SPATIALIZED AUDIO AND THE USERS EXPERIENCE
3D based audio and the affect for the listening experience.

Master Degree Project in Media, Aesthetics and Narration A1E
One year Level 30 ECTS
Spring semester 2023

Kristoffer Eriksson

Supervisor: Lars Vipsjö
Examiner: Ran Zhang
Abstract

With spatialized audio being tools for developers to enhance the experience of playing a game, how does the use of spatialized audio and different types of sounds affect the imagination and the feeling of being immersed in the sound environment? In this study, the use of spatialized audio and complete darkness explores the feelings and how someone perceives the different sounds being heard.

Keywords: HRTF, Audio, Spatialization, Immersion, soundscape
# Table of Contents

1 Introduction........................................................................................................................................... 1

2 Background............................................................................................................................................ 2

3 Problem .................................................................................................................................................. 7
   3.1 Method............................................................................................................................................. 7
   3.2 Method Revision ............................................................................................................................... 8
       3.2.1 Study Survey ............................................................................................................................ 10

4 Study Executions....................................................................................................................................... 12
   4.1 Survey Alterations .......................................................................................................................... 12

5 Analysis ................................................................................................................................................ 13
   5.1 Demographic analysis ..................................................................................................................... 13
   5.2 Artifact .......................................................................................................................................... 15

6 Conclusions.............................................................................................................................................. 24
   6.1 Summary ........................................................................................................................................ 24
   6.2 Discussion ...................................................................................................................................... 24
   6.3 Future Work ................................................................................................................................... 25

References.................................................................................................................................................... 26
1 Introduction

Video games are a medium that keeps evolving both in popularity and in the technology used to create the games, as well as in how to experience the games with technologies such as virtual reality, haptic feedback controllers, and even audio-based technology. These technologies allow developers to create various ways to experience these virtual environments.

With 3D based video games growing over the years, and with that the audio technology that have been developed and are in the forefront when it comes to research is spatialized audio panning meaning that the audio are in a 3D space and can be heard from more directions than regular stereo panning. Databases such as Head-Related Transfer Function has been utilized to improve how the audio is implemented.

This paper will explore how spatialized audio can be used in complete darkness and how the different sound scenarios is perceived.

An artifact was created for testing how different sounds in three different categories and how these sounds affect the user. These categories used was created with the thought of using different sort of sounds that would fit either a nature theme, a machinery theme and lastly an artificial or synthesised theme. The artifact was created with the Unity game engine and uses FMOD as the method of implementation for the spatialized audio.
2 Background

A tool that is related to spatialized audio implementation is something called head-related transfer function or as its shortened to HRTF, explained by Ruixing Wu and Guangzheng Yu (2016) HRTF measures the distance from the audio source to the listeners right and left ear and then reflects the overall effect by filtering from the shape of the listeners head, upper body, and the shape of their outer ears. Wu and Yu further explain how the ears acts in different ways depending where the audio is coming from within a 360 radius, the ears catch different amount of the information depending on distance, rotation and height from the audio source.

James Broderick, Dr. Jim Duggan and Dr. Sam Redfern (2018) describes how the human hearing can locate sounds in a 3D plane although only having two ears, this has to do with that the brain takes the sounds that reaches the ears and compares the difference of time, frequencies and also the changes that can happened with the geometry of the body. Broderick, Duggan, and Redfern describe HRTF as a measurement of these distinctions which can be recreated in a digital setting, however these measurements vary from person to person which makes it difficult to create a universal HRTF database that works the same on different people.

Spatialized audio used in video games can be utilized to provide information to the player depending on the game, with multiplayer first person shooters such as Counter-Strike: Global Offensive (Valve, 2012) and Overwatch (Blizzard Entertainment, 2016) the use of spatialized audio is used to inform the player about their surroundings such as nearby enemy players footsteps or if any weapons are being fired in the players proximity. With the competitive nature of these types of games being able to hear if there is something happening close by, can be helpful for the player, to take in information that could lead to winning or losing.

The horror game Alien Isolation (Creative Assembly, 2014) uses spatialized audio to create atmosphere and tension, during the game the sounds of the xenomorphs crawling through the ventilation system of the space station can be heard from time to time, providing the player with information that the creature may be nearby, which also creates tension for the player. Another modern horror game that uses spatialized audio in a similar way of providing audio-visual information and creating tension for the player is the remake of Resident Evil 2 (Capcom, 2019), where the player is being hunted by a seemingly immortal creature called Mr. X when in certain areas. The sounds effects that are connected to Mr. X are deep sounding footsteps walking around, as the player can´t hurt Mr. X during most of the game, the sound of the creature walking around together with the player fending themselves from zombies or are doing puzzle solving creates tension for the player.
In the Legend of Zelda: Breath of the Wild (Nintendo, 2017) spatialized audio is used as a guide for the player, in the game there are optional challenge areas called shrines that if completed the player can upgrade Links health or stamina. These challenge areas can be found across the game world and are either in the open or require a puzzle to be solved before getting access to them, the way audio is used to help guide the player to these shrines is through diegetic music. Karen Collins (2008, p.125,126) describes diegetic sounds as something that is taking place within the game world such as the footsteps of the character, whilst non-diegetic sounds are sounds that can be heard but doesn´t have a connection the game world itself such as background music. Collins further describes how non-diegetic sounds can become diegetic, if the player sees an instrument and interacts with it, the music that is being played have become diegetic which is what happens in Breath of the Wild.

