DEGREE PROJECT

Impact on player action based on the correlation between environmental game sounds and arousal level

Eddie Lundin

LULEÅ UNIVERSITY OF TECHNOLOGY

[This page intentionally left blank]

Abstract

This study investigates the relationship between environmental game sounds and arousal level, looking to see if/how it affects player action. Previous research has shown that music evoking different levels of arousal affect player lap time performance of a racing game. For this study a computer game was created containing two levels, one with high arousal environmental sounds, and another with low arousal environmental sounds. The two levels were in different environments but had the same task, which was to place specific objects in their corresponding box, with the same color. The amount of time it took for subjects to finish the game under each sonic condition was the data of greatest interest. The results showed no significant difference between subjects elapsed times. However, a significant difference could be seen of how subjects perceived the energy level between the two levels in consideration to the sounding environment. Having this information when designing environmental sounds for a game could be something to think about and use when sound designers make decisions about player perceptibility and where to direct the player mindset to.

Table of Contents

1. Introduction

In a video game, there is sound that functions as direct reflections of a visually obvious happening or object, for instance the sound of a sword hitting an object, a weapon firing, or an explosion. These types of sounds are in the center of attention and in most cases have an immediate and direct impact on user experience, behavior, and action. Sounds placed in the periphery of attention, like ambient or environmental sounds, are less noticeable to the player. However, it has been shown that music in games, which is one of those elements in the periphery of attention, have some effect on player action. Music may influence player decision, behavior and performance depending on qualities like tempo and how the music interacts with the game state. However, little research has been conducted on how environmental sounds impact user action in video games. To further explore sound placed in the periphery of a game, this study will focus on environmental sounds and see how it compares to music and its effects on player action.

2. Background

To start, it is important to consider how any sound can impact a listener's behavior. Serafin, Franinovic´, Hermann, Lemaitre, Rinott, and Rocchesso (2011) explains the concept of Sonic Interaction Design (SID) and how sound can "convey information, meaning, aesthetic and emotional qualities in interactive contexts" (Serafin et al., 2011, p. 87). They explore different ways sound can help people in everyday life situations and affect behavior. To develop the principles of SID, the authors consider interactions with sonic elements that are generally not directly reflected upon, for example the sound of a door closing and pouring water into a glass. The sonic feedback of these actions either confirms or direct you to completion, usually without paying attention to it. Similar to pouring water into a glass, they state that sound can decrease the learning curve when manipulating different interfaces, giving sonic feedback of an action and thereby learning the user how to control the object (Serafin et al., 2011).

The quality of the sound can affect users' perception of satisfaction and annoyance. Users prefer causal (natural) over arbitrary sounds because they easier relate to it, and which decreases misunderstandings, which may make sonic interactions more satisfactory. However, sonic interaction may help the user but not necessarily satisfy them (Serafin et al., 2011).

All this gives evidence to sounds' possible impact on user action. But, what is more interesting is that SID as a field of research explores interactive sounds that is not directly reflected upon since they are part of everyday life happenings (e.g., pouring water into a glass). We also interact with media and play games everyday. These experiences similarly involve sounds that are not objects of attention too.

2.1 Impact of sound in video games

Before considering a sounds potential impact on player action, the function of the sound in the game must be considered. Ekman (2005) has constructed a framework for sound in video games. The basis of this framework is the differences between diegetic and non-diegetic sound. When a sound is diegetic, it's part of the game world, as if, if it were real it would be possible for the characters to hear it in the fictive world (e.g., birds singing). Non-diegetic sound is instead not part of the game world and can only be heard by the actual player (e.g., music being played between cutscenes and music score in general). Because of the possible use of dynamic audio in video games (compared to a more linear form in movies) Collins (2008) explains how diegetic sound can be separated further into dynamic and non-dynamic sound. Dynamic sound can then also be further separated into interactive and adaptive activity, meaning the player's interactions cause it to change or other changes in the game state drive change. Similar to Serafin et al. (2011) presenting how sound is able to help and direct in everyday interactions, Collins (2008) explains how music in the form of interactive and non-diegetic is able to inform the player of an upcoming event (e.g., enemies approaching). It is considered interactive because the music event responds to whether the player is approaching or not. Collins (2008) also presents how music, when synced with enemies' movements, helps the player know when to make a move. These types of sonic directions, shows "that game sound can be a significant element of gameplay in at least some cases, and that it can function in many ways" (Collins, 2008, p. 128). Similarly to the last example, she explains how the use of acousmatic sound (having no origin visually) can have an effect on player decision making, for example in stealth games where the player may be notified, for example by a musical earcon cue, when an enemy is close. This has an affect of the player deciding to get out of there (Collins, 2008).

These impacts on players, Collins (2008) states, compared to audio turned off, shortens the learning curve, and may create a less frustrating gaming experience. Collins (2008) suggests that having a more linear sound design (e.g., nondynamic nondiegetic music) may be seen by the user as useless and substitutable and thus turn off the sound. Rogers and Weber (2019) found that a strong preference when playing "infrastructure-building" games (IBG) was using their personal music, with one of the reasons being that personal music is more enjoyable. Collins (2008) argues that to get players not turn off the music, it needs to be more integral in the game experience, like the examples mentioned above.

2.2 Music and effect on action

North and Hargreaves (1999) explore video games and the effects of musical preference on performance in a concurrent task. They explain humans´ have a limited capacity for cognitive processing, meaning that when exposed to a difficult task, it decreases the availability of cognitive processing for additional tasks. Similarly, Serafin et al. (2011) expresses the importance for sonic interaction to not exert cognitive overload. Therefore it is more effective to use sounds that are obvious and causal rather than arbitrary. North and Hargreaves (1999) conducted a study by having groups of participants playing a racing game. Subjects finished 5 laps. The different experimental conditions had music that evoked different levels of arousal and tasks of varying difficulty. Because of the risk of participants not paying attention to the music, North and Hargreaves (1999), recorded voices reading out two-digit numbers randomly with 10 second intervals during the test, and participants were asked to repeat it back. They measured task performance and music preference. The hypothesis was that music with high arousal and a concurrent task would result in worse performance than the opposite because of cognitive overload (North and Hargreaves, 1999).

The study showed that lap times were slower when exposed to high arousal than low arousal music. These results show that music and a concurrent task competes over cognitive space and therefore also affects the performance (North and Hargreaves, 1999). More importantly for the aims of the proposed research, it shows the possible impact on arousal (tempo and sound level) of music in consideration to user action.

