Bachelor's thesis

Information and Communications Technology

2023

Iulia-Andreea Alexandru

The impact of sound immersion on a user's ability to complete a task in a virtual learning environment

TURKU AMK
TURKU UNIVERSITY OF
APPLIED SCIENCES

Bachelor’s Thesis | Abstract

Turku University of Applied Sciences

Information and Communications Technology

2023 | number of pages

Iulia-Andreea Alexandru

The impact of sound immersion on a user’s ability to complete a task in a virtual learning environment

Immersion, through a rich audiovisual experience, is generally regarded as a strong catalyst for digital learning. As such, researchers often place a strong emphasis on sound when designing digital game-based learning (DGBL) solutions. However, some sounds can disrupt the learning process. This thesis analyzes how sound immersion in a DGBL context can have negative consequences on a user’s ability to complete tasks in a virtual learning environment (VLE). Two groups were subjected to two VLEs, a silent one and an audio-immersive one. The VLE required participants to complete a sorting task that collected in-app quantitative data. After the experiment, participants answered a questionnaire about their experience. The results show that the audio-immersive environment had a negative effect on participant performance in that the number of errors was significantly higher under sound stress ($p = 0.03754$). The completion time, however, was not significantly different. The questionnaire showed that men and women experienced stress differently. These results highlight the importance of careful consideration when selecting sounds for an engaging digital learning experience without inducing stress.

Keywords:

Sound immersion, serious games, digital game-based learning, psychology

Contents

Appendices

Figures

Tables

List of abbreviations (or) symbols

DGBL	Digital Game-Based Learning
HPA	Hypothalamic-pituitary-adrenal
NICU	Neonatal Intensive Care Unit
VLE	Virtual Learning Environment
VR	Virtual Reality

1 Introduction

The impact of video games on human psychology has been vastly studied ever since the popularization of the industry in the early 2000s. Initially created as a form of digital entertainment, use cases have grown into endless possibilities across two decades as digital interactive media expanded into multiple specialized industries. Ever since the beginning, video games have quickly enthralled both older and newer generations, including young children. Through the sustained success of digital gaming and the seemingly endless hours of engagement these games offer, pedagogues across the world turned their interests toward a new instructional method called digital game-based learning (DGBL). DGBL was implemented worldwide, across multiple subjects and skill levels. The goal of DGBL is to engage learners by incorporating educational content or learning principles into video games (Coffey, 2017). It encourages education by creating custom learning experiences where users combine information assimilation with practical assignments (Deubel, 2006), enhancing mental agility and stimulating the mind. DGBL has been found to lead to improved learning outcomes for students compared to traditional teaching methods (Chen & Tu, 2021). By engaging the user in a virtual learning environment (VLE) using DGBL principles, the user’s motivation, efficacy, and performance improves (Coffey, 2017), mainly due to the process of immersion.

Immersion is stated to be the precursor to Csikszentmihalyi’s (1990) concept of flow. Flow refers to “the state in which individuals are so involved in an activity that nothing else seems to matter” (Csikszentmihalyi, 1975, 1990). Csikszentmihalyi (1990) goes on to claim that users can enter a flow state by engaging into activities that have clearly defined goals, encourage progress, and respond with immediate feedback. The activity must not be too challenging in nature, or the user will enter an anxious state, lose confidence and motivation to engage. At the same time, the activity must not be too passive, or the user will become bored and once again lose the motivation to engage. A good balance between player ability and difficulty is required to achieve the perfect state of flow. A digital game that immerses the user and strongly evokes a psychological response similar to a real-world environment possesses high psychological fidelity. Audiovisual fidelity (including sound) is an important component of psychological fidelity that stimulates the senses (Petridis, 2012).

Sound as a form of immersion has been widely studied and the emphasis on sound immersion to engage users for a more enhanced experience (Kenwright, 2020) remains a key focus. Similar research for inducing immersion through sound (Salselas et al. 2020) advocates using sound to facilitate attention and stress-related emotional involvement in interactive environments. Much other current research tends to echo this advice that DGBL designers should incorporate sound as an immersion instrument in all DGBL applications.

Contrary to this, studies (Klatte et al., 2013; Monteiro et al., 2018; Münzel et al., 2018) demonstrated that a rich sound experience could become stressful and even dangerous in occupational and educational scenarios. While some thrive under stress, the majority react adversely to sensory stimulation while attempting to complete tasks. The overall experience is improved by sound immersion, but the stress factor could inhibit the learning process, rendering the user more prone to make errors and lose focus. The majority of research fails to isolate and focus on the impacts of sound immersion as stress and often lacks a nuanced understanding of the user experience. This thesis addresses the issue by specifically focusing on the effects of sound immersion as stress and considering multiple methods of measurement, including both objective and subjective measures as well as self-reported measures of stress. This thesis aims to demonstrate that audio immersion or stimulation (particularly defined as psychological noise) has an impact on a user’s ability to complete a task in a VLE by measuring task performance in two groups of participants. A VLE was developed that measures the time taken and the number of mistakes users make while performing a sorting task with and without exposure to environmental sound.

Chapter 2 of this thesis elaborates on literature and other findings in the field of sound immersion. Chapter 3 presents the thesis’ aims and objectives and is followed by an expanded view on the research protocol and prototype description in Chapter 4. The following Chapter 5 displays the technical development side of this study. Chapter 6 presents the results of this study, which will be further elaborated in the discussion section in chapter 7. Lastly, Chapter 8 focuses on conclusions and recommendations.

2 Literature review

2.1 Sound immersion in DGBL applications

It is important to note that the effectiveness of digital game-based learning can vary depending on a number of factors, including the specific game being used, the learning objectives being addressed, the characteristics of the learners, and the implementation context. While some studies have found that DGBL can be an effective way to facilitate learning and promote student engagement (Van Eck, 2006; Krakoç et al. 2008), others have identified challenges and limitations of using games in education. One challenge that has been identified in the literature is that of immersion. While immersion can be beneficial for learning in some cases, excessive or inappropriate immersion can also have negative consequences. For example, Gagné and Klimmt (2006) conducted a laboratory experiment in which participants played a violent digital game under different levels of immersion. The game was designed to manipulate the level of immersion by varying the level of sound (high versus low) and the level of interactivity (high versus low). Gagné and Klimmt (2006) found that immersion was negatively correlated with learning outcomes, suggesting that excessive immersion in a game may disrupt the learning process and reduce the effectiveness of the game as a learning tool. Specifically, the authors found that participants who played the game under high levels of sound and interactivity scored significantly lower on a post-test measuring knowledge about the game content than participants who played the game under low levels of sound and interactivity.

