programming tasks. Previous work has analyzed the self-reported perception or task completion of the participants (e.g., [10, 11]). We want to extend these metrics by additionally measuring the participants' EEG signals during our experiment. We propose an experimental design that not only includes different pre-selected music, but also a condition in which the participants can choose the music up to their own taste. We aim to analyze if the participants' performance is dependent on the personal taste or on characteristics of a musical record in general. These characteristics can be the presence or absence of lyrics, a regular or irregular rhythm, temporal changes or a steady pace. Furthermore, we see this work as a first step of utilizing EEG data to analyze the effect of different music styles on developers' productivity.

In recent years, the tools to measure EEG activity have become less invasive. In this study, we investigate whether one of these tools, namely the brain sensing headband Muse 2[1], is capable of measuring the task engagement and task difficulty of developers while listening to different music styles. The Muse 2 headband contains four sensors, two on the forehead and one behind each ear. Due to it's noninvasive design, it has been used in a research context before (e.g., [2]), however not in the context of measuring the effect of music on productivity. Our study aims to fill that gap and we treat it as first step on the path to developing objective metrics for developers' productivity and flow. More precisely, we aim to answer the following research questions:

- **RQ 1:** Can our experimental setup, using small programming tasks, replicate the results of previous work on the increased attention levels and heightened positive mood of music on developers' performance?
- **RQ 2:** To what extent do the EEG signals reflect the performance and self-reported flow of the participants when listening to different music styles?
- **RQ 3:** Does the personal music taste of the participants have an influence on the performance?

To answer these questions, we conducted a within subject user study in which we exposed the participants to different music styles and measured their EEG activity while they were solving small programming tasks.

In the end, the results indicate that the average self-reported flow score is higher for music conditions that the participants preferred. It is also higher in the music conditions that involved focus increasing or self-selected music. However, there does not seem to be any clear correlation between the personal music preferences and task engagement index, task difficulty index or average blink rate.

[1] https://choosemuse.com/muse-2/

CHI '16, May 07–12, 2016, San Jose, CA, USA
ACM xxx-x-xxxx-xxxx-x/xx/xx. . . $15.00
DOI: http://dx.doi.org/xx.xxxx/xxxxxxx.xxxxxxx

One potential application case of the findings of our study would be to convince employers that music can help to focus and with that improve the productivity and workplace happiness of the employees.

The main contribution of this paper is a novel experimental design that

- measures the effect of different music with an EEG-sensor,
- compares the music condition with office background noise to create a more real-life scenario,
- investigates how different music styles as well as the own music taste influence developers' performance.

The remainder of this work is structured as follows. We first present relevant related work. Section 3 describes the methodology of our study while section 4 discusses the data extraction, cleaning and analysis process. Section 5 shows the results which are then discussed in section 6. Before we conclude in section 8, we discuss the threats to internal as well as external validity in section 7.

RELATED WORK

The related work can roughly be categorized into two sections, literature that deals with the effect of music on work performance, and literature that uses electroencephalographic sensing in research to determine productivity factors.

The effect of music on work performance

Barton et al. conducted a survey among 2,242 developers of which the majority frequently listen to music when writing code or doing repetitive tasks. They found out that the main motivation to listen to music is to obscure background noises and regulate the mood. Additionally, they found evidence that the use of music is related to a person's character [1]. This is supported by Furnham who compared the cognitive performance of introverts and extroverts in different music styles [5]. Lesiuk conducted an experiment to find out whether listening to music during a programming tasks has an effect on anxiety and task achievement. The author found evidence that listening to music decreases the anxiety level of the participants, however there was no statistically significant difference in task achievements [10]. In addition, another study by Lesiuk measured the effect of music listening on state positive affect, work quality and time-on-task of developers. The author concluded that state positive affect and quality-of-work were lowest with no music environment, while time-on-task was longest when music was introduced in work environment and then removed [11]. Lastly, Hakke conducted a survey among 295 participants with the main goal to explore the influence of music on well-being [5]. This study reveals that many respondents believed that musically induced positive mood at work can lead to an improved performance, such as increased capability to deal with stressful situations, and well-being in general.

