The Role of Energetic Music in a Video Game: Analyzing its Effect on Immersion, Perception and Performance

Teodor Kostovski[1], Matjaž Kljun[1,3] and Klen Čopič Pucihar[1,2,3]

[1]University of Primorska, Faculty of Mathematics, Natural Sciences and Information Technologies, Koper, Slovenia
[2]Faculty of Information Studies, Novo Mesto, Slovenia
[3]Stellenbosch University, Department of Information Science, Stellenbosch, South Africa

Abstract

Energetic music has the potential to enhance immersion but can also negatively impact task performance. This study measured the effects of music on a video game suited for energetic music using both subjective (questionnaires) and objective (facial and physiological analysis) measures. Participants played under two conditions: with and without music. While the music condition led to higher reports of immersion, no significant difference between the two conditions was found. Performance scores were better without music, suggesting that energetic music, while increasing immersion, may disrupt performance by intensifying the task. These findings emphasize the importance of aligning immersion with task relevance, as the effects of music on performance are context-dependent. Further investigation into the balance between engagement, task perception, and performance is needed.

Keywords

energetic music, video game performance, immersion, gameplay perception, cognitive effects of music

1. Introduction

Music has been shown to have a significant negative effect on cognitive performance, especially on high-intensity tasks [1]. However, energetic music has been also shown to increase immersion, particularly in video games [2]. This challenges preconceived notions that higher immersion is directly related to higher performance and warrants further investigation into the exact relationship between music as a positive factor for immersion, yet a potential negative factor for performance.

A study by [3] suggests that music familiarity might play a role on how daunting a given task is perceived, which the researchers attributed to higher stress levels and as a result, the performance of the task. In detail, they found that music, specifically selected by the participant, had mostly a positive effect on task performance. Conversely, task performance was lower when the participant was exposed to unfamiliar, high-arousal music chosen by the experimenter.

HCI SI 2024: Human-Computer Interaction Slovenia 2024, November 8th, 2024, Ljubljana, Slovenia
89211084@student.upr.si (T. Kostovski); matjaz.kljun@upr.si (M. Kljun); klen.copic@famnit.upr.si (K. Čopič Pucihar)
0000-0002-6988-3046 (M. Kljun); 0000-0002-7784-1356 (K. Čopič Pucihar)
© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)
CEUR-WS.org/Vol-3866/paper5.pdf
CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

Another study [4] found that a slow-tempo song affected cognitive processing speed negatively, as opposed to a higher-tempo song, which increased cognitive processing speed. These seemingly contradictory conclusions suggest a nuanced relationship between music and task performance.

Given these findings, this study aims to explore the effect of energetic music on task immersion and perception while playing a vertically scrolling video game or vertical scroller type of a video game[1]. Specifically, we will investigate how varying levels of immersion impact task performance, and if higher immersion is indeed related to higher task performance. By understanding these dynamics, the research seeks to clarify the relationship between music, immersion, and cognitive performance in this specific context of a vertical scroller video game.

We have formulated the following research questions to guide our study:

RQ1 "*How does energetic music during gameplay of a scrolling dodge video game affect the level of immersion experienced by the players?*",

RQ2 "*How does energetic music during gameplay affect game perception (workload, pace, arousal)?*" and

RQ3 "*How does game immersion and perception affect overall task performance?*"

To address these research questions, the study aims to achieve the following objectives:

- Develop a vertical scroller video game designed to measure player immersion under different music conditions.
- Conduct a user study involving participants who will play the game with and without music.
- Collect data on player immersion and performance using a combination of questionnaires, eye tracking, facial expression, skin conductivity, and game scores.
- Evaluate the impact of energetic music on immersion, and how immersion is related to overall task performance.

2. Literature Review

2.1. Music and Gameplay Performance

The relationship between video game immersion and music is already well documented. It is known that music and sound effects play a large role in video game immersion, but studies into this topic show inconclusive objective results, suggesting a more nuanced relationship between the music provided in the video game and video game immersion. One study [2] found that participants reported a higher level of immersion when playing a high intensity, fast paced video game, which included music and/or sound effects compared to the same video game without music and sound effects. Interestingly, the only statistically significant results were from subjective measures, such as the Game Experience Questionnaire (GEQ), while objective measures, including psychophysiological recordings of electrodermal activity (EDA) and facial muscle activity (EMG), did not yield significant results.

[1]Vertically scrolling video game https://en.wikipedia.org/wiki/Vertically_scrolling_video_game

This lack of objective conclusion presents an opportunity for this study to utilize objective measures of immersion, such as pupil size, electrodermal activity, in order to find a more objective determiner of immersion.