2.2 Types of digital sound and their psychological effects

Only in recent years researchers have realized the importance of sound and the crucial role it plays in the psychology of digital interactive media and engagement. Kenwright (2020) states that even in a digitally interactive context, sound is the most important human sense after sight. In the same study, Kenwright (2020) places sounds in two different categories: passive and active. Passive sounds refer to the ambient or background and their purpose is to create an atmosphere. The sounds are usually subtle and soft. They work in synergy with the other senses and are capable of generating a spectrum of moods. Active sounds, also referred to as “stimulating sounds” are sudden, powerful, sometimes extreme and their purpose is to engage the user’s conscious awareness. “These sounds are classed as non-linear as they exceed the normal musical range of an instrument (or the vocal cords of living creatures)” (Kenwright, 2020; Cheng, 2014). Active sounds stimulate the senses and may contribute to short-term changes in the users’ thinking patterns and emotions. Both passive and active sounds are required, in different volumes, to create a coherent, realistic environment. Kenwright (2020) concludes that sound remains an under-appreciated tool in interactive environments and advocates for specialists to focus and include more sound in future applications.

Generally, the mutual conclusion of recent is that (a) sound remains underutilized in interactive environments, (b) sound immersion has many effects on learning and (c) strong support in favor of using a wide range of sounds in interactive environments. However, it is important to note that the specific effects of sound on emotional response and task performance may vary depending on the individual, the specific context in which the sound is used, and the type of sound being used. Psychologically, the users’ brain can regard sound as pleasant, stressful, or neutral due to cognitive appraisal. Cognitive appraisal “refers to the personal interpretation of a situation that ultimately influences the extent to which the situation is perceived as stressful” (Campbell et al. 2013). In interactive digital games, sound highly supports cognitive appraisal through realism, strengthening engagement and amplifying immersion, creating a positive user experience (Ekman, 2008).

2.3 Fidelity

In a 2008 study, psychological fidelity is described as “the degree of perceived realism, including factors such as emotions, beliefs, and the self-awareness of participants...” (Dieckmann et al. 2008). Concerning digital games, Petridis (2012) further develops this concept into two categories: audiovisual fidelity and functional fidelity. Audiovisual fidelity focuses on rendering, animation, and sound, while functional fidelity refers to technical aspects such as scripting, supported AI techniques and game physics. A third category is added and described (Alexander et al. 2005) as psychological fidelity in video games, addressing emotional content, time pressure and noise. Noise, relating to sound immersion in video games, is “a gameplay disruption, either from within the game itself of the player’s immediate surroundings (including actual noise, interruption, or distraction) that causes the player to disengage from the game” (Ravyse et al. 2020). Psychologically, noise can activate different stress reactions that impact focus, memory, and performance (Westman & Walters, 1981).

2.4 Noise

Noise has been vastly studied outside of the digital gaming context for many years for its effects on health and mental well-being. Monteiro et al. (2018) discovered that high auditory stimulation in an occupational environment increases the levels of stress and discomfort. The participants had an increase in errors and reaction time when subjected to noise pressure. Both steady-state background noise (passive) as well as impulsive noise (active) contribute to increased circulating cortisol, causing an acute physiological stress reaction and decreasing performance in mental work (Radun et al. 2022). Not only does noise affect users on a mental level but also on a physiological one, as chronic mental stress (generated by noise pollution) is associated with cardiovascular risk factors (Münzel et al. 2018).

Pertaining to digital video games, in a quasi-experiment (Hébert et al. 2005) it has been discovered that even though the sound isn’t regarded as noise, it could still trigger a stress reaction in users that have been exposed to built-in game music. “These findings suggest for the first time that the auditory input contributes significantly to the stress response found during video game playing” (Hébert et al. 2005). Due to everything stated above, in a DGBL context, it is safe to say that sound immersion is expected to impact learning and task performance.

Sound as a form immersion has been proven to affect users psychologically and cognitively by eliciting a wide array of feelings and states of mind. (Daniels, 1995) Sensory game elements, such as sound, are important factors that encourages players’ involvement in a DGBL, and thus, plays a crucial role in increasing and retaining learners’ engagement. (Byun & Loh, 2015) Byun and Loh (2015) further claim, alongside Bishop et al. (2008) that specialists tend to ignore the audio component when designing virtual learning environments. There is only limited information available on the use of sounds in DGBL applications. (Bishop et al. 2008) Hence, most studies in the DGBL field fail to focus on or even address the audio element as a booster for engagement (Byun & Loh, 2015).

3 Aims & objectives

The main purpose of this thesis was to investigate whether audio stimulation has an effect on a user's ability to complete a task within a virtual learning environment by measuring task performance. Task performance was measured through a randomized control-group experiment in which both quantitative and qualitative data were collected. Both groups completed a simple task in a VLE and then qualitatively reported on the experience in a post-test questionnaire. The independent variable in this experiment is represented through the noise of a crying baby present in Group B's version of the VLE. Quantitative data was collected by examining in-game user performance through variables.

The research's objective was to compare the results between the groups, note differences between the experiences of all participants and identify key factors in the results. This study also takes into consideration the possibility of certain results being influenced by categorical variables such as gender, parental status or headphone use. Data was grouped and analyzed based on all the variables stated above and the following hypothesis were tested:

H_0: Auditory stimulation has no impact on task performance, Group B exhibiting similar behavior as Group A.

H_A: Auditory stimulation has an impact on task performance, Group B having different results and exhibiting a different behavior than Group A.

4 Methodology

4.1 The prototype

This study uses a neonatal intensive care unit (NICU) virtual reality (VR) environment created by Turku University of Applied Sciences (TUAS) for the IC x-RP project collaboration between Finland, Tunisia, and Germany. For the purposes of the IC x-RP project, the NICU hospital scene was highly decorated. Decorated environments that contain numerous items are often distracting, hindering one's focus on the main assignment. Therefore, to best isolate the sound variable for the purpose of this study (i.e., to minimize the effect of confounding factors), most ornamental purposeless visual components were removed from the scene. Furthermore, due to performance and test distribution concerns, the VR implementation was removed so that the learning environment could be used on a regular desktop PC.

The virtual environment used for this study is set in an undecorated NICU recovery room adjacent to a medication preparation room. The player spawns in the recovery room with three incubators (Figure 1).