Electroencephalographic sensing to determine work performance

In [3], Fritz et al. presented multiple psycho-physiological measures to assess task difficulty of software developers, among them electroencephalographic (EEG) sensing. As a follow up, Fritz and Müller focus on extracting a developer's cognitive and emotional state with the help of biometric sensing to increase productivity [4].

Telford and Thompson as well as Holland and Tarlow both [14, 6] investigated the relation of blinking and cognitive processes. They found signs that the blink rate of people is lower when they are focused on solving a problem.

Sammler et al. [13] combined the elements of music and biometric sensing in their work on music and emotions. They measured the EEG data while exposing their participants to pleasant and unpleasant music conditions to investigate the effect of music-induced emotions on the EEG measures.

METHODOLOGY

To measure a potential effect of music on developers' performance, we performed a user study with 15 participants. Each participant had the option to perform the programming tasks on their own laptops and we additionally asked them to bring their own headphones. With that, we aim to avoid measuring an effect due to using e.g. an unfamiliar keyboard layout or uncomfortable headphones. The following sections describe the overall procedure as well as each individual component of our study design.

Procedure

The study followed a within subjects design. Each participant was exposed to four different, randomly assigned, music conditions while performing coding tasks in the game CodeCombat[2]. After an introduction to the study and the fitting of Muse headband, we conducted a training session that consisted of the first four levels of the game. Each condition lasted for six minutes in which the participant solved coding tasks while listening to music over headphones. After the condition, we recorded how many levels the participant finished. We presented the solutions to the levels that the participant did not complete, as the levels build up on each other. We then asked the participants to fill out a short questionnaire to gather information about their self-perceived flow. Finally, each condition ended with showing the participants a fish tank video for two minutes to record the baseline EEG-data and avoid carrying over potential frustration from the previous coding tasks. After all four conditions were completed, we conducted a semi-structured interview and a survey to gather demographic information. The interview, survey and flow assessment questions are provided in the supplemental material[3].

Programming Tasks

We used the game CodeCombat for the programming tasks. CodeCombat is a game-based computer science program where students type real code in the Python programming language and see their characters react in real time. Figure 1 shows the CodeCombat interface. The participant types the code in the editor on the right part of the screen and directly sees the effect when executing it in the left part. In each level

[2] https://codecombat.com/

[3] https://github.com/kathriwa/HASE19_Music_Project

new basic programming concepts, like loops or conditional statements, are introduced. Participants were solving the training session as well as the four programming tasks by typing simple commands. This is closer to the real world tasks of a developer than the drag and drop action that many other programming games focus on. The first programming task was a training session that consisted of the first four levels of the game where we let participants to get familiar with the game and the coding environment in order to minimize a potential learning effect. The following four tasks consisted of 6-7 levels each.

Figure 1: CodeCombat interface

Music Selection

In total we used four different music conditions: focus, distracting, own choice and office noise. The music selection for each condition was based on the knowledge from related literature [1, 5]. The first two conditions are opposing music styles, one with which we aim to increase the focus, and one that is supposed to distract the participants from the programming task. As music is highly subjective and to avoid measuring an effect because the participants dislike the music in the aforementioned two conditions, we add a condition in which the participants choose the music up to their taste. The final condition serves as baseline. Previous studies compared their results to the participants' performance during silence. However, developers who work in shared offices will rarely have the choice between music and silence [1]. Therefore, we played a track with office background noises during this condition. The order of the music conditions was random. In the following paragraphs, we explain how we defined each condition.

Distracting

According to Barton et al. [1], music containing vocals in the participant's native language is more distracting than instrumental music. We recruited participants with an adequate command of the English language and therefore opted to choose music with English vocals in the distracting condition. To counterbalance potential differences in the English skills of our participants, we additionally required the music to have an irregular rhythm and temporal changes. These features increased the complexity of a song and with that the potential to attract the participants' attention [5]. We selected the songs *A Modern Leper* by *Frightened Rabbit* followed by *Our Retired Explorer (Dines With Michel Focault in Paris, 1961)* by *The Weakerthans*.