One of North and Hargreaves (1999) suggestions about further research was to see if complexity or familiarity of music could have an influence on performance. This was among other things investigated by Cassidy and Macdonald (2010) where they did a similar study to North and Hargreaves (1999) but also explored the effects of listening to personal selected music and among other things, participants' performance on accuracy while playing. Contrary to North and Hargreaves' (1999) study, Cassidy and Macdonald (2010) found that participants made faster lap times listening to high arousal than low arousal music, and fastest when listening to self-selected music. Compared to North and Hargreaves (1999), this study separated arousal and tempo which made it possible to analyze them individually. The experimenter-selected music was chosen based on a pre-study using the Circumplex model of emotion (in which emotions can be defined in terms of valence and arousal) and independent participants in order to determine perceived arousal potential. For self- selected music, there was no constraints to what music the participants were to choose, but most of it could be labeled as pop music and varied in tempo 70-150 bpm. (Cassidy and Macdonald, 2010).

Cassidy and Macdonald (2010) found that participants were faster when exposed to the faster music, and slower when exposed to the slower music. This was the case regardless of being exposed to high or low arousal music. They also found that performance was more inaccurate when exposed to fast tempo in high arousal music. However, no effect of tempo was found in low arousal music. As mentioned earlier, self-selected music resulted in best performance on lap time, but they also found this to result in highest accuracy.

The contradictory results suggest the more research is needed. Moreover, the methods for conducting these tests need further development. North and Hargreaves (1999) had numbers randomly read out every 10 seconds in gameplay that participants was asked to repeat, reducing the risk of participants not paying attention to the music. Cassidy and Macdonald (2010) chose not to use this method, and thus risk this happening. The problematic of using this method is that it can assumably reduce the ecological validity of the study. On the other hand, they based this method on previous studies, among them one by Wolfe (1983, via North and Hargreaves, 1999), where a similar experiment was made but found no effect on music, and thus argued that participants avoided attending to the music on a concurrent task. Similar to the debate on whether or not it's necessary for experimenters to include a way assuring that participants pay attention to the music while playing, an understanding of how it translates to environmental sounds is necessary considering it usually being placed in the periphery (like music).

2.3 Retail stores and effect of music

Interestingly, the results of Cassidy and Macdonald (2010) seems to add evidence to some extent research on the impact of music in retail stores. Milliman (1982, via Knoferle et al., 2011) found that slower music in a supermarket resulted in it taking longer for customers getting in and out the store, thus leading to greater sales. He also found that patrons were spending more time and purchasing more alcohol in a restaurant when exposed to slower music (Milliman, 1986, via Knoferle et al., 2011)

2.4 Impact of environmental sounds

The proposed research considers environmental sounds, not music. Dall' Avanzi and Yee-King (2019) explored environment soundscapes in video games to determine, "How much time or recourses should be used to replicate an element that is stochastic and unpredictable in nature, in order to convey a satisfactory experience?" (Dall' Avanzi & Yee-King, 2019, p. 1). The purpose of their study was to determine how different levels of detail affect the player's perception of immersion. There is disagreement about this in the prior literature (via Dall' Avanzi & Yee-king, 2019). Some authors question the importance of a realistic and detailed recreation of the environment, while other lean to a more perceptual realism. To gain better understanding, Dall' Avanzi and Yee-King (2019) conducted a study by having two groups of participants play two different versions (Version A & Version B) of the same game, where the only difference was the environment soundscapes. Version A aimed to create the most detailed and realistic environment as possible, and argued that this "follows an approach more indicative of industry practices" (Dall' Avanzi & Yee-King, 2019, p. 7). Version B was less detailed and used recorded loops of elements from Version A. The results showed that neither version was preferred by their subjects. If no significant difference in preference exists, Dall' Avanzi and Yee-King (2019) believe sound designers could dedicate more time on assets like sounds for user interactions (Dall' Avanzi & Yee-king, 2019).

It can be suggested that the study of Dall' Avanzi & Yee-king (2019) gives an example on affecting the players perception of reality by attempting to make environmental sounds feel realistic without actually sounding realistic. However, it also suggests and predicts environmental sounds being difficult to utilize (like music) when conducting experiments. Even though this information is highly relevant, environmental sounds and how it impacts user action has not yet been explored.

Much research exists on peoples' preferences on environmental soundscapes in non virtual worlds (e.g., Wang and Zhao, 2019; Miller, 2013). However, there is little research on how environmental sounds affect user action. It's been shown that music can impact user action, therefore the same tactics could be employed in the creation of environmental sound. It may be inferred that music consists of five different attributes, being harmony, melody, rhythm, tempo, and mode. In the case of environmental ambiences in games, harmony may be likened to the usage of drones as a mood provoking tonal effect in soundscapes. Further, melody may be likened to melodic sounding animals, a good example is bird songs, for its prominent qualities and varying tonality (see, Ratcliffe et al., 2013). Rhythm may as well be likened to mechanical sounds, thus being stable and carrying repetitive qualities. The tempo of environmental sounds in games may affect how long the time between events are perceived. Modes might have an effect in ambiences in the same way that modes have in music, mainly on emotions. Based on these assumptions, it's suggested that, through inspiration from previous research on how music impact user action, investigate how it's linked to environmental sounds.

2.5 Aims and purpose

Considering studies about tempo and arousal manipulation of music and the effect on player action, this thesis will investigate similar principles but with focus on environmental sounds instead of music. The research question for this thesis is: How does arousal in environmental game sounds impact player action? Tempo is also a factor when considering manipulation of environmental sounds but will be imbedded in arousal as a correlation can be seen between them, separating them (which have been done in research on music) may also be difficult as environmental sounds do not have as much distinct variables (like music) that may be manipulated. The findings of this study will potentially make sound designers for video games think of ambiences as a more integral element, and as a chance to impact player action through a peripheral medium, like music. At the very least, this will expand sound designers' knowledge of sound placed in the periphery of attention.

3. Method

To answer the research question, an experiment was conducted in form of a video game. Subjects were asked to play the game, and at the same time they were exposed to either high or low arousal environment sounds. The purpose of the experiment was to see if and how subjects behave differently depending on the level of arousal in the environmental sounds they are exposed to. Quantitative and qualitative data was collected during the experiment, the amount of time it took for subjects to finish the game under each sonic condition was the data of greatest interest.

3.1 Experimental design

To be able to see the effects of being exposed to different arousal levels within subjects, the experiment was made so that subjects played the game two times, once exposed to a high arousal condition (HA), and once exposed to a low arousal condition (LA). Subjects were not informed that the time it took to complete the game task was being recorded. This was to make it possible for subjects to play as they felt like in the moment (e.g., relax and taking it easy while playing), and thus see if the level of arousal was part of the gameplay behavior. Similarly, it was decided to make the gameplay fairly easy so that all subjects regardless of skill or game experience could finish the task without facing challenges that could take extra time and minimize any potential impact of the sound design. To minimize the risk of becoming familiar with the game task and thus performing significantly faster the second time, and being confused by a change in the ambient soundtrack without any visual change in the environment, the same game task was set in two different environmental settings within the game. One setting was a farm and the other was a traffic atmosphere. These two locations made for two game levels.