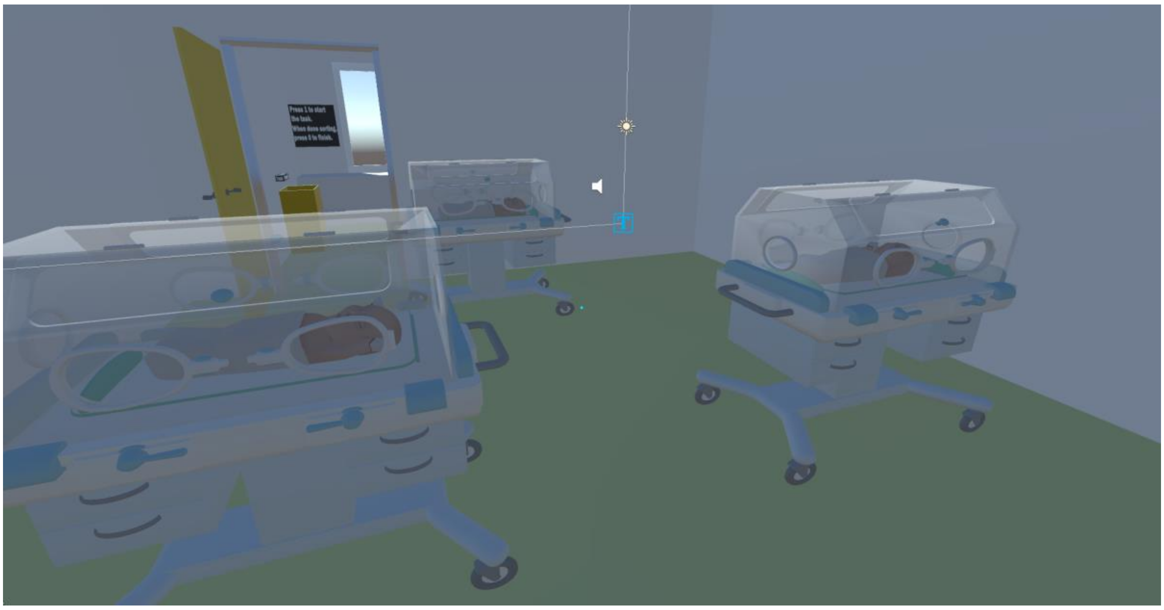

Figure 1. The NICU recovery room.

Each of the incubators contains a baby (or patient). The player has access to the virtual babies for the duration of the experiment. With the help of the W, A, S and D keys, the player then proceeds to enter the medication preparation room that has a work counter, three labeled drawers, a trash bin and instructions panels indicating how to continue (Figure 2).

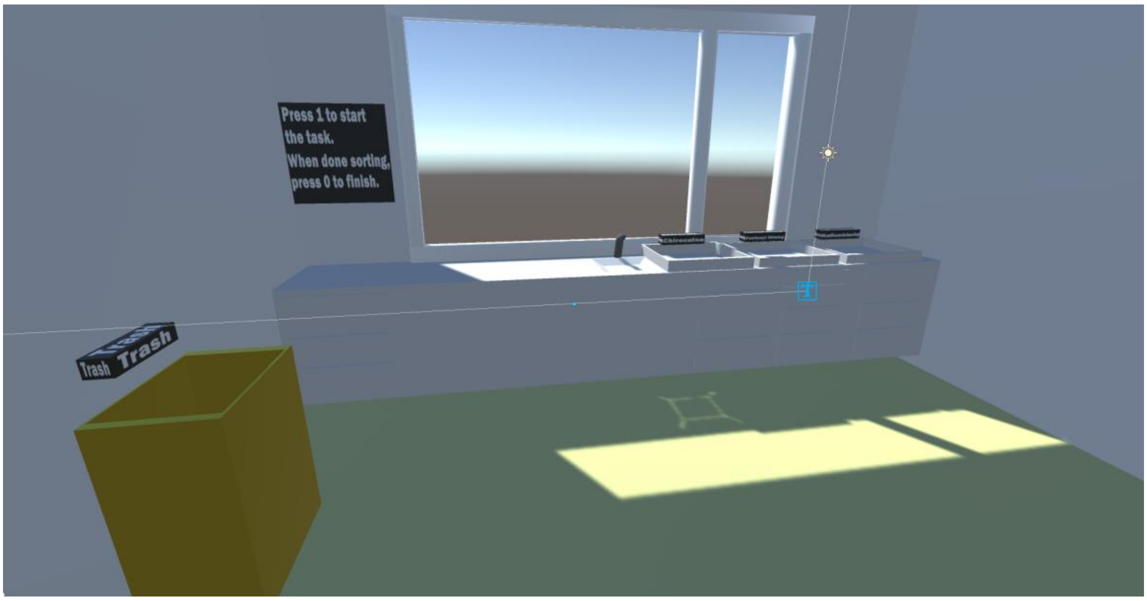

Figure 2. The NICU medication preparation room.

When pressing the alphanumeric key 1, the task starts by randomly placing 10 medication packages from 5 different packaging prefabs (two medication packages of each type) onto the work counter next to the drawers. Each participant was given a set of basic instructions that contained the key binding information and task description. The participants were left to explore, start, and finish on their own time with little to no interruption. To remove any possibility of added pressure or inducing stress, the participants were timed discreetly. The errors the groups do are also accounted for, but the data is hidden until the end of the task. The task consisted of sorting 10 randomly generated medication boxes of 5 different types into the correct drawers. Two out of the five medication types were control objects and had to be discarded into a bin. The control objects aimed to verify that the user’s attempt at completing the task was genuine was used as an emphasized measure of maintaining focus. By aiming the light blue cursor at the medication and pressing left-click on the mouse, the objects can be picked up and dropped. The player must place the genuine medication boxes into their respective drawers and place the control packages into the bin. If the player makes an error, the app will add a point to the mistake counter. Mistakes are counted when a user places: (a) any of the three genuine medication types in the bin; (b) any of the two control medication types into a drawer; and (c) any of the genuine medication types into an incorrect drawer. Once the player is done, pressing the alphanumerical key 0 ends the task and a panel appears on the screen with information on the number of mistakes made as well as the how long (measured in seconds) it took to complete the task from start to end. This is the quantitative data for this study. The groups received two different versions of the application, one version contains an audio stressor and the other does not. The stressor consists of a newborn cry that starts at the same time as the medication packages spawn and continues for 140 seconds.

4.2 Experiment protocol

Two groups of 20 people each, of any age or gender, were tested by completing an assignment in custom virtual learning environment through a control-group experiment. To combat bias and avoid familiarity of the scenario, the selection of the participants was conditioned. No group member had any background in any medical or healthcare related field. The groups received two different versions of the application. After progressing through the application, all participants were asked to complete a post-test questionnaire. The study collected both quantitative and qualitative data. The people were split in the two groups as following:

a. Group A is the group that experienced the application without a sound stressor.

b. Group B is the group that experienced the application with a sound stressor.

The quantitative results were compared using tables containing central tendency measurements, analyzed using a t-test and a Mann-Whitney U test and displayed as box plots using five-number summaries. In terms of qualitative results, each statement of the Likert scale was analyzed for recurring patterns and significant clusters of responses.