	Average	Min	Max
Age	24 years	21 years	40 years
Level of English Skills	7.8/10	7/10	10/10
Programming Experience	5.7 years	1.5 years	12 years
Python Experience	1.36 years	0 years	4 years
Software Developer Experience	2 years	0 years	10 years

Table 1: Overview of the participants

Focus-oriented

The music selection process in the focus-oriented condition aimed to fulfill the opposite criteria as in the distracting condition. We decided for an instrumental song with a regular, slow-paced rhythm. In the focus increasing condition, we played *Sigur Ros - Sleep 1*.

Own choice

We asked each participant to send us three songs that they would listen to when coding. If they do not listen to music while coding, we asked them to select songs that they like. The songs were played in the order that the participant sent them to us. Where possible, we used the Spotify API[4] to extract the genres of the songs. An overview of the songs and their genres of this condition is provided in the supplementary material.

Background Noise

As each participant might work in a different environment, we opted to play standard office background noises in this condition[5].

Participants

For our experiments, we recruited 15 participants (6 male, 9 female). All but two participants are enrolled in a master or PhD program at the Institute for Informatics of the University of Zurich. The other two participants are one MSc and one BSc student from ETH. Their age ranges from 21 to 40 with the average age being 24. They have an average overall programming experience of 5.7 years with the minimum being 1.5 years and the maximum 12 years. Nine of the participants reported that their Python programming experience is one year or less. The software development experience ranges from $<$ 1 year (5 participants) to 10 years (1 participant).

Table 1 gives an impression of the average participant as well as the maximum and minimum values. The level of English skills was assessed on a scale from 1 (very bad) to 10 (native speaker). We additionally asked for the native language which 7 participants answered with *German*. The remaining participants' native languages were *Dutch, English, French, Italian, Persian, Russian, Serbian,* and *Spanish* with one answer each.

DATA ANALYSIS

This section describes the data analysis process from the data collection over the cleaning to the analysis. Details about the collected data will be given in the respective subsections.

Data cleaning

For each participant we extracted the same collection of data. We removed all values from the recordings for which the signal

[4] http://organizeyourmusic.playlistmachinery.com/index.html
[5] https://www.youtube.com/watch?v=D7ZZp8XuUTE

strength had an hsi-precision above 2, which indicates a poor connection of the sensors.

As each person has a different personal power spectrum distribution [3], the data was normalized by calculating a baseline from the final minute of the fish tank recording after each condition. This baseline was subtracted from the absolute values which were then used for all further analysis.

We aim to investigate the effect of music on focus and flow. As it takes certain time to get into a flow state, we only used the final minute of each recording for our analysis.

EEG data

The use of Muse headband records a plethora of data points[6]. For our analysis, we focused on the blink rate, as well as the absolute alpha (α), beta (β) and theta (θ) frequency bands. Literature suggests that a decrease of alpha and an increase of theta are indicators for an increase in attentional demand and working memory load [3]. Furthermore, these values allow us to calculate the Task Engagement Index (TEI) as well as the Task Difficulty Index (TDI). For the recording of the EEG data, we used the application *Muse Direct*.

Brain wave values are a logarithm of the sum of the Power Spectral Density of the EEG data over some frequency range, for example Alpha frequency. Since it is a logarithm, the values are on a log scale, and the units are Bels. The Muse device records these values for each of it's four sensors [7].

Task Engagement Index

The Task Engagement Index is defined as $\frac{\beta}{(\alpha+\theta)}$ [9, 3]. It is based on the observation that with an increase in task engagement, θ decreases, while α is blocked and β increases in relative power [3].

Task Difficulty Index

Based on the same observations that led to the TEI, the Task Difficulty Index is defined as $\frac{\theta}{(\alpha+\beta)}$ and serves as additional measure of task difficulty.

Blink rate

The Muse device not only captures the EEG values mentioned above, but it also records when a participant blinks. Fritz et al. [4] mentioned that the blink rate can give insights on stress and anxiety levels, visual attention and mental workload. In stressful situations, the blink rate can increase significantly while it decreases if the participant is more focused on the task [4]. It serves us as an additional measure to support the TEI and TDI values computed from the EEG data.