3.2 Pre-study and sound design

Before implementing the environmental sounds in the game level, 4 different linear soundtracks/ambiences were made, resembling the conditions later assembled in the experiment. Subjects (who would not be included in the main study) were asked to rate the soundtracks on a scale of 1-5 for different attributes; level of energy (1 = Calm, 5 = Energetic), complexity (1 = Simple, 5 = Complex), and sounds being perceived as soft/harsh (1 = Smooth & Soft, 5 = Rough & Sharp). The pre-study only consisted of listening to the linear soundtracks, it was therefore no gameplay involved. Subjects first compared and rated the attributes for HA and LA ambiences of farm environment, and then did the same procedure for traffic environment. The soundtracks were about 1 minute and 20 seconds long, and subjects were able to switch between the two soundtracks as much as they wanted. The four soundtracks were evaluated in total: Farm HA, Farm LA, Traffic HA, and Traffic LA. When designing the soundtracks, factors that were most considered were the time between sonic events, and how much energy there is between 2-4 kHz. According to McDermott (2012), frequencies in the range of 2-4 kHz is contributing to a feeling of annoyingness, since people perceive that range 30 dB stronger than other frequency ranges. The farm environment included different types of birds, wind-swooshes, and crickets. In the high arousal condition of this environment, it also contained elements like the sound of a tractor on idle, a rooster, and bees buzzing. The traffic environment included various ambiences of different cities, church bells, car on idle and distant crowd talking. They were composed using sound effects from Soundly Pro (2022) and edited and processed in Logic Pro X 10.4.8 (2022). The time between sounding events was made shorter in the high arousal versions than the low arousal versions, which made the ambiences denser, and more complex. In the high arousal versions, the number of elements that sounded closer to the player was more than in the low arousal versions. Generally, there were not many elements that sounded close to the player in the low arousal versions, instead prerecorded ambiences were often more used, and they were suitable for the corresponding environment. It was also made sure that the energy between 2-4 kHz was boosted on most elements in the high arousal versions, while set on a lower level in the low arousal versions.

The pre-study showed that subjects somewhat perceived the HA soundtracks as HA and the LA soundtracks as LA. However, the perceived differences between the HA and LA conditions were small. Therefore, a second pre-study was conducted after some corrections were made on the soundtracks. This was made by adding additional elements in the high arousal versions, like more farm animals and machinery, a dog barking, and distant traffic noise. The same procedure and group of subjects for the first pre-study was used in the second one. The results then showed a clearer difference between the different soundtracks, and these soundtracks became templates for the non-linear soundtracks implemented in the game level for the main experiment. The soundtracks can be listened at the following url:
https://drive.google.com/drive/folders/1ZzwXqhMwvitZhXgAE92PHNKwN4VtyTxi?usp=sharing

3.3 The game

The game was built and sound was implemented using Unreal Engine 4.26 (2022). It was made using the first person shooter (FPS) template, and where the standard controls WASD and the mouse was used to move and look around. By holding down the left mouse button, the subjects were able to pick up objects. The subjects were not able to jump or sprint, which minimized the different ways of playing the game, making the data easier to analyze. A demo environment made by Assetsville (2022) was used for the game world, including various props and assets that could be set out in the world.

There were two levels made in the game using the same template being placed in different environments, one being played at a farm, and one in a traffic atmosphere. For each level there was a number of objects set out within a limited area, half of the objects were green, and the other half were red. There were also two boxes in both green and red placed in the area. The task was to place the objects in their corresponding box, with the same color. Subjects had to collect a total of four objects for each box to complete the task. A number of physical obstacles were placed in the levels. This was to make the game a bit more challenging for the player. The obstacles were placed at different locations in the two levels. This was made so that subjects wouldn´t get accustomed to the levels. Before playing the two test conditions, subjects played a trial level where no ambiences were heard, but only the sound of the player character´s footsteps and object collisions. This trial was included to make subjects comfortable with the controllers before playing the actual test games.

Figure 1. Overview of farm environment

Figure 2. Overview of traffic environment.

3.4 Audio for game

The audio assets used in the second pre-study were used for the main study game. The audio was implemented in such a way that the frequency and distribution of sonic elements in the pre-test was maintained in the audio implementation. The prerecorded ambiences were set to play in 2D, while the foley sound elements were scattered around the play area as positional sounds. It was also made sure that the positional sounds could be heard from any location in the play area. This made the sound experience similar for all subjects no matter where they moved. The balance of the different elements was set to resemble the pre-study soundtracks as much as possible.

3.5 Subjects

A total of 24 subjects was used for the experiment and were between the age 19-34 years old. All the subjects on some occasion in their life had played a computer game where the controls of WASD and mouse was used. This was a criterion for participating in the experiment. All subjects were students at Luleå University of Technology, studying either music, audio engineering or journalism. *Figure 3* shows a pie chart of how often subjects play computer games. The question that was asked was "How often do you play computer games?" The results show that most subjects play at least either once a month or once a week. It can also be seen that none of the subjects play less than once a year.

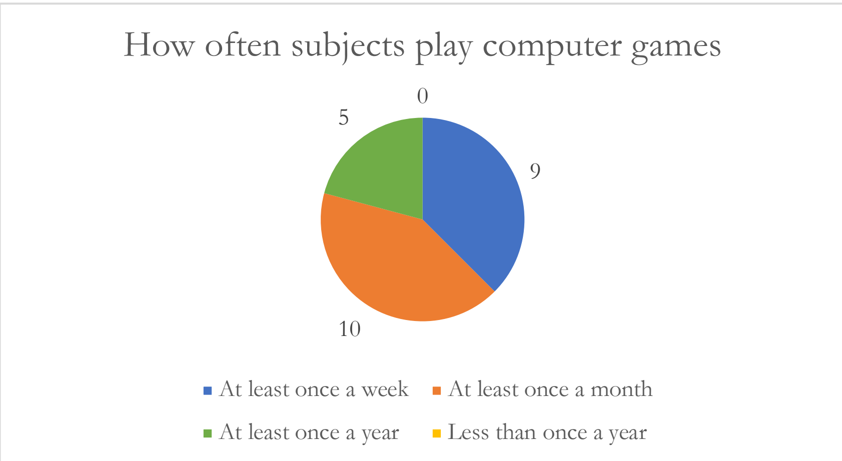

Figure 3. A pie chart of how often subjects play computer games.