4.3 The post-test questionnaire

After the participants completed the assignment within the application, they were asked to fill in a post-test questionnaire (Appendix 1). Section A of the questionnaire asked the participants to enter their application metrics (number of mistakes and task completion time). As this experiment aimed to trigger acute psychological stress response in Group B's participants, sexual identity was requested to investigate if women respond differently than men to the sound of a crying baby. Section B of the questionnaire intended to verify the legitimacy of the user's answer and identify the presence of possible confounding variables. While question one was there to determine the level of immersive stimulus, question two verifies whether the participant had heard anything at all—those who played the stressor free version should answer "no." Questions three and four checked whether participants were conditioned (or desensitized) to hearing a cry of a newborn or a child by asking the participant's parental status or recent contact with very young children.

In Section C a Likert questionnaire was used in which respondents rated their agreement or disagreement with statements on a scale: strongly agree, agree, neither agree nor disagree, disagree, or strongly disagree. These responses were assigned numerical values from one to five, allowing for quantitative analysis of the data. This section served to collect qualitative data and mainly focuses on how the participants felt about the experiment. The first two statements check if the user understood the assignment. Statements three through seven validated the mood that the user was in before and during the assignment. Lastly, statement eight verified if the answers submitted reflect the participant's mood throughout the experiment, or if their answer stemmed from misunderstanding or confusion. At the same time, it investigated if confusion is correlated with self-reported distress.

5 Technical development

This study utilizes a VR environment of a NICU developed by Turku University of Applied Sciences. The original environment was adapted to fit the study by reorganizing assets and removing all the redundant components such as plugins, scripts, and decorative models. The following 3D assets, previously not part of the scene, were made in Blender: genuine and control medication boxes, the newborn baby, the drawers with each matching label and the trash bin.

Measuring the task performance is achieved through a user interface (UI) canvas containing a timer that counts in seconds and a mistake counter. This canvas is initially hidden but appears when the experiment is over. The timer counts from 0 as the user presses 1 to start the main task and it stops when the user presses 0 to end the task. The mistake counter operates through a mistake manager script that instantiates the main function of the script *AddMistake();* each time an incorrect collision occurs. The function then adds “1” to the counter.

The application’s main feature is its collision system. Since proper collision settings are necessary for the experiment to function correctly, each of the drawers and the trash bin has its own custom collider and collision scripts. Initially, the player moving around the environment could trigger the *AddMistake();* function and display an invalid number of mistakes. As a result, the drawers are programmed to ignore collision with the player and only account for correct or incorrect collision with the medicine boxes. All the medicine boxes are prefabs tagged accordingly for the collision detection script to determine if the collision is correct or incorrect. A correct collision will destroy the medicine box game object and an incorrect one will add a mistake point to the mistake counter. Additionally, an invisible game object with its own collider has been included in the environment. The purpose of this game object is to block the physics of the medicine box spawner, as it has been noticed that at times the spawner could shoot one or many of the medicine boxes in the labelled drawers, adding invalid mistakes to the counter. The existence of these invisible game objects limits the area in which the medicine boxes can spawn (Figure 3).

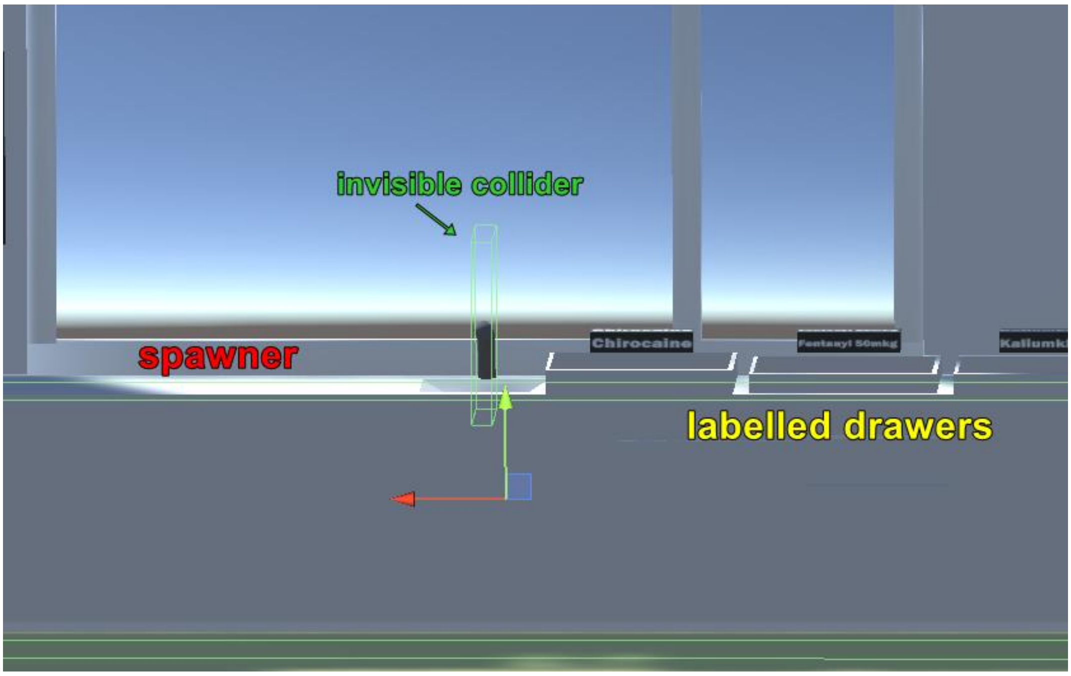

Figure 3. Representation of the invisible collider dividing the spawner area and the labelled drawers.

For the user to be able to pick up and drop the objects, a script called *PlayerPickUpDrop* was created. After establishing a grab point and a pickup distance, a ray can be cast towards the grab point using Raycast physics. The grab point is represented by the light blue cursor. By pressing the left mouse button when the ray hits a game object that is tagged as a medicine package, the object transforms its original position to the grab point. When pressing the left mouse button again, the grab point releases the object, and the object gains a new position. Any of the medicine boxes can be picked up and dropped as desired using this mechanic.

In Group B’s version of the application, sound immersion was added. Two empty game objects have been created: *PlaySound* and *NewbornCry*. An audio source component was added to *NewbornCry*, alongside an audio clip, disabling *Play On Awake* and adjusting the volume. Then, a script was added to the *PlaySound* game object that triggers the *NewbornCry* sound. The audio clip starts playing as soon as the player starts the task by pressing 1 and plays for 140 seconds. The audio clip is represented by a crying newborn and has been carefully selected from various free-to-use online sound libraries in a way that is not too impersonal nor too disturbing.