Self-assessed flow

We are aware of the limitation that getting into a state of complete flow during a 6-minute-task is unlikely. However, Martin and Jackson [12] addressed exactly this limitation and developed a questionnaire to assess task absorption and enhanced subjective experience during short tasks. This questionnaire is based on the original 36-item flow scale by Jackson and Marsh [8]. Martin and Jackson reduced this by only using one item out of each of the nine flow dimensions that are rated on a 7 point Likert scale [12].

The self-assessment questionnaire gives us the opportunity to see whether the self-reported flow is congruent with our findings from analyzing the EEG data. Additionally it gives us the opportunity to put our results into relation with previous work that was relying on self-perceived metrics.

We have computed the average flow score for each participant and each condition. After normalizing the data we put them into relationship with the extracted data from the interviews. More specifically, we have created a summary containing every flow score for the music which the participants liked, disliked or were indifferent to. This resulted in a table that indicates whether the musical conditions preferred by the participant are correlated with the self-reported flow score.

Furthermore, we have computed a summary for each musical condition from the flow score. Thus, creating a matrix that should show how each musical condition influenced the self-reported flow score on average.

Interview

At the end of each experiment we conducted a semi-structured interview. The interview questions can be found in the supplementary material.

The goal of this interview was to get more information on the participants' music listening habits during work and in their leisure time. Furthermore, we used it to assess the perceived difficulty of the programming tasks.

We first independently coded one third of the transcripts, followed by a discussion of the codes and a generation of a codebook. Later we have finished coding of the rest of the interviews.

The outcomes were a summary matrix of how many people disliked, liked or were indifferent towards the music conditions played, and a summary of what conditions were considered difficult for the participants. In case a participant explicitly mentioned that they enjoyed a music condition or explicitly said they disliked it, we have noted it as 'liked' and 'disliked' codes. If the participants did not mention the music condition or mentioned that they felt indifferent towards it, we have coded that as 'indifferent'. As for difficulty of the levels, we have used the same method as for the music preference coding. However, here the participants specified whether the condition was hard, easy or they felt indifferent.

RESULTS

This section presents the results of our study.

After removing recordings that did not fulfill the requirement of having a good signal connection for 2/3 of the time, 12 participants were left for our analysis.

Our hypothesis is that the task engagement increases with the focus increasing music type while at the same time the task difficulty and blink rate decreases. For the distracting music

[6]see https://musemonitor.com/FAQ.php#Algorithms for more information

type and office background condition, the assumption would be opposite.

The results show a high variation between the different participants and also between the different measures which will make a between subject comparison difficult.

To validate our results, we performed an exploratory repeated measures ANOVA (RM-ANOVA) with our four music conditions as within-participants factor and the TDI, TEI and blink-rates as dependent measure.

Task Engagement Index (TEI)
Figure 2 shows the average TEI values of each condition for all participants that were not removed due to bad EEG signal. The labels for each bar indicate the opinion that the respective participant had on that music condition. This information was taken from the interviews.

There is no visible pattern for the task engagement between the conditions or the personal preferences. For some participants, the TEI value decreases with the e.g., distracting music (P03, P10), while for others the task engagement in this condition increases (P01, P02).

The RM-ANOVA test returned $F_{3,33} = 1.0713$ which means there are no statistically significant differences between the conditions.

Task Difficulty Index (TDI)
Figure 3 visualizes the average TDI values of each condition for all participants with proper EEG signal. As mentioned above, the labels indicate the participants' opinion of the music.

For the majority of the participants, the task difficulty results are close to the baseline. Half of the participants (P03, P04, P07, P08, P09, P13) had a lower TDI value than in the baseline condition while for P08 and P13 the high TDI values for the distracting and own choice respectively stand out.

The RM-ANOVA test returned $F_{3,33} = 1.8000$ and with that no statistical significant differences between the conditions.

Blink rate
The average blink rates per second are shown in Figure 4 using the same notation as above. There is no clear pattern visible between the music conditions. Some participants (e.g., P07) have a clear distinction between the music conditions, others (e.g., P02, P03) have very similar results. One third of the participants (P01, P05, P07 and P15) had a lower blink rate in the conditions with music that they liked.