2.2. Music and Cognitive Performance

Research has already extensively documented the effects of music on cognitive performance, showing varied results based on the type of music and task. One study [5] found that background music can impair performance on complex cognitive tasks, especially those that require memory and comprehension. On the other hand, other research [6] found that calming music had a positive impact on cognitive tasks, followed by no music, and finally energetic, high-intensity music, which had a negative effect. Another study [7] found that music had a positive effect on cognitive performance, but only when referring to simple, repetitive tasks.

Interestingly, one article [8] suggested that the relationship between music and cognitive performance is more complex and dynamic than just observing music affects on immersion. It suggests that the impact of music on immersion and performance depends on various external factors.

2.3. Impact of Music on Task Perception

Several studies [1, 3] noted that unfamiliar, high-intensity music increased stress levels, suggesting that energetic music might induce a sense of panic, pressure and a need to perform, which directly affect how a task is perceived. While other studies [9] found that calm music lowered anxiety levels in participants. While these studies produced significant findings, it is worthwhile to discuss the aspect of compatibility of music with a given task. For instance, in one study [1], fast-paced music was played while assessing reading comprehension, raising questions about the appropriateness of the music for the given task.

Consequently, our study aims to address the issue of musical compatibility and nuance by creating an environment where highly energetic music is appropriately matched to the task intensity, such as in a vertical scroller video game. In this type of game the player views the action from a top-down perspective, with the background scrolling from the top to the bottom of the screen, setting a rhythm for gameplay, and requiring the player to respond rapidly to the scrolling environment. By selecting music that aligns with the nature of the task, we tried to ensure that the music enhanced rather than detracted from the task experience. As a result, to effectively explore the impact of music on task perception and performance, we have developed a video game specifically for this user study. This game, serves as the experimental environment for our study.

3. Design and Implementation of the Video Game

The game developed for this user study is titled Dance Dance Mania, a Dance Dance Revolution inspired vertical scrolling and timing-based game designed with a colorful 8-bit retro aesthetic to maintain player engagement. We developed our own game to avoid the familiarity with the existing games which could benefit some uswers, and to better control the game exectution (same scenario, the increasing of the pace, time of the game run, etc.). With existing games this was not possible and too many other variables would be introduced.

In the game developped players need to press the correct arrow keys in sync with arrows falling down the screen when these overlap with the gray arrows on the bottom of the screen as shown in Figure 1. The game starts slowly, with arrows taking 4.5 seconds to reach the bottom and appearing every 2 seconds, allowing players to adjust. As the game progresses, the speed increases, with arrows moving 6 times faster (0.75 seconds to the bottom) and generated at 2 arrows per second, creating a more challenging experience within the 120-second (2 minutes) round.

The scoring system is simple: a “Perfect” hit (correct timing of pressing the arrow key when the falling arrow completely overlaps with the gray arrow) adds 2 points to the score, a “Good” (slightly mistimed pressing of the arrow key when the falling arrow slightly overlaps with the gray arrow) adds 1 point, while a “Wrong” key press or a “Missed” arrow deducts 1 point. At the end of each round, the player’s score and performance details (number of Perfects, Goods, Wrongs, and Misses) are displayed. Both the game’s music and sound effects follow the retro theme, creating a cohesive audiovisual experience.

A pre-determined sequence of arrow generation ensures all participants face the same difficulty level. Participants play four rounds of the game, experiencing either music in the first and third rounds or the second and fourth rounds, based on their assigned starting condition.