6 Results

The purpose of this study was to examine whether audio stimuli have an impact on users' ability to complete a task in a VLE. Both quantitative and qualitative data were collected and analyzed from both groups. Quantitative data was collected during the experiment within the application and qualitative data was collected post-experiment through the post-test questionnaire specifically designed for this study.

6.1 Quantitative data from the VLE: Time-to-completion and number of errors.

Group A consist of participants that completed the experiment in a silent environment, while Group B consists of the participants that completed the experiment in an audio-immersive environment. Tendency measurement tables were created to represent the quantitative data based on the measurements of time-to-completion and number of errors variables. The raw data collected was constructed into tables where sample size and the means of the time-to-completion and number of errors were calculated. The data was compared between (a) Group A and Group B; (b) men and women in Group A; and (c) men and women in Group B. Increases or decreases in tendencies were also calculated. For the statistical evaluation of the time-to-completion scores between Group A and Group B, a single tailed t-test was performed. As the error margins were getting high, the statistical evaluation of the number of errors scores between the groups was done using a Mann-Whitney U test for a more accurate analysis. Box plots using five-number summaries were created for the purposes of data visualization of the differences between Group A and Group B.

6.1.1 Group A and Group B

A t-test was conducted to compare the means of time-to-completion samples of Group A and Group B. The p- value for this test was 0.3175, which is statistically insignificant. The mean difference between Group A and Group B was -17.60, with a 95% confidence interval ranging from -53.57 to 18.37. This confidence interval gives an estimate of the range within which the true mean difference is likely to fall, based on the sample data. The t-value for this test was 1.0280, and the degrees of freedom (df) were 18. The standard error of the difference was 17.120.

A Mann-Whitney U test was conducted to determine if there is a statistically significant difference between the means of number of errors. The results of the test show the u-value of 26, the z-score of -1.77643 and the corresponding p-value of 0.03754, which is statistically significant.

Based on the tendency table assessment, the time-to-completion mean of Group B is 22.56% greater than the mean of Group A. At the same time, participants of Group B were 56,52% more likely to make errors (Table 1).

Table 1. Time-to-completion and number of errors means of Group A and Group B.

Group	**Sample size**	**Time-to-completion mean**	**Number of errors mean**
A	10	78	2.3
B	10	95.6	3.6

Data visualization is displayed as the following five-number summary box plots. Outliers were detected and excluded by calculating Tukey's fences for a more accurate representation (Figure 4).

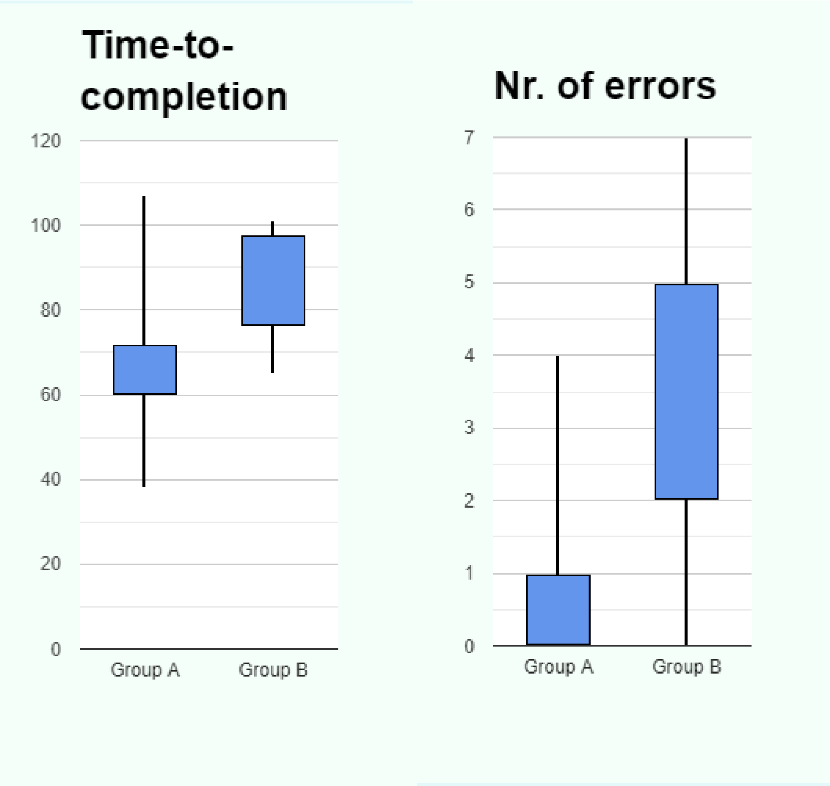

Figure 4. Time-to-completion and nr. of errors box plots comparison between Group A and Group B.

6.1.2 Men and women in Group A

On average, women in Group A took 81,46% more time to complete the assignment than men in Group B. Also, women were 4 times more likely to make errors, the increase being 442,17% (Table 2).

Table 2. Time-to-completion and number of errors means of men and women in Group A.

Group A	Sample size	Time-to-completion mean	Number of errors mean
Men	6	58.83	0.83
Women	4	106.75	4.5

6.1.3 Men and women in Group B

In Group B, men’s time-to-completion score was 11,02% greater than of women. Men were also 23,08% more likely to make errors (Table 3).

Table 3. Time-to-completion and number of errors means of men and women in Group B.

Group B	Sample size	Time-to-completion mean	Number of errors mean
Men	7	98.9	3.9
Women	3	88	3

6.2 Headphone users

The following section showcases the time-to-completion and number of errors means by headphone usage in Group B, being the group that experienced the audio-immersive environment (Table 4). Participants who have used headphones during the assignment are 26% faster than participants who did not. At the same time, they were 57,14% less likely to make errors. Three out of five headphone users were women.

Table 4. Self-reported headphone usage along time-to-completion and number of errors means.

Headphone usage	Sample size	Time-to-completion mean	Number of errors mean
Yes	5	84.6	2.8
No	5	106.6	4.4

6.3 Parental status or contact with children under three in the last five days

No participant belonging to Group B has reported any recent contact with children under 3. Also, all of the participants in Group B are non-parents.

6.4 Qualitative data from the post-VLE questionnaire

Participants in both groups have rated eight different statements on a scale from one to five. A low rating of one or two corresponds with disapproval of the statement, while high rating of four or five is associated with agreement and relatability. A rating of three remains neutral. Key differences were identified between the ratings of the silent environment testers (Group A) and audio-immersive environment testers (Group B) as well as the ratings between the genders of each group.