3.6 Listening Test

All subjects played one HA condition and one LA condition. Subjects played one farm condition and one traffic condition. The experimenter planned the order of the combinations in advance so that there would be an even distribution of subjects playing each of the following combinations:

HA farm/LA traffic; LA farm/HA traffic; HA traffic/LA farm and LA traffic/HA farm. The switch in environment and high-low arousal conditions was to prevent order effects. Experiment took place in a room located at Luleå University of Technology in Piteå. Subjects played on a computer (Windows (Microsoft, 2020) operating system) listening with a pair of headphones (beyerdynamic DT-990 Pro). They were not able to change the volume for the game. This was to make sure that all subjects could hear the environmental sounds clearly and had equal conditions. A set of instructions was presented to the subjects before the trial level started. The instructions can be seen in the Appendix of this paper. After playing the first level, subjects answered on some demographic questions, and then preceded to the second level. When both levels were completed, subjects answered on a qualitative questionnaire. The questionnaires were answered on a second computer, next to the gaming computer. This made it possible for the experimenter to simultaneously prepare for the next level.

3.7 Collected Data

The quantitative data collected was the time it took for subjects to finish each level. Again, it was not mentioned to the subjects that time was being recorded. The demographic questionnaire was given to subjects between the first and second test. It was meant to be a kind of break that could help to neutralize any change in arousal level caused by playing the first test. The questionnaire took approximately 2 minutes to complete and consisted of questions about their age, what they studied, and how often they play video games.

The qualitative questionnaire was taken after they had completed both levels. Subjects rated the following on a scale of 1-5: "How would you compare the energy levels of the two sound environments?" (1 = Level 1 was more energetic, 5 = Level 2 was more energetic); and "How would you compare the level of complexity of the two sound environments?" (1 = Level 1 was more complex, 5 = Level 2 was more complex). Before subjects answered the questions about how they experienced the sound environments, it was asked whether they paid any attention to it while playing. If they answered yes, they were told to continue with remaining questions, otherwise they would continue to the last question, which was whether they had any other thoughts about the game.

4. Results & Analysis

The high arousal farm environment is referred as "FHA", and the low arousal to "FLA". The high arousal traffic environment is referred as "THA", and the low arousal to "TLA". An independent t-test (two-tailed) was conducted to find statistically significant differences of time elapsed between levels of the same environment. On top of this, a paired sample t-test was also conducted to find any significant differences of time elapsed between the main conditions (HA/LA). A null hypothesis was conducted for the analysis, and as excepted, no significant difference of time elapsed could be found. Furthermore, a chi square test was conducted to see if there were any statistically significant differences in perceived energy level. All statistical calculations used a significance level of α-value 0.05.

4.1 Time results for level FHA & FLA

Table 1 shows time elapsed of subjects experiencing the FHA and FLA conditions. The mean, median, and standard deviation (STDEV) value can also be seen for each condition. The number column presented in the table is not representative of the actual order of the subjects. 12 subjects played each condition. Subjects who played the FHA condition did not play the FLA condition. *Figure 4* shows a box plot of time elapsed where the mean value, upper and lower quartile, and the outlier can be seen.

Table 1. Time (in seconds) elapsed on condition FHA and FLA.

Number	Time Elapsed FHA (in seconds)	Time Elapsed FLA (in seconds)
1	111	88
2	85	105
3	121	84
4	108	157
5	83	85
6	248	131
7	68	96
8	102	93
9	132	155
10	91	86
11	114	105
12	97	147
Mean	113.333	111
Median	105	100.5
STDEV	45.973	28.489

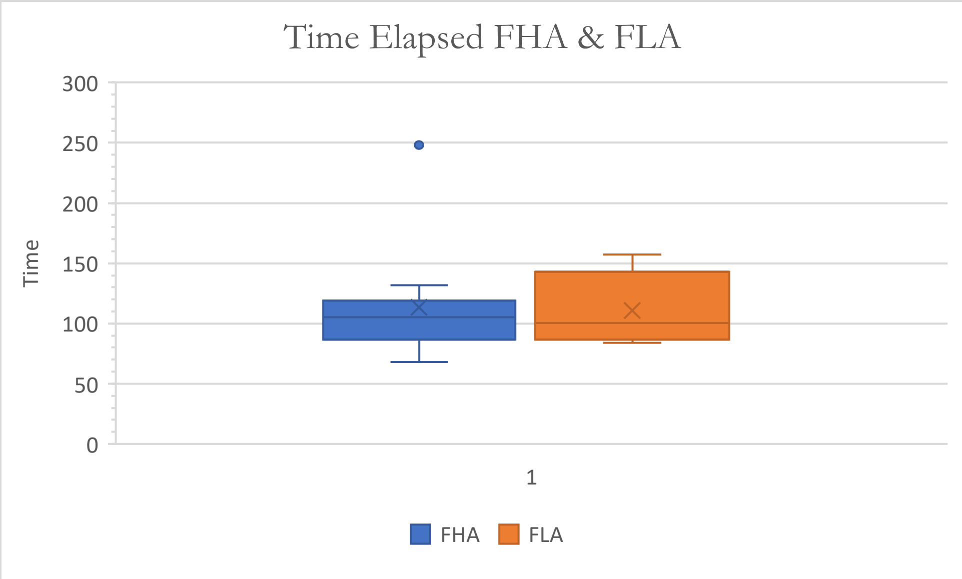

Figure 4. Playing time (in seconds) on condition FHA and FLA.

4.1.1 Analysis of results for level FHA & FLA

An outlier with a time of 248 seconds is seen on the FHA condition in *Figure 4*. The outlier was included in analysis. An independent t-test was executed showing a t-value of 0.128 (two-tailed). Compared to the critical t-value of 2.074, it can be stated that no statistically significant difference is found between the two conditions. Therefore, the null hypothesis cannot be rejected.

4.2 Time results for level THA & TLA

Table 2 shows time elapsed for subjects experiencing the THA and TLA conditions. The mean, median, and standard deviation value can also be seen for each condition. The number column presented in the table is not representative of the actual order of the subjects. 12 subjects played each condition. Subjects who played the THA condition did not play the TLA condition. *Figure 5* shows a box plot of time elapsed where the mean value, upper and lower quartile, and the outlier can be seen.

Table 2. Time (in seconds) elapsed on condition THA and TLA.