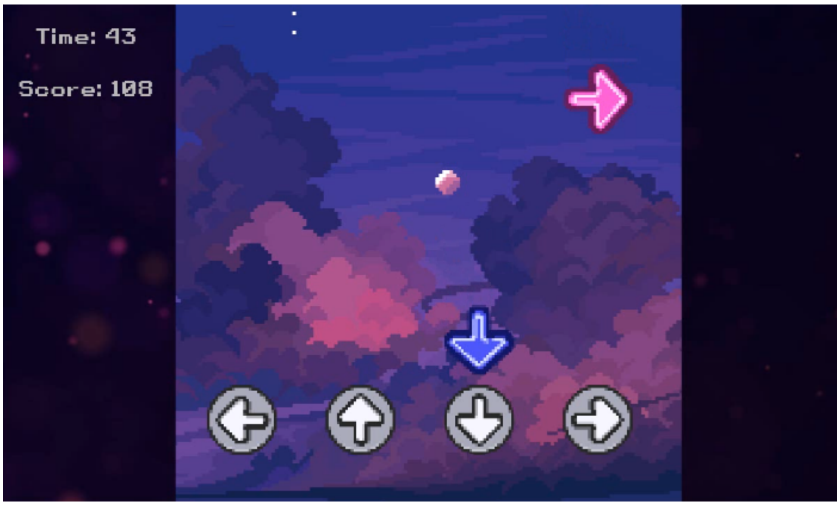

Figure 1: Screenshot of the video game during a gameplay session. The image showcases the game interface and the current state of the player’s progress. The arrows scroll top to bottom and when the blue ↓ arrow hits its corresponding grey area on the bottom the user needs to press the ↓ arrow on the keyboard. If the keyboard button is pressed when the blue arrow fully covers the corresponding area (“Perfect” score) user gets two points, if it just touches it (“Good”) 1 point, and if it is pressed while not touching the corresponding area or the wrong key (“Wrong”) or no key (“Miss”) is pressed one point is deducted. New arrows appear on the top and move downwards at a rate of one arrow every 2 seconds with the falling speed increasing from 4.5 seconds from top to bottom to 0.75 seconds at the end of a 2 minutes round.

4. Methodology

4.1. User Study Design

Participants were recruited through university boards and word-of-mouth. A total of 10 participants, aged between 20 and 23, signed up for the study. There was no criteria for participant selection, and no incentives were offered for participation. All participants were university students. They were informed that the study involved playing a video game under differing conditions, but information regarding the study's expectations, hypothesis, and research questions was not disclosed.

The study was conducted in a controlled environment to minimize external variables. The setup involved the use of iMotions software to collect facial expressions, eye tracking data, and galvanic skin response together with heartbeat. Participants were randomly assigned to one of two conditions, while ensuring that each had an equal number of participants.

4.2. Procedure

Upon arrival, participants were briefed on the study's procedure. They were then asked to sign informed consent forms, which explained their rights, the voluntary nature of their participation, and the confidentiality of their data.

Next, the experimenter showcased a 10 second demo of the game, to help participants understand the controls, the mechanics, and what different sound effects refer to. This demonstration did not involve the use of music. Following that, the experimenter explained different hardware that would be used to monitor the participants immersion level. The Shimmer 3 GSR Unit was applied to the participant's non-dominant hand to avoid interference with gameplay controls.

Next, participants played four rounds of Dance Dance Mania each lasting 2 minutes. In one condition the participants experienced music and sound effects in the first and third round of the game and only sound effects in the second and fourth round. In the other condition participants experienced music and sound effects in the second and fourth round of the game and only sound effects in the first and third round. After the second round, participants were given a brief break to relax.

We opted for four rounds of the game to make sure the participants got familiar with the game during the first two rounds and knew what to expect in the *music condition* and *no-music condition*. We then only took the results of the third and fourth round. This ensured that the majority of the learning happened during the first two rounds and the learning effect is smallest between round 3 and 4. Following the third round, participants completed the NASA-TLX and the Immersive Experience Questionnaire (IEQ), documenting their experience specific to that round. After the fourth round, they completed the same questionnaires, this time focusing on their experience in the final round. Finally, participants filled out a custom-made questionnaire about their music and gaming history, as well as demographic information including age and gender. At the end of the study, the participant's game scores were saved and labeled accordingly. For a quick summary of the study procedure, refer to Table 1.

Table 1
Table depicting study procedure

Time	Order music/no-music	Order no-music/music
	Consent form and game demo	
	Attach GSR unit	
	Calibrate eye tracker	
2 minutes - round 1 (R1) gameplay	Music condition	No-music condition
	Calibrate eye tracker	
2 minutes - round 2 (R2) gameplay	No-music condition	Music condition
5 minute break	Participant walks around and relaxes	
	Calibrate eye tracker	
2 minutes - round 3 (R3 gameplay)	Music condition	No-music condition
Up to 10 minutes	NASA TLX and IEQ questionnaires	
	Calibrate eye tracker	
2 minutes - round 4 (R4) gameplay	No-music condition	Music condition
Up to 10 minutes	NASA TLX and IEM questionnaires	
Up to 10 minutes	Participant demographic and musical history questionnaire	

4.3. Data Collection Instruments

4.3.1. Questionnaires

The questionnaires completed after round 3 and round 4 are the aforementioned NASA-TLX task (work)load index[2] and the Immersive Experience Questionnaire (IEQ) developed by [10]. NASA-TLX is a widely recognized subjective, multidimensional assessment tool used to evaluate perceived workload, allowing for the assessment of a task, system and other performance-related factors.