The most significant difference between Group A and Group B is related to the statement “I got irritated while completing the task”. All participants in Group A strongly disagreed with the statement, while in Group B, the majority (70%) agreed. Similarly, Group B also reported more distress during the experiment than Group A through the statement “I was not distressed while completing the task”. All participants belonging in Group A rated this statement high. Group B’s responses were mixed: 30% giving it a high rating, 20% giving it a low rating and 50% remaining neutral. Another significant result has to do with the ability to maintain focus. Group B rated the statement “I managed to maintain focus while completing the task” lower than Group A. The perception of simplicity of the assignment has also been affected, Group B rating the assignment as being more difficult than Group A through “I found the main task to be easy enough to complete” statement. Most contrasting responses regarding the statements above came from the men in Group B.

Men belonging to Group B responded diversely from both men in Group A and women in Group B. Compared to men in Group A, men in Group B reported more distress, irritation and difficulty, less focus, all in accordance with what was described in the previous paragraph. Additionally, they also reported more feelings of confusion, rating the statement “I felt like the task was confusing” way higher than men in Group A. In contrast with women in Group B, who have experienced the identical environment containing the same audio stimulus, men in Group B reported more distress, difficulty, and confusion regarding the task. Even though women in Group B claimed to be subjected to more stress and irritation during the assignment than people in Group A, they did not report loss of focus nor increased difficulty of the task. At the same time, women in both groups have reported extra pre-test stress by rating the statement “I was mainly relaxed before I started this task” lower than men in both groups.

Furthermore, more similarities between the groups and genders were observed. All participants agreed with the “I understood what I needed to do” statement. No participant wanted to give up on the task. The levels of irritation between the genders of Group B were documented as similar. All statements have also been rated similar between the genders of Group A, except the one regarding pre-test stress. A summary table (Table 5) of the questionnaire averages regarding the genders between both groups is shown below.

Table 5. Summary of the responses across genders in both groups.

		Group A		Group B	
Nr.	Statement	Men	Women	Men	Women
1	"I found the main task to be easy enough to complete"	4.8	4.8	3.9	4.7
2	"I understood what I needed to do"	4.8	4.8	4.4	4.7
3	"I managed to maintain focus while completing the task"	5	4.8	3.4	4.3
4	"I was mainly relaxed before I started this task"	4.3	3.8	4.7	4
5	"I was not distressed while completing the task"	4.8	5	3	3.7
6	"I got irritated while completing the task"	1	1	3.7	3.3
7	"I wanted to give up on the task"	1	1	1.3	1
8	"I felt like the task was confusing"	1.2	1.8	2.1	1

Headphone users of Group B rated the simplicity of the task higher than non-headphone users. Similarly, they also reported losing less focus but being more distressed. The levels of irritation between headphone users and non-headphone users across group B were similar.

No significant data has been observed regarding parental status or recent contact with very young children between the groups, as there were no parents participating in Group B and only two parents participating in Group A.

7 Discussion

7.1 Stressful audio immersion and its impact on learning

Both the quantitative and qualitative data collected displayed a trend in which Group B’s task performance and experience is affected by the sound stimulus present in the environment. The results of the performed t-test on the time-to-completion samples indicate that there is not a statistically significant difference between the means of Group A and Group B, with a p value of 0.3175. These results suggest that any observed difference between the groups could be due to chance. However, participants of Group B took longer to complete the assignment and made more mistakes on average than participants of Group A. At the same time, Group B reported more feelings of distress, irritation, and lack of focus than Group A. This confirms other research (Klatte et al., 2013; Monteiro et al., 2018; Münzel et al., 2018) which state that noise has a negative impact on task performance and ability to concentrate by increasing stress levels and is further supported by the Mann-Whitney U test that was performed on the nr. of errors means. The results of this test show that there is a statistically significant difference between the means, as indicated by the u-value of 26 (u<27), which is significant at a p-value of less than 0.05, and the z-score of -1.77643 and the corresponding p-value of 0.03754. These findings suggest that the observed difference between the means of the two groups is unlikely to have occurred by chance. This tendency is also evident in the graphical representation of the data (Figure 4).

Elevated levels of stress do affect learning on a greater scale, irrespective of whether the reaction is acute or long term. Heightened stress during an educative process mostly consolidates memory of the stressor itself, rather than the information to be acquired (Gewirtz & Radke, 2010). Gewirtz & Radke (2010) further claim that a prolonged stress reaction also impacts memory formation altogether, but only for a limited amount of time until the body decompresses the physiological effects of stress. It is plausible to assume that immersive sound (defined as psychological noise) has a negative effect on learning altogether, including in a DGBL context, as 100% of the participants in Group B were troubled by the sound stressor being present in the environment, regardless of headphone use or gender.

7.2 Women and men react differently to stress

Even though the levels of irritation reported were similar between the genders in Group B, women performed better than men under acute stress by being faster at completing the assignment and committing less errors, while men dealt with higher levels of distress, confusion, and loss of focus. This difference can be explained by examining the basic physiological differences between men and women such as sex hormones and the interaction these hormones have with stress. The physiological mechanism controlling and regulating the body's reaction to stress is the interaction between the hypothalamus, pituitary gland, and adrenal gland, also known as the hypothalamic-pituitary-adrenal (HPA) axis. According to Gillies & McArthur (2010) the HPA axis is heavily influenced by sex hormones, particularly estrogen which helps in regulating neural stress responses (Handa et al. 2012). At the same time, Reschke-Hernández's study (2016) suggests that men showed a greater increase in cortisol than women subjected to identical acute stress-inducing experiments. Cortisol is the main body's stress hormone influenced by the functions of the HPA axis that allows the body to stay on high alert (Thau et al. 2022). "During a stressful event, an increase in cortisol can provide the energy required to deal with prolonged or extreme challenge" (American Psychological Association, 2018). The sound stressor chosen for this experiment (a newborn crying) was assured to trigger a reaction and present a challenge to all participants.

An infant's main method of eliciting parental care is through crying (Konner, 2010). "The adult responsiveness to these signals is necessary to optimize the chances of survival for the individual, and for the species" (Piallini et al. 2015). This responsiveness is strongly associated with the HPA axis-related stress reaction. “Adults, despite of their parental status or sex, are genetically and evolutionary programed to respond to human infants…” (Piallini et al. 2015).

In a 2013 study, Duthie and Reynolds claim that the HPA axis undergoes dramatic changes during pregnancy and postpartum, mainly due to significant hormonal changes. As a result, the vulnerability and responsiveness to infants is heightened in women (Duthie & Reynolds, 2013). No participant in Group B reported being a parent, consequently it is unclear if additional factors influenced women’s responsiveness in the experiment. Besides pregnancy and postpartum, women also go through hormonal changes depending on their menstrual cycle phase. “The subjective and neuroendocrine stress responses significantly vary across the menstrual cycle” (Duchense & Pruessner, 2013). Moreover, the hormonal fluctuation during the menstrual cycle is also influenced by certain medicine such as oral contraceptives. It is unknown if any of the women were administered any hormonal medication.