Number	Time Elapsed THA (in seconds)	Time Elapsed TLA (in seconds)
1	114	88
2	120	94
3	167	141
4	103	102
5	89	76
6	138	90
7	81	98
8	81	87
9	70	93
10	121	95
11	82	83
12	111	160
Mean	106.417	100.583
Median	107	93.5
STDEV	28.079	24.63

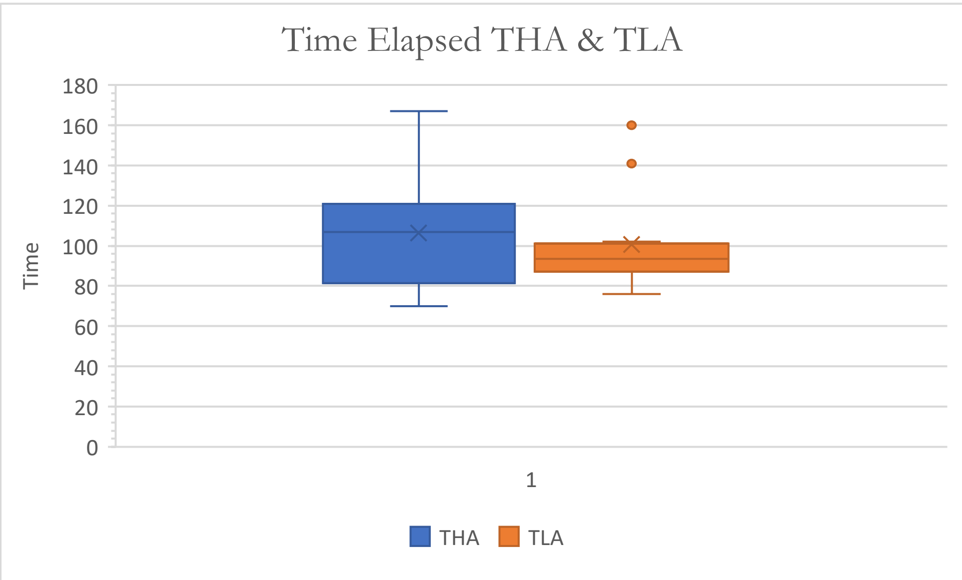

Figure 5. Playing times (in seconds) on condition FHA and FLA.

4.2.1 Analysis of results for level THA & TLA

Two outliers with a time of 160 and 141 seconds is seen on the TLA condition in *Figure 5*. The outliers were included in analysis. An independent t-test was executed showing a t-value of 0.464 (two-tailed). Compared to the critical t-value of 2.074, it can be stated that no statistically significant difference is found between the two conditions. Therefore, the null hypothesis cannot be rejected.

4.3 Time results for condition HA & LA

Table 3 shows time elapsed for subjects experiencing a HA and LA condition. Results marked in blue shows subjects that have played a HA condition first, and results marked in orange shows subjects that have played a LA condition first. The table also shows the mean, median, and standard deviation value for each condition group. *Figure 6* shows a box plot of time elapsed where the mean value, upper and lower quartile, and the outlier can be seen.

Table 3. Time (in seconds) elapsed on condition group HA and LA.

Test Subjects no	**Time Elapsed HA (in seconds)**	**Time Elapsed LA (in seconds)**
1	111	98
2	85	87
3	121	93
4	108	95
5	83	83
6	248	160
7	114	96
8	120	93
9	167	155
10	103	86
11	89	105
12	138	147
13	68	88
14	102	94
15	132	141
16	91	102
17	114	76
18	97	90
19	81	88
20	81	105
21	70	84
22	121	157
23	82	85
24	111	131
Mean	109.875	105.792
Median	105.5	94.5
STDEV	37.422	26.582

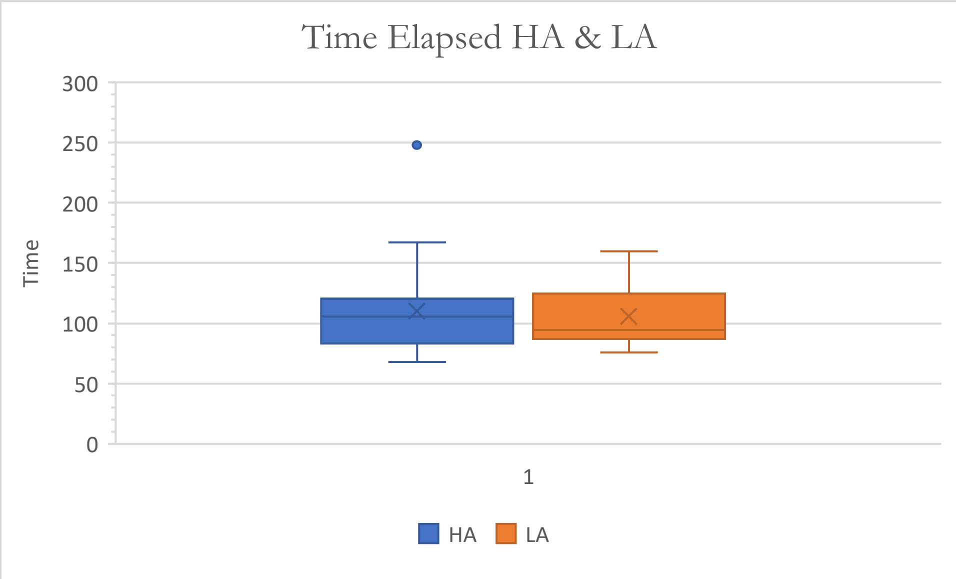

Figure 6. Playing times (in seconds) on condition group HA and LA.

4.3.1 Analysis of results for condition group HA & LA

One outlier with a time of 248 seconds is seen on the HA condition in *Figure 6*. The outlier was included in analysis. A paired sample t-test was executed showing a t-value of 0.784 (two-tailed). Compared to the critical t-value of 2.069, it can be stated that no statistically significant difference is found between the two conditions. Therefore, the null hypothesis cannot be rejected.

4.4 Perceived energy level results

Figure 7 shows a pie chart of how subjects perceived the energy level between the main conditions HA and LA. The question that was asked was "How would you compare the energy levels of the two sound environments?".

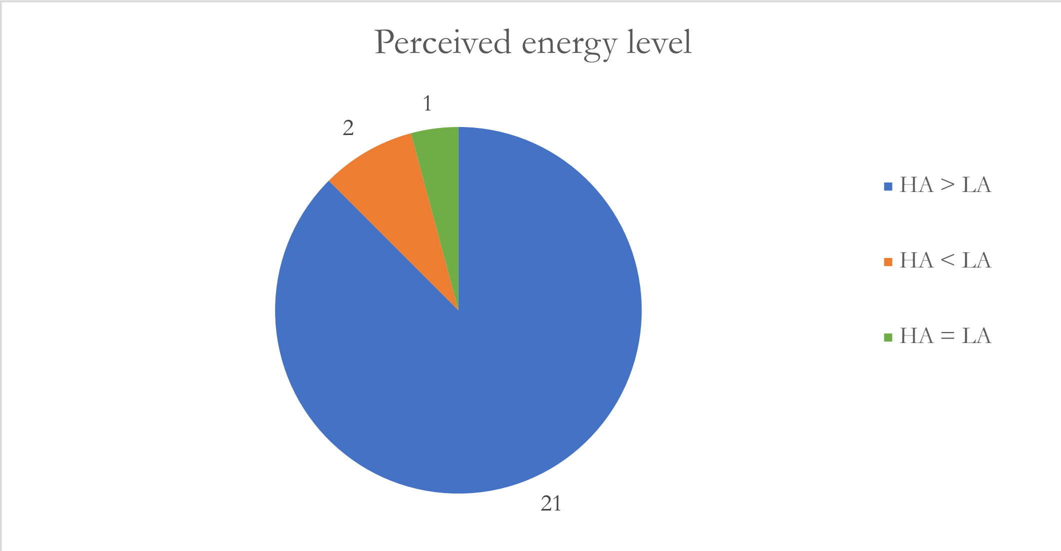

Figure 7. Perceived energy level between the main conditions. "HA > LA" (Perceived HA as higher than LA); "HA < LA" (Perceived LA as higher than HA); "HA = LA" (Perceived both as equal).