The IEQ was slightly changed by omitting questions not pertinent to this particular game, such as questions about in-game events and story. Most questions were rated on a 7-point Likert scale, while others required participants to select discrete answers such as “Yes”, “No”, or “Maybe”.

The questionnaire related to the participants’ music and game history included items designed to identify potential outliers and provide a foundation for future research.

4.4. Measuring Physiological Response

In order to measure Physiological response, the following iMotions products were used:

- Affectiva AFFDEX [11] - utilizing a Logitec webcam to capture facial expressions at 720p. For each timestamp, Affectiva provides a set of facial expression and a value ranging from 0 to 100 for each facial expression (a higher value means that particular emotion is more expressed or noticeable), For some emotions that can be either negative or positive (i.e Valence), a value ranging from -100 to 100 is provided instead (less than 0 means negative valence, more than 0 means positive valence)

[2]NASA-TLX https://en.wikipedia.org/wiki/NASA-TLX

- Tobii Pro Spectrum screen for eye tracking purposes. At each timestamp, we are provided with multiple values related to eye movement. For the purpose of this study, we will use pupil size parameter.
- a Shimmer 3 GSR Unit attached to the participants non-dominant hand. The Shimmer measures electrodermal reactions and heart rate. The main electrodermal reaction used in this study is GSR Conductance, which reflects changes in skin conductance related to physiological arousal.

4.5. Game Scores

For each participant, data was collected for the number of “Perfects”, “Goods”, “Misses”, and “Wrongs” within each round of gameplay, as well as the final score for that round. Additionally, the round number and whether or not the round included music were recorded.

5. Data Analysis Methods

5.1. Analysing Video Game Immersiveness

NASA-TLX was score was derived using a weighted sum of the participants’ responses, with the weights provided by a manual on how to calculate NASA-TLX scores [12]. For IEQ negatively scored items (i.e., items indicating a lack of immersion) were inverted (e.g., a score of 1 becomes 7, a score of 2 becomes 6, etc.), and then the responses were summed. The method used for calculating the results from the questionnaires is based on practices in previous research [13].

We will utilize the following metrics provided by the aforementioned iMotions tools:

- Engagement: Reflects the level of involvement with the task.
- Valence: Indicates the positive or negative emotional tone.
- Attention: Measures how attentively a participant is focused.
- Brow Furrow: Associated with concentration or confusion.
- Eye Widen: Often a sign of surprise or heightened attention.
- Lip Press: Can indicate stress or concentration.
- Lip Suck: May be a sign of nervousness or deep thought.
- Mouth Open: Could signal surprise or high arousal.
- Smile: Reflects positive engagement or enjoyment.

Higher values in these metrics generally suggest greater immersion.

TA larger pupil size captured by Tobii Pro Spectrum eye tracker is typically associated with higher levels of arousal, which, in the context of our study, correlates with increased immersion. Elevated skin conductance and heart rate captured by Shimmer 3 GSR Unit are also often associated with higher levels of arousal, which may indicate increased either immersion or nervousness in the task.

To effectively utilize these values in our data analysis, we used the mean of each measurement across the entire duration of each round. By averaging these values, we can provide a clearer comparison of emotion levels both between participants and across different conditions, such as rounds with music and rounds without music.

As by the study procedure described in subsection 4.2 we used a within-subject design. The influence of music on task performance is measured in the same way as immersion. Prior to data analysis, Shapiro-Wilk tests and Q-Q plots will be used to determine whether the different results follow a normal distribution across all participants. If a given immersion or performance result follows a normal distribution, we will use a paired t-test. On the other hand, if a performance and immersion result does not follow normal distribution, we will use a Wilcoxon test.

6. Results

6.1. Questionnaire Results

Prior to data analysis, Shapiro-Wilk tests and Q-Q plots confirmed a normal distribution of the questionnaire results.

For the within-subject design, NASA-TLX workload scores and IEQ immersion scores were compared between the same participants, for rounds 3 and 4. The mean NASA-TLX score when participants experienced *music condition* was slightly higher (M = 979, SD = 367.1) compared to when they experienced *no-music condition* (M = 976, SD = 313.7). Despite this small difference, the paired t-test results ($t(9) = 0.088$, $p = 0.9312$ $d = 0.008$) showed no significant difference in the NASA-TLX score between the two conditions. Conversely, the mean Immersion score when participants experienced *music condition* was larger (M = 50.8, SD = 7.6) compared to participants who experienced *no-music condition* (M = 45.8, SD = 6.8). The paired t-test results ($t(9) = 1.811$, $p = 0.10$ $d = 0.691$) suggested a marginally significant difference between the two conditions, with a large effect size.