In the sample size, women in Group B also outperformed women in Group A, who were not subjected to any audio stimulation. Research on the topic of risky decision making under stressful conditions suggests that on average, people with elevated cortisol levels in their blood (subjected to stress) make safer decisions (Dreyer et al, 2022). Dreyer et al. (2022) further claims that “low-to-moderate elevations of stress are associated with riskier decision making whereas higher elevations are associated with safer decisions” (Dreyer et al, 2022). The biological levels of stress have not been measured so it remains unclear if women in Group B’s stress levels peaked over a threshold that triggered safer decision making, which influences the outcome of the experiment. It is also unclear why the men have not exhibited the same behavior and why the results are so contrasting.

7.3 Headphone users and possible deeper immersion

In Group B, headphone users reported similar levels of stress compared to non-headphone users but performed better, by taking less time to complete the task and making less mistakes. It is a widely researched fact that headphones enhance the player's experience in a video game, intensifying immersion (Iverson, 2018). Sound as immersion engages the user into the digital activity (Kenwright, 2020) and encourages learning. It is possible that headphones users achieved better results due to the fact that headphones contributed to their immersive experience and motivated them to complete the assignment. However, it is to be considered that three out of five participants that used headphones were women, and as stated above, women generally performed better than the men in this group. Therefore, a correlation between headphone users' excellence and overall women's performance in Group B could be assumed.

7.4 Limitations

As with most of the research, the design of the current study is subject to limitations. The sample size in this study is small. A larger sample size might clarify or even modify the data presented in the results. Additionally, the qualitative and quantitative information was collected directly from the participants through a questionnaire. Therefore, self-reported data cannot be independently verified and needed to be taken at face value. Physiological stress has not been physically measured and the actual levels of circulating cortisol in participants is unknown. At the same time, due to the correlational nature of the study, causality cannot be inferred.

Regarding the survey, a deficiency in the existing methods has been identified. Supplementary questions that clear some of the confounding variables (i.e., use of hormonal medication) are missing. As these variables extend way beyond this study's scope, additional research is needed to further examine their effects and significance.

8 Conclusions and recommendations

The purpose of this thesis was to find out whether sound immersion has any effect on a user's ability to complete a task in a VLE. Task performance has been measured through in-game variables and a post-experiment questionnaire about each participant's impression about the task and their performance. The data was analyzed and conclusions were drawn. Auditory stimulation in the form of stress impacts task performance within a VLE negatively, with the possibility of hindering the learning process. Sudden (active) sounds or sounds defined as noise experienced as immersion overall decreases task performance and focus while increasing irritability. Even though the t-test identified no statistical significance between the time-to-completion scores of the groups, the Mann Whitney U test displayed the contrary about the means of the number of errors. The qualitative results of the post-test questionnaire also pointed towards the same conclusion. This study has identified that task performance is also influenced by headphone usage as well as gender, as men and women’s task performance and stress reports differ from one another.

It is imperative to determine which types of sound should be utilized in sound design to create an immersive experience (encouraging learning) without triggering psychological and physiological stress reactions, inducing unpleasant feelings, and hampering education. To ensure the appropriateness of sounds in a DGBL, user testing with a diverse group of learners should be conducted to identify any potential negative reactions. At the same time, it is essential to consider the context and the main objective of the application. This ensures that the chosen sounds align with the theme and tone of the game and support the learning objectives of the material. Furthermore, it is also important to identify the target demographic for future specialized DGBL applications considering that people react to certain auditory experiences differently based on a variety of factors. Additional research and innovation are needed to create more efficient DGBL systems as well as explore neuroscientific perspectives as to how these systems may affect people of different groups.

References

American Psychological Association. (2018) “Stress effects on the body”. Available at: https://www.apa.org/topics/stress/body [Accessed 11.12.2022].

Alexander, A. L., Brunyé, T., Sidman, J., & Weil, S. A. (2005). “From gaming to training: A review of studies on fidelity, immersion, presence, and buy-in and their effects on transfer in pc-based simulations and games”. DARWARS Training Impact Group, 5, pp. 1–14.

Bishop, M. J., Amankwatia, T. B., & Cates, W. M. (2008). ”Sound's use in instructional software to enhance learning: A theory-to-practice content analysis.” Educational Technology Research and Development, 56, pp. 467-486.

Byun, J. & Loh, C. S. (2015). ”Audio engagement: Effects of game sound on learner engagement in digital game-based learning environments”. Computers in Human Behavior, 46, pp. 129-138.

Campbell, T. S., Johnson J. A. & Zernicke K. A. (2013). “Cognitive Appraisal”. In: Gellman, M.D., Turner, J.R. (eds) Encyclopedia of Behavioral Medicine. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-1005-9_1115

Chen, C.C. & Tu, H.Y. (2021). “The Effect of Digital Game-Based Learning on Learning Motivation and Performance Under Social Cognitive Theory and Entrepreneurial Thinking”. Front. Psychol. 12:750711. doi: 10.3389/fpsyg.2021.750711

Cheng, W. (2014) “Sound Play: Video Games and the Musical Imagination”. New York, Oxford University Press.

Coffey, H. (2017). “Digital game-based learning”. Available at https://www.medicina.univr.it/documenti/Avviso/all/all036209.pdf [Accessed: 10.11.2022].

Csikszentmihalyi, M. (1975). “Beyond Boredom and Anxiety: The Experience of Work and Play in Games”. San Francisco, Jossey Bass.

Csikszentmihalyi, M. (1990). “Flow: The Psychology of Optimal Experience.” New York, Harper and Row.

Daniels, L. (1995). “Audio Vision: Audio-Visual Interaction in Desktop Multimedia” Annual Conference of the International Visual Literacy Association, Tempe, Arizona. Available at: https://files.eric.ed.gov/fulltext/ED380063.pdf [Accessed: 12.11.2022)

Deubel, P. (2006). “Game on!”. T.H.E. Journal (Technological Horizons in Education), 33(6), pp. 30-35.

Dieckmann, P. & Rall, M. (2008). “Designing a Scenario as a Simulated Clinical Experience. The TuPASS Scenario Script”. Clinical Simulation, pp. 541-550.