4.4.1 Analysis of results for perceived energy level

When observing the pie chart of *Figure 7*, most subjects perceived HA condition as more energetic than LA. A chi-squared (χ^2) test was executed by dividing the value of subjects perceiving HA as equal to LA (1 subject) into the remaining response groups ("HA>LA", and "HA<LA"). This resulted in a χ^2-value of 15.042, which is larger than the critical value of 3.841. Therefore, it can be stated that a statistically significant difference is found between the two response groups.

4.5 Perceived level of complexity results

Figure 8 shows a histogram of how subjects perceived the level of complexity between HA and LA. The question that was asked was "How would you compare the level of complexity of the two sound environments?".

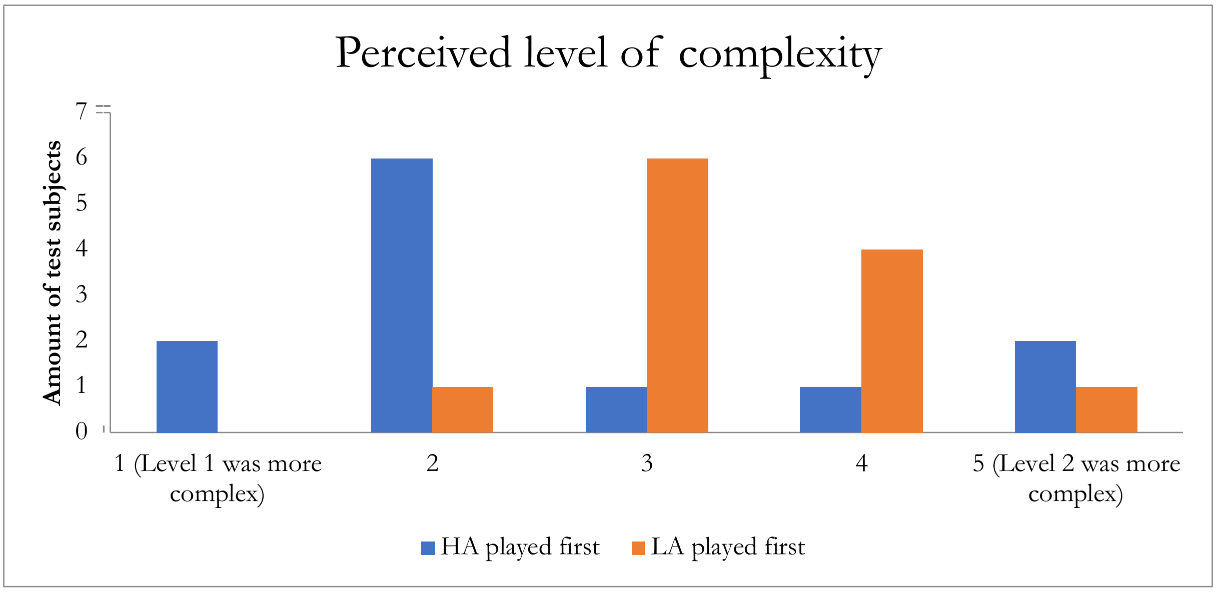

Figure 8. Perceived level of complexity between subjects playing a HA condition first and subjects playing a LA condition first.

4.5.1 Analysis of results for perceived level of complexity

The results seen in *Figure 8* indicate there is not a clear pattern in perceived level of complexity. It can be assumed that the meaning of "complexity" has been interpreted in different ways.

4.6 Responses from subjects playing the game

Table 4 shows a selection of responses from the questionnaire written by subjects. All responses can be seen in the Appendix of this paper. The selection was made from what was seen as rewarding for the study, therefore general compliments, or suggestions about potential improvements of the game mechanics was not included for the results analysis. The question that was asked was "Do you have any other thoughts about the game?".

Table 4. Selection of responses from the questionnaire. “FHA/TLA” = Subject played level FHA first and then TLA, “THA/FLA” = Subject played level THA first and then FLA, and “FLA/THA” = Subject played level FLA first and then THA.

Test Subject no & Playing Condition	**Responses**
1 - FHA/TLA	“Experienced level 2 as more relaxed than level 1, maybe because it was just the other one I did but I also think the sound recorded.”
3 – THA/FLA	“It was easy to distinguish the different objects but it took me a while to start thinking about the sound because you are so focused on finding the objects and ensuring that it is the right color.”
5 - FHA/TLA	“I thought more about the sound in level 1 than in level 2. Above all, the sounds of the birds were more prominent in level 1, while I experienced that the sound in level 2 was more natural and not something I consciously thought about. I immersed myself more in the game in level 2.”
7 - THA/FLA	“I think the church bell worsened my focus.”
14 – FLA/THA	“Sometimes it felt like you had to go a little too close to pick up the items”
15 - THA/FLA	“Interesting. The sound in level 2 was much more pleasant. Another thing I experienced was that when I collected the objects, I tended to take objects that were far away from the boxes, although I later realized that there were objects much closer that I missed completely.”

4.6.1 Analysis of comments from subjects playing the game

A number of subjects seem to have experienced the LA condition as more pleasant, natural, and relaxed, compared to the HA condition. Some subjects thought that the sounds of HA were more prominent, and one subject thought their focus worsened when being exposed to the church bells in the THA condition.

5. Discussion

After analyzing the results of time spent on each level, it seems that environment sounds do not have an impact on player performance. This lack of effect can be seen regardless of subjects being exposed to environment sounds with high or low arousal. However, interestingly, the mean values presented in *Table 3*, shows a lower mean time for the LA (105.792 seconds) conditions than HA (109.875 seconds), though there is not a statistically significant difference between the two. This trend is somewhat similar to the results of North and Hargreaves (1999), showing that subjects performed a faster lap time while exposed to low arousal music, indicating that players tend to perform worse when exposed to high arousal music. They suggested that this was due to cognitive overload.