6.2. Physiological Data Analysis

Prior to data analysis, Shapiro-Wilk tests and Q-Q plots confirmed a normal distribution of the aggregated data captured by iMotions. We observed no discernible differences in the mean values of emotions between the two conditions. While some emotions exhibited higher mean values under the *music condition*, others showed higher mean values under the *no-music condition*. Utilizing paired t-tests, we found no significant differences between the two conditions for any of the analyzed emotions, additionally, the effect sizes were very small. The most notable result was observed for the “Lip Press” expression. Here, the mean value for the expression when participants were exposed to *music condition* was slightly higher (M = 4.27, SD = 7.46) compared to when they were exposed to *no-music condition* (M = 2.65, SD = 4.61). The paired t-test yielded: $t(9) = 1.72$, $p = 0.12$, Cohen’s effect size $d = 0.08$. Even then, the effect size is minimal. For detailed statistical results of all emotions refer, to Table 2

6.3. Game Score Analysis

Utilizing the game scores, we can determine whether music has a significant impact on video game performance.

Table 2
Statistical Results for Emotion Categories.
A = Mean in the *music condition*, *B* = Standard deviation in the *music condition*, *C* = Mean in the *no-music condition*, *D* = Standard deviation in the *no-music condition*.

Emotion	A	B	C	D	t	p	d
Joy	0.43	0.87	0.23	0.47	0.67	0.52	0.27
Engagement	4.14	7.12	4.38	9.11	-0.23	0.816	0.02
Valence	-2.51	5.07	-1.83	6.01	-0.59	0.56	0.12
Attention	97.42	0.77	97.46	0.99	-0.12	0.91	0.03
Brow Furrow	0.55	1.59	0.11	0.20	0.93	0.37	0.28
Eye Widen	0.38	0.77	0.68	1.97	-0.71	0.49	0.09
Lip Press	4.27	7.46	2.65	4.61	1.72	0.12	0.08
Lip Suck	8.79	14.23	9.79	20.59	-0.30	0.77	0.04
Mouth Open	0.98	1.48	1.13	2.42	-0.2	0.84	0.07
Smile	0.79	1.08	1.44	3.36	-0.68	0.51	0.22
Left Pupil	4.22	0.65	4.34	0.50	-1.37	0.20	0.17
Right Pupil	4.37	0.73	4.52	0.46	-1.20	0.26	0.18
GSR Cond.	4.25	4.58	4.10	4.05	0.59	0.56	0.02

Prior to data analysis, Shapiro-Wilk tests and Q-Q plots confirmed a normal distribution of the aggregated data. A within-subject design was used to analyze gameplay performance between the two conditions: As mentioned, in order to minimize the learning curve effect, we opted to only use rounds 3 and 4. We will now refer to rounds 3 and 4 as such: the rounds where people experienced music as the *music condition* and the rounds where people did not experience music as the *no-music condition.*

- Score: The mean scores in the *music condition* was actually lower (M = 248.5, SD = 54.05) than in the *no-music condition* (M = 256.4, SD = 48.7). Despite this, the paired t-test results ($t(9) = -0.94$, $p = 0.37$, Cohen's effect size $d = 0.20$) showed no significant difference between the scores.
- The number of "Goods": the mean amount of "Goods" was actually higher in *the music condition* (M = 46.8, SD = 19.42) compared to *no-music condition* (M = 39.3, SD = 14.06). The paired t-tests showed a borderline significant difference between the two conditions, and a small to medium effect size ($t(9) = 1.78$, $p = 0.10$, Cohen's effect size $d = 0.45$).
- The number of "Perfects": The mean amount was smaller in the *music condition* (M = 110.9, SD = 24.50) compared to the *no-music condition* (M = 119.1, SD = 21.24). The paired t-test results showed no significant difference ($t(9) = -1.72$, $p = 0.12$, Cohen's effect size $d = 0.35$.
- The number of "Misses": The mean amount was slightly higher in the *music condition* (M = 9.6, SD = 10.00) than in the *no-music condition* (M = 10.01, SD = 8.44). Alongside this, the paired t-test showed no significant difference($t(9) = 0.26$, $p = 0.80$, Cohen's effect size $d = 0.06$).
- The number of "Wrongs": The mean amount was slightly higher in the *music condition* (M = 13.2, SD = 14.80) than in the *no-music condition* (M = 12.1, SD = 12.12). The

paired t-test showed no significant difference($t(9) = 0.34$, $p = 0.74$, Cohen's effect size $d = 0.08$).