Similar to the measures above, the RM-ANOVA did not show any statistically significant differences between the conditions ($F_{3,33} = 0.3857$).

Interview
Out of the 15 interviewed participants, 9 participants regularly listen to music when they are coding. The main reason for the participants to listen to music at work is to block background noise (8 mentions). Other reasons include that they just want to listen to music, to avoid boredom when doing repetitive tasks or to improve focus.

	Focus	Distracting	Self-selected	Office sounds
Liked	5	5	6	7
Indifferent	7	4	9	5
Disliked	3	6	0	3

Table 2: Overview of the participants music preferences

	Flow score
Liked	3.78
Indifferent	-1.42
Disliked	-2.36

Table 3: Summary of the flow score based on music preference of the participant

Table 2 summarized the responses to the question which music the participant preferred most/least for the music condition. The majority of people who indicated a negative preference named the distracting condition as least preferred music. The office sounds were liked by a surprisingly high number of participants.

When it comes to the music preferences, five participants stated that it depends on their moods when selecting the music for either programming or in their leisure time. Only few participants named specific genres that they prefer, the majority stated that the music selection also depends on their mood.

Self-perceived flow
The sum of average flow score for each music condition that a participant liked, disliked or was indifferent towards can be seen in table 3. On average, participants who mentioned that they enjoyed a music condition also reported higher flow score for that condition. Table 4 shows that on average the flow score was higher during the focus increasing condition and the own music selection, and contrary, the self-reported flow was lower when distracting music or office noise were played.

Task Completion
On average the participants finished 3.1 levels during the first two tasks no matter what music was played, 2.6 levels during the third task and 2.9 levels during the final task. The table with the number of completed levels for each individual participant is part of the supplementary material. During the focus increasing music an average of 3 levels were completed, in the distracting condition 2.3 levels were finished, while 3.36

	Flow score
Focus	1.97
Distracting	-2.58
Own selection	3.08
Office noise	-2.47

Table 4: Summary of the flow score for each music condition on average

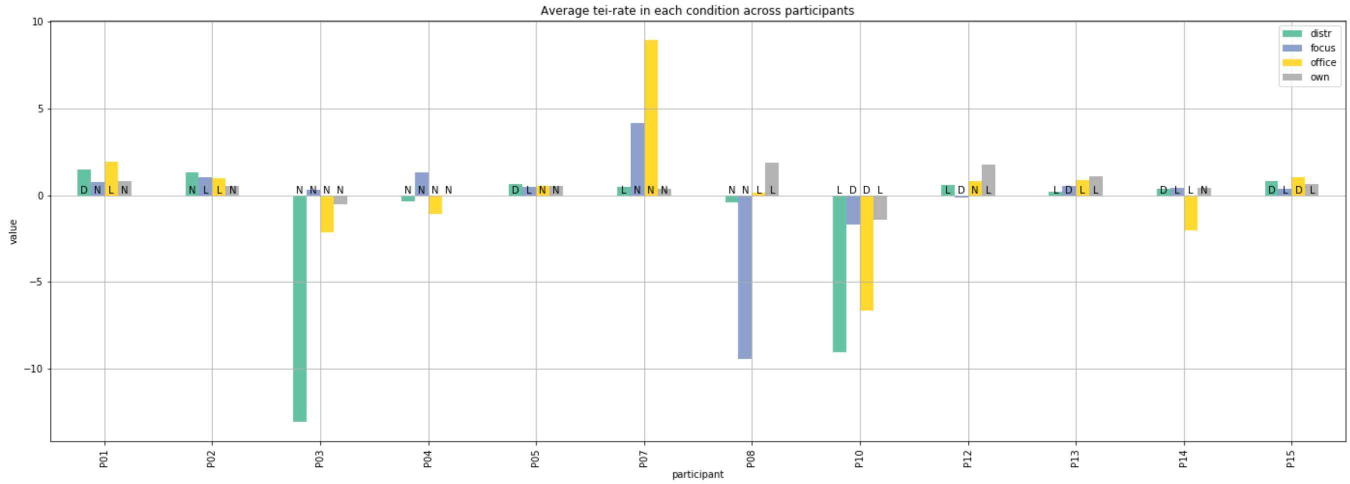

Figure 2: The average TEI values for each participant and each condition including their opinion on the music condition (D=dislike, N=neutral, L=like)