Similar to North and Hargreaves (1999) who used backward counting for some subjects while playing, making the game more difficult, it can be assumed that some subjects from this study found the game to be a bit difficult too. As seen in *Table 4*, some subjects thought: "Sometimes it felt like you had to go a little too close to pick up the items" (14 – FLA/THA); "it took me a while to start thinking about the sound because you are so focused on finding the objects and making sure it is the right color" (3 – THA/FLA); "I think the church bell worsened my focus" (7 – THA/FLA). Also, this can be compared to the study by Cassidy and Macdonald (2010), where the backward counting design was not used, assuming the game, in this case, was not as difficult as in North and Hargreaves' study (1999). This may then explain why Cassidy and Macdonald (2010) found high arousal music to result in faster lap time, because the concurrent task may not have required as much cognitive space and thus leaving more space for high arousal music. This in turn can be likened to the environment in retail stores, which usually is not seen as a difficult task either. It can be suggested that the listener's view of time, whether the goal is to finish as fast as possible or not, and task difficulty, all influence how the task will be executed, quickly or slowly. However, this study did not have a goal that was to finish the level as fast as possible, unlike the study of North and Hargreaves (1999), which may be one of the reasons why the results were not significant. On the other hand, this may also suggest that qualities like game design, mechanics, and ambience design, simply were not right for this type of experiment.

Even though exposure to high or low arousal did not affect subject performance in relation to time, 21 of 24 subjects perceived the HA sound environments as more energetic than the LA. This also corresponds with subjects' impressions about the game. As seen in *Table 4*, some subjects thought: "The sound in level 2 was much more pleasant" (15 – THA/FLA); and "Experienced level 2 as more relaxed than level 1" (1 – FHA/TLA). It can be suggested that environmental sounds do impact perceived arousal level, and since the difference is perceptible, it may indirectly affect how players experience the game. For example, it may affect their mindset and their level of concentration, even though it does not affect their performance. As seen in *Table 4*, some subjects thought: "I think the church bell worsened my focus" (7 – THA/FLA); and "I experienced that the sound in level 2 was more natural and not something I consciously thought about. I immersed myself more in the game in level 2." (5 – FHA/TLA). Having this information when designing environment sounds for a game could be something to think about and use when making decisions about how to steer player's perceptions. If a designer plans to create environment sounds that have characteristics as either high or low arousal, factors to consider may be tempo (the time between sounding events), presence (amount of energy between 2-4 kHz), and the distance between listener and different sonic elements (how prominent it sounds). There might also exist much more factors to explore, but this can be seen as a starting point for sound designers.

As seen in *Figure 8*, it can be assumed that the term "complexity" was interpreted and used in different ways by subjects. Unfortunately, it is not possible to investigate in what way subjects interpreted the term in this study. It can be suggested that having an explanation for how "complexity" should be interpreted could have been beneficial for clearer results. On the other hand, this might have directed the subjects too much in answering in a certain way.

One factor that may have affected the results is the main game mechanics. It can be assumed that bugs or different problems with the game could have affected the results. For instance, a bug could have been experienced by subjects when player was picking up an object but stood too close to it. The game would then start to glitch for a moment, which would have affected time elapsed. Another possibility is the differences in game design between traffic- and farm-environment, like obstacle placement and general atmosphere may have had an effect. Even though the plan was to make the two environments as close in game difficulty as possible, the results show that mean times for traffic environment is lower than for farm environment (Traffic env = 103.5 seconds, Farm env = 112.167 seconds). The difference in game design between the two environments may therefore have affected the way subjects prepared for the second level differently, thus affecting time performance. *Figure 9* shows a box plot of time elapsed in the farm environment and traffic environment where the mean value, upper and lower quartile, and the outlier can be seen.

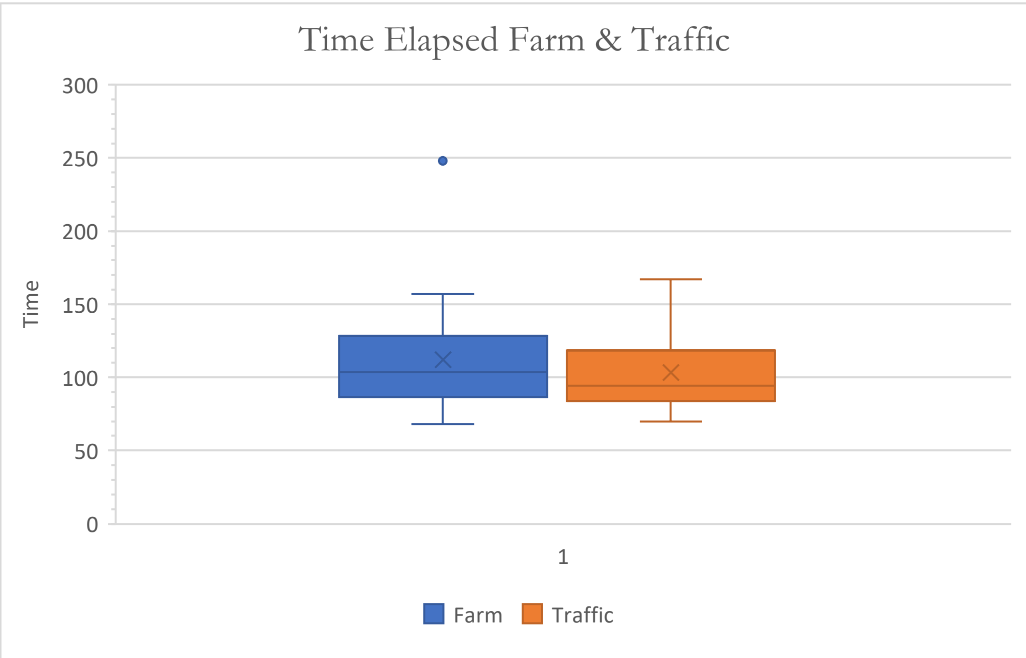

Figure 9. Playing times (in seconds) of farm and traffic environment.

On the other hand, the experiment was designed for subjects to play how ever they desired, as long as they, at some point, completed the task. However, the way to that point, probably looked different between subjects. Some may have been impatient and prepared to finish the game as fast as possible and may not have paid any attention to the background sound. Some may have played around with the different bugs in the game, and some may have just focused on the game task. Thus some might been more aware of the sound and easily affected by the environment sounds. It can be suggested that these different ways of playing would be minimized if the goal was to finish the task as fast as possible, similar to the racing games in North and Hargreaves (1999) and Cassidy and Macdonald (2010). This would eliminate the idea of playing freely. In the case of this game, it may have resulted in clearer results on playing time and would then further resemble the experimental design used by North and Hargreaves (1999).

6. Conclusion

This study investigates the relationship between environmental game sounds and arousal level, looking to see if/how it affects player action. A computer game was created containing two levels, one with high arousal environmental sounds, and another with low arousal environmental sounds. The two levels were in different environments but had the same task, The amount of time it took for subjects to finish the game under each sonic condition was the data of greatest interest. The results showed no significant difference between subjects elapsed times. However, a significant difference could be seen of how subjects perceived the energy level between the two levels.