7. Discussion

7.1. Result Interpretation

Even though the mean NASA-TLX workload score was slightly higher in *music condition*, the difference in workload scores was statistically insignificant, and the effect size was negligible. Thus, it is reasonable to assume that music has no effect on the perceived workload. On the other hand, the mean values of the immersion scores were notably higher in *music condition*. Despite the non-significant difference, the lower p-value and larger effect size suggest that music might enhance the immersive experience of the game. With that in mind, there is a possibility that a Type II error occurred (i.e a false negative) due to the small sample size, indicating that the study might not have had sufficient power to detect a significant difference.

When analyzing the physiological data, the different emotions that would indicate immersion, did not follow a discernible trend or pattern. Engagement, attention, valence, and smile counts were higher in the *no-music condition*, while the number of joy expressions, lip presses, and brow furrows were higher in the *music condition*. Due to the non-significant differences and small effect sizes, it can be assumed that music either has no significant effect on immersion or that the participants did not express immersion through facial expressions.

Additionally, the higher GSR conductance in the *music condition*, although insignificant, might indicate heightened emotional responses such as fear, excitement, anxiety, or stress. This supports the idea that energetic music, especially unfamiliar music, might increase nervousness. This could either indicate heightened immersion levels or a break in immersion. Based on self-reported immersion scores, GSR conductance could be a supporting factor for music's impact on immersion.

The game scores were higher in *no-music condition*, and while these differences are also non-significant, The fact that 9 out of 10 participants played better when there was no music involved, as well as obtained the smaller amount of "Misses", and "Wrongs" and the larger amount of "Perfects" strongly supports the idea that music, especially unfamiliar music is distracting and can lead to more errors, even if it matches the taks at hand. There is also a possibility that a Type II error occurred in the significance testing, due to the small amount of participants.

7.2. Addressing Reseach Questions

Research question 1: *How does energetic music affect the level of immersion experienced by the player*

- The difference in immersion scores in the self reported immersion questionnaires, as well as the difference in GSR conductance, is a strong indicator that music positively influences the level of immersion experienced by the player.

Research question 2: *How does energetic music during gameplay affect game perception (workload, pace, arousal)?*

- Due to the conflicting results, namely, the NASA-TLX workload score yielding no significant difference, supporting the idea that energetic music has little to no effect on game perception, while the GSR Conductivity indicated higher levels of arousal, which could mean that the participants were more stressed, anxious, or in panic due to the game being perceived as faster, or more difficult, we cannot come to a definite answer. In a subsequent questionnaire, participants were asked if the music changed their perception of the pace or intensity of the game, to which 7 answered that the pace of the game felt faster, 2 answered that the pace of the game felt slower, while 1 answered that it felt the same. While this result does not provide statistical significance, it does provide insight into the relation between energetic music and task perception, but further research is needed.

Research question 3: *How does immersion and game perception affect game performance?*

- While it is clear that the presence of music negatively impacted game performance, the exact reasoning behind this effect can be explored in various ways. One possible interpretation is that participants perceived the game as more difficult or fast-paced when music was present, leading to a higher propensity for mistakes. This notion is slightly supported by the findings related to the second research question. A third interpretation is that participants did indeed experience higher immersion levels, but this immersion was related more to the overall game experience rather than the specific gameplay itself. The presence of music might have led participants to feel that they should enjoy the game fully, rather than focus on performing the task to the best of their abilities. This suggests that music might have influenced their emotional engagement with the game environment rather than their cognitive engagement with the gameplay mechanics. The emotional engagement, driven by energetic music, could have made the participants feel more connected to the game setting, yet distracted them from optimal gameplay performance. Additionally, energetic music is often described as distracting, thus leading to a divided attention scenario. Participants might have found themselves paying more attention to the music, which created an immersive auditory environment, but at the expense of their concentration on the game's mechanics and objectives.

7.3. Comparison with Previous Studies

Our findings closely align with some of the previous studies. Namely, research suggesting that music negatively impacts game performance, and that energetic, unfamiliar music is seen as distracting [1, 3]. Concurrently, our finding align with previous research concluding that there is no significant difference in performance between *no-music condition* and *music condition* [4].