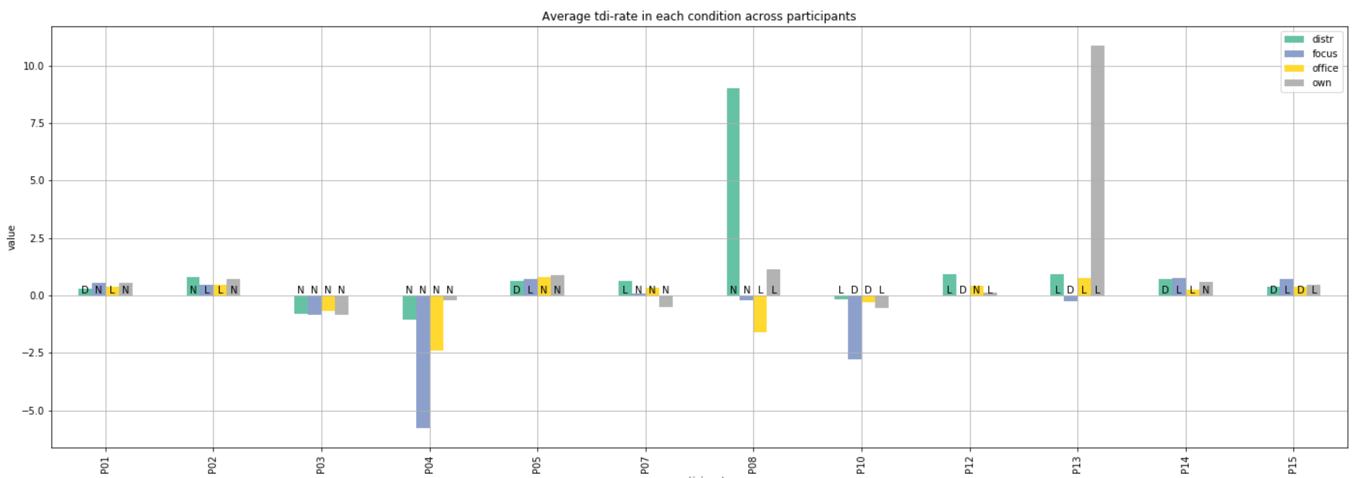

Figure 3: The average TDI values for each participant and each condition including their opinion on the music condition (D=dislike, N=neutral, L=like)

and 3.09 levels were completed during the office background noise and own selection respectively.

DISCUSSION

This section discusses the results that were described in the previous section and answers our research questions. Although the EEG data did not return statistically significant results between the conditions, we want to see if we can find ideas in the interview and flow survey data that could explain the EEG results.

One can see on first glance in Figures 2 - 4 that the results contain outliers. Although participants had a preference for music, there is no consistent pattern that reflects this in the EEG data. One explanation could be that the questions we asked the participants was focused on whether they like the music, not whether they felt focused. To support this assumption, we would need to further investigate the individual responses of the flow survey, which we left for future work.

The low effect of the office background condition could be explained with the fact that the track that we played did not contain clearly spoken words. That assumption is backed by the interview data. Many participants stated that they listen to music at work because it distracts them when other people in the office talk. P09 even stated that s/he enjoyed this condition. *[...] I was thinking maybe I should play it. I guess because there was no clear speaking that you can really get to know what they are talking about, because my own problem is when people in like very small room talking topics that I have ideas about I kind [of] get involved.* In contrast to that, P10 stated: *It was fine, except from the [office noise]. I hate those people around me. I can't really concentrate when I hear voices and I never work without my headphones. I tried, but it doesn't work.*