Dreyer, A. J., Stephen, D., Human, R., Swanepoel, T.L., Adams, L., O’Neill, A., Jacobs, W.J. & Thomas, K.G.F. (2022) “Risky decision making under stressful conditions: men and women with smaller cortisol elevations make riskier social and economic decisions”. Frontiers in Psychology, 13, 810031. https://doi.org/10.3389/fpsyg.2022.810031

Duchesne, A., & Pruessner, J. C. (2013). “Association between subjective and cortisol stress response depends on the menstrual cycle phase”. Psychoneuroendocrinology, 38(12), pp. 3155–3159. https://doi.org/10.1016/j.psyneuen.2013.08.009

Duthie, L., & Reynolds, R. M. (2013). “Changes in the maternal hypothalamic-pituitary-adrenal axis in pregnancy and postpartum: influences on maternal and fetal outcomes”. Neuroendocrinology, 98(2), pp. 106–115. https://doi.org/10.1159/000354702

Ekman, I. (2008) “Psychologically motivated techniques for emotional sound in computer games.”, Proceedings of the Audio Mostly Conference AM ’08, pp. 20–26.

Gagné, M. & Klimmt, C. (2006). “The effects of immersive computer games on aggression and prosocial behavior: A two-year longitudinal study”. Aggressive Behavior, 32(3), pp. 282-293.

Gewirtz, J.C., & Radke, A. (2010). “Effects of stress on learning and memory”. Encyclopedia of Behavioral Neuroscience, 1, pp. 461-468.

Gillies, G. E., & McArthur, S. (2010). “Estrogen actions in the brain and the basis for differential action in men and women: a case for sex-specific medicines”. Pharmacological Reviews, 62(2), pp. 155–198. https://doi.org/10.1124/pr.109.002071

Handa, R. J., Mani, S. K., & Uht, R. M. (2012). “Estrogen receptors and the regulation of neural stress responses”. Neuroendocrinology, 96(2), pp. 111–118. https://doi.org/10.1159/000338397

Hébert, S., Béland, R., Dionne-Fournelle, O., Crête, M., & Lupien, S. J. (2005). “Physiological stress response to video-game playing: the contribution of built-in music.” Life Sciences, 76(20), pp. 2371–2380.

Iverson, S. (2018). “Exploring the concepts of immersion and why headphones enhance the player experience in 21st century online competitive games”. University of Huddersfield. Available at: https://eprints.hud.ac.uk/id/eprint/34839/1/Iveson%20THESIS.pdf [Accessed 12.12.2022].

Kenwright, B. (2020). “There’s more to sound than meets the ear: sound in interactive environments”. IEEE Computer Graphics and Applications, 40 (4), pp. 62–70.

Klatte, M., Bergström, K. & Lachmann, T. (2013). “Does noise affect learning? A short review on noise effects on cognitive performance in children”. Front. Psychol. 4:578. doi: 10.3389/fpsyg.2013.00578.

Konner, M. (2010). “The evolution of childhood: Relationships, emotion, mind”. Belknap Press/Harvard University Press.

Krakoç, B., Eryılmaz, K., Turan Özpolat, E. & Yildirim, I. (2022). ”The Effect of Game-Based Learning on Student Achievement: A Meta-Analysis Study.” Technology, Knowledge and Learning, 27, pp. 207–222.

Monteiro, R., Tomé, D., Neves, P., Silva, D. & Rodrigues, M. A. (2018). “The interactive effect of occupational noise on attention and short-term memory: A pilot study.” Noise Health. 2018 Sep-Oct, 20(96) pp. 190-198. doi: 10.4103/nah.NAH_3_18. PMID: 30516172; PMCID: PMC6301086.

Münzel, T., Sørensen, M., Schmidt, F., Schmidt, E., Steven, S., Kröller-Schön, S. & Daiber, A. (2018). “The Adverse Effects of Environmental Noise Exposure on Oxidative Stress and Cardiovascular Risk.” Antioxid Redox Signal. 2018 Mar

20, 28(9), pp. 873-908. doi: 10.1089/ars.2017.7118. PMID: 29350061; PMCID: PMC5898791.

Piallini, G., De Palo, F., & Simonelli, A (2015). “Parental brain: cerebral areas activated by infant cries and faces. A comparison between different populations of parents and not”. Front. Psychol. 6:1625. doi: 10.3389/fpsyg.2015.01625

Petridis, P., Dunwell, I., Panzoli, D., Arnab, S., Protopsaltis, A., Hendrix, M., & de Freitas, S. (2012). “Game engines selection framework for high-fidelity serious applications”. International Journal of Interactive Worlds, 2012, pp. 1–19.

Radun, J., Maula, H., Rajala, V., Scheinin, M. & Hongisto, V. (2022). ”Acute stress effects of impulsive noise during mental work” Journal of Environmental Psychology, Volume 81, 101819, ISSN 0272-4944, https://doi.org/10.1016/j.jenvp.2022.101819.

Ravyse W., Blignaut, S. & Botha-Ravyse, C. (2020). “Codebook Co-development to Understand Fidelity and Initiate Artificial Intelligence in Serious Games”. International Journal of Game-Based Learning, 10(1), pp. 37-53.

Reschke-Hernández, A. E., Okerstrom, K. L., Bowles Edwards, A., & Tranel, D. (2017). “Sex and stress: Men and women show different cortisol responses to psychological stress induced by the Trier social stress test and the Iowa singing social stress test”. Journal of neuroscience research, 95(1-2), pp. 106–114. https://doi.org/10.1002/jnr.23851

Salselas, I., Penha, R. & Bernardes, G. (2020). “Sound design inducing attention in the context of audiovisual immersive environments”. Personal and Ubiquitous Computing, 25, pp. 737-748.

Thau, L., Gandhi, J. & Sharma, S. (2022). “Physiology, Cortisol”. National Library of Medicine. StatPearls. Available at: https://www.ncbi.nlm.nih.gov/books/NBK538239/ [Accessed 11.12.2022]

Van Eck, R. (2006). “Digital game-based learning: It's not just the digital natives who are restless.” EDUCAUSE Review, 41(2), pp. 16-30.

Westman, J. C. & Walters, J. R. (1981). “Noise and stress: a comprehensive approach.” Environmental Health Perspectives, 41, pp. 291–309.

The post-test questionnaire

A. Participant input

0. Name:
1. Gender: M/F/Non-binary
2. Time:
3. Nr. of Mistakes:

B. Yes and no questions

1. Did you use headphones?
2. Did you hear a newborn cry while completing the task?
3. Are you a parent?
4. Have you interacted with children younger than 3 in the last 5 days?

C. Likert Scale

Please rate each of the statements given below on a scale of 1 to 5, where:

1: Strongly disagree

2: Disagree

3: Neither agree nor disagree

4: Agree

5: Strongly Agree

1. I found the main task to be easy enough to complete.
2. I understood what I needed to do.
3. I managed to maintain focus while completing the task.
4. I was mainly relaxed while completing the task.
5. I was not distressed while completing the task.
6. I got irritated while completing the task.
7. I wanted to give up on the task.
8. I felt like the task was confusing.