Our findings also align with research suggesting that music helps with the immersive experience [2]. The discrepancy between higher immersion and lower performance observed in this study can be better understood through previous research. Previous studies have established a significant correlation between immersion and performance [14, 15, 16], but this relationship depends on the immersion being directly relevant to the task at hand. In our study, the high levels of immersion were induced by intense, high-tempo music. This form of immersion was maybe not directly related to the gameplay task, leading to the observed lack of improvement in performance [10, 17, 18, 19]. Moreover, the energetic nature of the music likely introduced additional stress and nervousness for participants.

Thus, it appears that immersion only supports task performance when it is closely tied to the task itself. In cases where the immersive element, such as high-tempo music, is not task-related, it may not contribute positively to performance and can even hinder it.

Additionally, task immersion is shown to be most impactful to performance when the task relies heavily on working memory, spatial reasoning, and the ability to mentally manipulate and store information (e.g trying to memorize a pattern, or an image) [14, 15]. Moreover, studies have shown that immersion is least significant when the task performed is highly reliant on motor skills, reaction time and the ability to synchronize physical movements with visual inputs [16], such as the video game in our study.

7.4. Implications of Findings

First, the results suggest that simply increasing immersion through elements like high-tempo music does not inherently enhance performance. Instead, the effectiveness of immersion is highly context-dependent. So, designers of immersive experiences should be mindful of how immersion elements relate to the specific task.

Second, the impact of different types of music on task performance empathizes the need for careful selection of audio stimuli. While calming music has been shown to improve performance, energetic and intense music might increase stress levels and lower efficiency. This information can guide the selection of background music in various settings, including workplaces, educational environments, and therapeutic settings, to optimize user experience and performance.

Regarding stress management, the negative impact of high-tempo music on performance underlines the importance of managing stress in immersive environments. Designers should be aware of how immersive elements, such as music, affect stress levels and overall well-being. providing options for users to select their preferred auditory environment could mitigate potential negative effects.

In practice, understanding that immersion must be relevant to the task can improve the design of virtual environments and other immersive applications. By coordinating immersive elements with the specific goals of the task, practitioners can create more effective and engaging environments that enhance user performance and satisfaction.

7.5. Limitations of Study

There were a few limitations of the study. The most impactful limitation is the small sample size. This likely hindered statistical power, increasing the risk of Type II errors, and incorporated a large amount of participant variability, which could be a confounding variable. This additional variability might have influenced the results, making it difficult to isolate the effects of music. An additional limitation would be the hardware used. During the user study, the Shimmer unit failed to detect a heart rate, resulting in the heart rate values being unusable, which would have otherwise been a useful statistic. The study focuses on a single type of game and one specific type of music (energetic, high-tempo). While this focus allows for controlled testing, it limits the generalizability of the findings to other game genres or music types. Future research could include different genres of music or different types of games (e.g., strategy, puzzle, adventure games).

8. Conclusion

This study aimed to analyze the impact of energetic music on game immersion, game perception and performance. and whether there is a reasonable connection between them. We collected a combination of subjective questionnaires (NASA-TLX, IEQ) and objective measures (eye tracking, skin conductivity, facial expressions) as indicators of immersion. Additionally, the effect of music on game performance was assessed through in-game scores. The statistical analysis found no statistically significant differences in mean values for most emotional and physiological measures between the *music* and *no-music conditions*. Additionally, the analysis found no statistically significant differences in game scores. Despite this lack of statistical significance, several key observations emerged: the mean values of immersion in the immersion questionnaire were higher in the *music condition* than in the *no-music condition*. This suggests that participants felt more immersed when music was present. However, this increased sense of immersion did not translate into improved performance, as the *no-music condition* had higher performance scores on average.

These findings lead to a deeper examination of the relationship between immersion and performance. The observed discrepancy, where increased immersion was associated with lower performance, suggests that the energetic music might have been detrimental to performance. This prompted a hypothesis that immersion, to positively impact performance, needs to be directly related to the task at hand. Previous research supports the notion that immersion can enhance performance when it is closely aligned with task requirements. In contrast, when the immersive element, such as high-tempo music, is not relevant to the task, it may not contribute positively and could even impair performance.

The relationship between immersion, perception and performance needs to be further investigated by exploring how different types of immersion, particularly those tailored to align with the nature of the task, influence performance. Studies could investigate various immersive elements and their direct relevance to specific tasks to better understand how immersion can be optimized for different activities. For instance, examining how task-specific immersion affects performance could reveal more about the conditions under which immersion enhances task outcomes.

Exploring individual differences, such as personal music preferences, baseline stress levels, or familiarity with the task, could offer insights into why some individuals experience varying effects compared to others. Expanding research to include different contexts beyond video games, such as educational environments, workplaces, or therapeutic settings, would help understand how immersive elements like music impact performance in a variety of real-world scenarios.