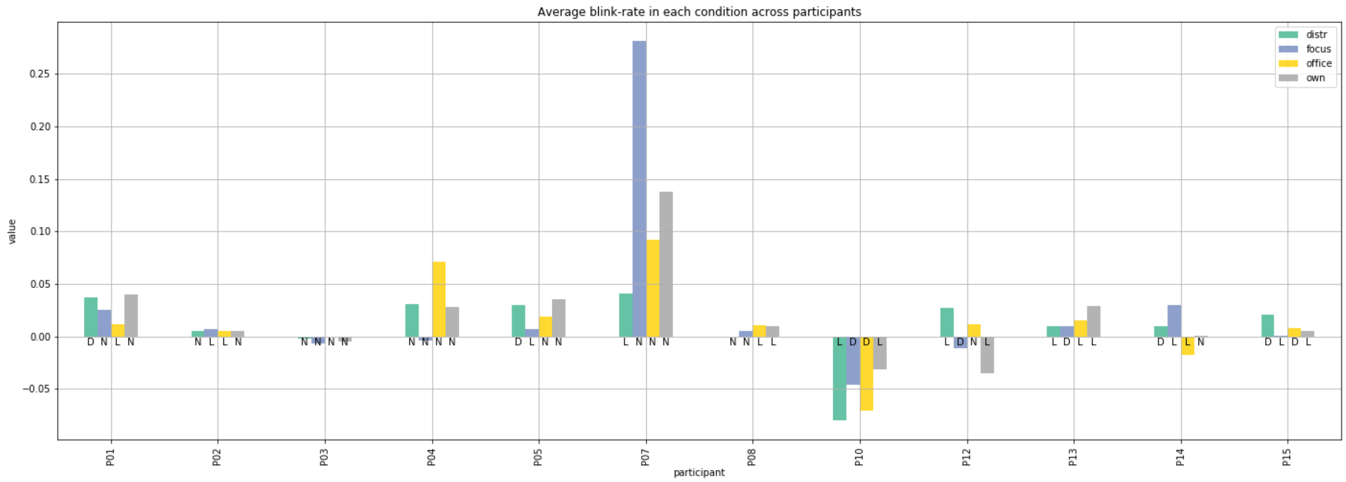

Figure 4: The average blink rates per second for each participant and each condition including their opinion on the music condition (D=dislike, N=neutral, L=like)

. The TEI values for P10 were also lowest for the distracting and office noise conditions. This strengthens our assumption that the individual preferences and personalities can have a strong effect on the self-perceived focus.

We answer our research questions as follows.
RQ 1: *Can our experimental setup, using small programming tasks, replicate the results of previous work on the increased attention levels and heightened positive mood of music on developers' performance?*
The selection of our metrics cannot report any results on the effect of music on the mood of the developer. However, many participants stated that their mood is the main selection criteria for the music that they listen to. Unfortunately, our EEG analysis also did not show any significant effect on the increased attention level.

RQ 2: *To what extent do the EEG signals reflect the performance and self-reported flow of the participants when listening to different music styles?*
As the results for the EEG signals do not contain significant differences between the conditions, we cannot answer this question.

RQ 3: *Does the personal music taste of the participants has an influence on the performance?*
There is no evidence in the EEG data that suggests an overall better performance when coding with the self-selected music. However, none of the participants reported a strong dislike of our music selection which was our intention when choosing the music. It would be interesting to see the results when selecting more extreme music.

Overall, our results indicate that the effect of music depends on the individual person and that the self-perception does not necessarily correlate with the EEG data.

THREATS TO VALIDITY

As with all experimental studies, there are several threats to validity to this work. This section discusses the threats as well as the methods that we applied to limit the threats.

Internal Validity

In terms of internal validity, our biggest threats concern the selection of the music and programming tasks.

Music Selection

None of the authors has extensive knowledge in music theory and it was out of scope of this project to consult an expert. We addressed this limitation by extracting features for distracting and focus increasing music from existing literature and had a test run to confirm the desired effect on the participant.

Programming Tasks

Due to the gamification character of our programming tasks, the levels build up on each other. We aimed to group the programming tasks for each condition to have a similar level of difficulty. However, this was not 100% possible. To further see if certain programming tasks were more difficult than others, we additionally asked the participants in the interviews to name the programming tasks that were the most difficult and easiest.