7. Areas for further research

It would be interesting to investigate how arousal level in relation to environment sounds affect player performance when the goal of the game is to finish it as fast as possible. This could be conducted using similar experiment and sound designing methods used in this study, thus creating a scenario similar to North and Hargreaves (1999). Further research on what environment sounds is perceived as high/low arousal, and investigating what additional attributes can be used when designing these elements would possibly make this area of research more comprehendible and easier to replicate. It would also be interesting to further investigate player perception in relation to environment sounds and arousal. Thus searching to see what environmental sounds may affect factors like immersion and distraction.

8. References

Apple Inc. (2022). *Logic Pro X* (Version 10.4.8). [Computer Software]. Mac App Store. https://www.apple.com/logic-pro/

Assetsville. (2022). *Assetsville Town* (Unknown version). Epic Games Inc. https://www.unrealengine.com/marketplace/en-US/product/assetsville-town

Cassidy, G. G., & Macdonald, R. A. R. (2010). The effects of music on time perception and performance of a driving game: Effects of music on time perception. *Scandinavian Journal of Psychology*, *51*(6), 455–464. https://doi.org/10.1111/j.1467-9450.2010.00830.x

Collins, K. C. (2008). Gameplay, genre, and the functions of game audio. In *Game sound: An introduction to the history, theory, and practice of video game music and sound design* (pp. 123-137). MIT Press.

Dall'Avanzi, I., & Yee-King, M. (2019). Measuring the impact of level of detail for environmental soundscapes in digital games. *Journal of the Audio Engineering Society*, 10202. Retrieved from http://www.aes.org.proxy.lib.ltu.se/e-lib/browse.cfm?elib=20335

Ekman, I. (2005). Meaningful Noise: Understanding Sound Effects in Computer Games. *University of Tampere.*

Epic Games. (2022). *Unreal Engine 4* (Version 4.26). Epic Games Inc. https://www.unrealengine.com/en-US/

Knoferle, K. M., Spangenberg, E. R., Herrmann, A., & Landwehr, J. R. (2012). It is all in the mix: The interactive effect of music tempo and mode on in-store sales. *Marketing Letters*, *23*(1), 325–337. https://doi.org/10.1007/s11002-011-9156-z

McDermott, J. H. (2012). Auditory Preferences and Aesthetics. In *Neuroscience of Preference and Choice* (pp. 227–256). Elsevier. https://doi.org/10.1016/B978-0-12-381431-9.00020-6

Miller, N. (2013). Understanding Soundscapes. *Buildings*, *3*(4), 728–738. https://doi.org/10.3390/buildings3040728

North, A. C., & Hargreaves, D. J. (1999). Music and driving game performance. *Scandinavian Journal of Psychology*, *40*(4), 285–292. https://doi.org/10.1111/1467-9450.404128

Ratcliffe, E., Gatersleben, B., & Sowden, P. T. (2013). Bird sounds and their contributions to perceived attention restoration and stress recovery. *Journal of Environmental Psychology*, *36*, 221–228. https://doi.org/10.1016/j.jenvp.2013.08.004

Rogers, K., & Weber, M. (2019). Audio Habits and Motivations in Video Game Players. *Proceedings of the 14th International Audio Mostly Conference: A Journey in Sound*, 45–52. https://doi.org/10.1145/3356590.3356599

Serafin, S., Franinovic´, K., Hermann, T., Lemaitre, G., Rinott, M., and Rocchesso, D. (2011). Sonic interaction design. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, *The Sonification Handbook*, chapter 5, p. 87–110. Logos Publishing House, Berlin, Germany.

Soundly AS. (2022). *Soundly Pro* (Version 2020.4) Soundly AS Inc. https://getsoundly.com

Wang, R., & Zhao, J. (2019). A good sound in the right place: Exploring the effects of auditory-visual combinations on aesthetic preference. *Urban Forestry & Urban Greening*, *43*, 126356. https://doi.org/10.1016/j.ufug.2019.05.018

Appendix 1 - Instructions used for the listening test, original language (Swedish)

”Hej och välkommen!

Du kommer att få spela ett spel som är uppdelat i två delar. Efter det att du spelat klart den första delen kommer du att få svara på ett antal demografiska frågor. Därefter kommer du att få spela del två. När du spelat klart båda delarna kommer du få svara på några ytterligare frågor.

I spelet finns ett antal föremål utplacerade; hälften av föremålen är gröna och andra hälften röda. Det finns även två lådor utplacerade i både grön och röd färg. Din uppgift är att placera föremålen i sin tillhörande låda, med samma färg. Du ska totalt samla ihop fyra föremål för respektive låda. När du gjort det, är uppgiften slutförd.

Kontrollerna för att röra på dig är WASD-tangenterna och musen. Du använder den vänstra musknappen för att plocka upp föremål.

Innan vi sätter igång med den fösta delen kommer du att få spela en testomgång där du får lära dig kontrollerna samt bli bekväm med spelets fysik. Målet med testomgången är detsamma som huvudtestet. Säg till när du har samlat ihop alla föremål och känner dig redo att starta del 1.”

Appendix 2 - All responses from subjects playing the game

Test Subject no & Playing Condition	Responses
1 - FHA/TLA	"Experienced level 2 as more relaxed than level 1, maybe because it was just the other one I did but I also think the sound recorded."
2 – FLA/THA	"Jumping on space would have been nice"
3 – THA/FLA	"It was easy to distinguish the different objects but it took me a while to start thinking about the sound because you are so focused on finding the objects and ensuring that it is the right color."
5 - FHA/TLA	"I thought more about the sound in level 1 than in level 2. Above all, the sounds of the birds were more prominent in level 1, while I experienced that the sound in level 2 was more natural and not something I consciously thought about. I immersed myself more in the game in level 2."
6 - FLA/THA	"Very nice"
7 - THA/FLA	"I think the church bell worsened my focus."
8 – TLA/FHA	"Sweet! and fun listening test"
10 – FLA/THA	"Comfortable"
12 – TLA/FHA	"It was a nice bug if you stood on the object before grabbing it. Almost that you could fly around. This may have been a reason why it took longer to complete the second round so it was fun to play around with just that thing."

14 – FLA/THA	"Sometimes it felt like you had to go a little too close to pick up the items"
15 - THA/FLA	"Interesting. The sound in level 2 was much more pleasant. Another thing I experienced was that when I collected the objects, I tended to take objects that were far away from the boxes, although I later realized that there were objects much closer that I missed completely."
16 – TLA/FHA	"Good with the training round to learn the physics of the game"
20 – TLA/FHA	"Impressed!"
22 – FLA/THA	"Appreciated the feeling of how the sound adapted to how one moved in the game."
23 – THA/FLA	"Very nice graphics and nice environments the courses took place in. It was very fun to play and easy to understand."
24 – TLA/FHA	"The church bell in part 1 was fun to spin around and listen to!"