Acknowledgments

This research was funded by the Slovenian Research Agency, grant number P1-0383, P5-0433, IO-0035, J5-50155 and J7-50096. This work has also been supported by the research program CogniCom (0013103) at the University of Primorska.

References

[1] W. Thompson, E. Schellenberg, A. Letnic, Fast and loud background music disrupts reading comprehension, Psychology of Music 40 (2012) 700–708. doi:10.1177/0305735611400173.

[2] L. Nacke, M. Grimshaw-Aagaard, C. Lindley, More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game, Interacting with Computers 22 (2010) 336–343. doi:10.1016/j.intcom.2010.04.005.

[3] G. Cassidy, R. MacDonald, The effects of music choice on task performance: A study of the impact of self-selected and experimenter-selected music on driving game performance and experience, Musicae Scientiae 13 (2009) 357–386. doi:10.1177/102986490901300207.

[4] H.-M. Lin, S.-H. Kuo, T. Mai, Slower tempo makes worse performance? the effect of musical tempo on cognitive processing speed, Frontiers in Psychology 14 (2023) 998460. doi:10.3389/fpsyg.2023.998460.

[5] S. Dobbs, A. Furnham, A. McClelland, The effect of background music and noise on the cognitive test performance of introverts and extraverts, Applied Cognitive Psychology 25 (2011) 307 – 313. doi:10.1002/acp.1692.

[6] S. Hallam, J. Price, G. Katsarou, The effects of background music on primary school pupils' task performance, Educational Studies - EDUC STUD 28 (2002) 111–122. doi:10.1080/03055690220124551.

[7] C. Tse, Effect of music on cognitive performance in different contexts, To be published (2024). doi:10.58445/rars.1051.

[8] B. Dalton, D. Behm, Effects of noise and music on human and task performance: A systematic review, Occupational Ergonomics 7 (2007) 143–152. doi:10.3233/OER-2007-7301.

[9] M. V. Thoma, R. La Marca, R. Brönnimann, L. Finkel, U. Ehlert, U. M. Nater, The effect of music on the human stress response, PloS one 8 (2013) e70156. doi:10.1371/journal.pone.0070156.

[10] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, A. Walton, Measuring and defining the experience of immersion in games, International journal of human-computer studies 66 (2008) 641–661. doi:10.1016/j.ijhcs.2008.04.004.

[11] M. Bishay, K. Preston, M. Strafuss, G. Page, J. Turcot, M. Mavadati, Affdex 2.0: A real-time facial expression analysis toolkit, in: 2023 IEEE 17th international conference on automatic face and gesture recognition (FG), IEEE, 2023, pp. 1–8. doi:10.1109/FG57933.2023.10042673.

[12] T. Science, Test science, 2019. URL: https://testscience.org/, accessed: 2024-08-15.

[13] J. Rigby, S. Gould, D. Brumby, A. Cox, Development of a questionnaire to measure immersion in video media: The film ieq, 2019. doi:10.1145/3317697.3323361.

[14] J. Gisler, C. Hift, A. Kunz, V. Holzwarth, Designing virtual training environments: Does immersion increase task performance?, in: 2020 International Conference on Cyberworlds (CW), IEEE, 2020, pp. 125–128. doi:10.1109/CW49994.2020.00026.

[15] M. Slater, V. Linakis, M. Usoh, R. Kooper, Immersion, presence, and performance in virtual environments: An experiment with tri-dimensional chess, ACM Virtual Reality Software and Technology (VRST) (1999). doi:10.1145/3304181.3304216.

[16] T. Rose, K. B. Chen, Effect of levels of immersion on performance and presence in virtual occupational tasks, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, volume 62, SAGE Publications Sage CA: Los Angeles, CA, 2018, pp. 2079–2083. doi:10.1177/1541931218621469.

[17] N. Perham, J. Vizard, Can preference for background music mediate the irrelevant sound effect?, Applied Cognitive Psychology 25 (2011) 625 – 631. doi:10.1002/acp.1731.

[18] D. Sanseverino, A. Caputo, C. G. Cortese, C. Ghislieri, "don't stop the music," please: The relationship between music use at work, satisfaction, and performance, Behavioral Sciences 13 (2022) 15. doi:10.3390/bs13010015.

[19] M. C. Muhmenthaler, B. Meier, Different impact of task switching and response-category conflict on subsequent memory, Psychological research 85 (2021) 679–696. doi:0.1007/s00426-019-01274-3.