Duration of Each Condition

With four conditions in a within subjects study, we had to balance between having short enough conditions to avoid measuring an effect due to fatigue of the participants while giving the participant enough time to get into a flow-like state. The six minute duration makes it difficult to get into a full flow state, however any longer duration would have increased the overall duration of the experiment to an unreasonable extend.

External Validity

In terms of external validity, the main threat comes from the participant selection. Our participants are predominantly master students from the University of Zurich. To evaluate how well the results of this study generalize, the study would have to be replicated with different user groups consisting of professionals at different experience levels.

CONCLUSIONS AND FUTURE WORK

In this paper, we presented a novel user study that aimed to measure the effect of music during small programming tasks by evaluating the task difficulty and task engagement derived from EEG data measured in a novel way using Muse headband device. We performed a within-subject user study with 15 participants and exposed them to different music styles while performing small programming tasks.

While this particular study design did not yield statistically significant results from the EEG data, it opened an interesting path for further exploration. Future work could investigate whether longer programming tasks would yield better results. A direction to explore the differences in the measurements between the participants would be a connection to personality measures. Previous work already found evidence that different music has a different effect on people with different personalities (mentioned e.g. by Lesiuk [11]). Our results show big variations between participants. Their different personalities could be one potential reason for that.

REFERENCES

[1] Laura E Barton, Gülipek Candan, Thomas Fritz, Thomas Zimmermann, and Gail C Murphy. 2019. The Sound of Software Development: Music Listening Among Software Engineers. *IEEE Software* (2019).

[2] Karen Anne Cochrane. 2019. Reconnecting the Body and the Mind: Technology to Support Mindfulness for Stress. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, DC05.

[3] Thomas Fritz, Andrew Begel, Sebastian C Müller, Serap Yigit-Elliott, and Manuela Züger. 2014. Using psycho-physiological measures to assess task difficulty in software development. In *Proceedings of the 36th international conference on software engineering*. ACM, 402–413.

[4] Thomas Fritz and Sebastian C Müller. 2016. Leveraging biometric data to boost software developer productivity. In *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, Vol. 5. IEEE, 66–77.

[5] Adrian Furnham and Kathryn Allass. 1999. The influence of musical distraction of varying complexity on the cognitive performance of extroverts and introverts. *European Journal of Personality* 13, 1 (1999), 27–38.

[6] Morris K Holland and Gerald Tarlow. 1975. Blinking and thinking. *Perceptual and motor skills* 41, 2 (1975), 403–406.

[7] Interaxon. 2018. Technical Documentation. (2018). https://web.archive.org/web/20181105231756/http://developer.choosemuse.com/tools/available-data

[8] Susan A Jackson and Herbert W Marsh. 1996. Development and validation of a scale to measure optimal experience: The Flow State Scale. *Journal of sport and exercise psychology* 18, 1 (1996), 17–35.

[9] III Lawrence J Prinzel, Pope Alan T, Freeman Frederick G, Scerbo Mark W, and Mikulka Peter J. 2001. Empirical analysis of eeg and erps for pyschophysiological adaptive task allocation. (2001).

[10] Teresa Lesiuk. 2000. The effect of music listening on a computer programming task. *Journal of Computer Information Systems* 40, 3 (2000), 50–57.

[11] Teresa Lesiuk. 2005. The effect of music listening on work performance. *Psychology of music* 33, 2 (2005), 173–191.

[12] Andrew J Martin and Susan A Jackson. 2008. Brief approaches to assessing task absorption and enhanced subjective experience: Examining 'short' and 'core' flow in diverse performance domains. *Motivation and Emotion* 32, 3 (2008), 141–157.

[13] Daniela Sammler, Maren Grigutsch, Thomas Fritz, and Stefan Koelsch. 2007. Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology* 44, 2 (2007), 293–304.

[14] C. Telford and N. Thompson. 1933. Some factors influencing voluntary and reflex eyelid responses. *Journal of Experimental Psychology* 16 (08 1933), and 4 second intervals show smaller averages. DOI: http://dx.doi.org/10.1037/h0071694