Early in the game the player will meet a character near one of these challenge shrines, named Kass that has an interest in shrines, one of the noteworthy characteristics of this character is that he is playing an accordion and when the player is interacting with him a specific melody is played with the accordion as the instrument. When the player is running around and exploring the world of Breath of the Wild the accordion melody can be heard and if the player then follows the sound, they will soon find Kass which then provides a riddle of how to unlock a shrine that is near that area. Having the player meet the character Kass near a shrine and having the music of the accordion being played creates an audio-visual contract between the character and the music, creating a connection between the music of the character and shrines, which then when the player is out exploring the world and are hearing the feint melody of the accordion which is then informing them that there is a shrine somewhere nearby.

In Hellblade: Senua’s Sacrifice (Ninja Theory, 2017) the use of audio is unique as it provides for the player what is happening in the game world however it also communicates with the character that the player is controlling, this is in the form of dialogue from voices being heard by the character Senua. This is something that Kristine Jorgensen (2007, p.76) describes as trans-diegetic, which is when the line between something diegetic or non-diegetic isn´t clear enough, “it is characteristic for trans-diegetic sounds that they cannot be posited as clearly diegetic or extradiegetic. Instead, they seem to place themselves somewhere in between the two” (Jorgensen, 2007, p.76). An example of audio being trans-diegetic is during a puzzle where the player has to focus on different sounds and where they are coming from and one of the voices says “Shh. Can you hear it? Varlan’s song! You have to focus! Focus!” which can be interpreted as the voice is communicating with Senua, however it can also be interpreted as the voice is communicating directly to the player telling them what must be done for the puzzle to be solved.

The way how spatialized audio can be utilized to provide information in video games are not limited to 3D games it can also be used when it comes to 2D games, in the platformer Ori and the Will of Wisp (Moon Studios, 2020) spatialized audio is used to create depth of the world and to inform the player of their surroundings, such as if something is above the player. In an interview Dolby with the audio engineer for the game Kristoffer Larson explained that with programs such as Wwise (AudioKinetic, 2006) which is a third-party audio engine connected to the game engine, allows the audio engineer to create sound experiences for games that utilizes spatialised audio functions for people without them having to invest in 5.1 stereo systems, and can instead be experienced with headphones.
In movies sound where split up into three different types, which is onscreen, offscreen and non-diegetic which were presented in a model (figure 1) by Michel Chion (1995, p.73-79), onscreen is described as sounds that are clearly connected to an object that is showed on screen, whilst offscreen sounds belongs to objects that aren´t shown onscreen however it is still connected to the world. the last on non-diegetic are sounds that aren´t connected to the world of the movie such as a narrator.

**Figure 1** Michel Chion´s onscreen-offscreen-nondiegetic model.

Chion later updated the model (figure 2) to include three more categories, that enriched the way sounds can be categorised within movies and helped place sounds that didn´t clearly fit into one of the three categories in the first model, the new categories introduced were on-the-air, internal and ambient. On-the-air is a category which is for audio that is coming from a telephone or a radio, where the source of the audio isn´t on the screen though there is an object that receives the audio. Internal is described as sound that are connected to a character directly, such as inner dialogue or their heartbeat or breathing. The category ambient is for sounds that are present during a scene however there is no question asked where the sounds is originating from such as birdsong.
Jorgensen (2010) criticises the way the diegetic and non-diegetic terms is used as explanation of sounds in video game as the terms originated from the movie medium, the critique arrives from how the two mediums differentiate from each other, with movies being a linear experience whilst games are interactive making the line of what is diegetic and what is non-diegetic vaguer. An example Jorgensen provides about the indistinct line between diegetic and non-diegetic is in the Elder Scrolls III: Morowind (Bethesda, 2002), where the player is walking through a forest and then hears the music change, which warns the player that there is an enemy approaching them. The music itself doesn’t have a source in the game world, yet the character is being controlled by the player to act upon the approaching enemy. What has happened in this scenario is that a non-diegetic sound has affected a diegetic event in the game, these types of interactions between player and the non-diegetic sounds is what creates ambiguity between what is diegetic and what is non-diegetic in games.

Based on Chion’s model a model for categorizing sounds in games was created called IEZA (figure 3), this model separates the sounds into four categories which are interface, effect, zone and affect, these four categories are divided by diegetic and non-diegetic however it also adds setting and activity to the grouping. Andreas Westerberg and Henrik Schoenau-Fog (2015) describes interface as when the player interacts with something that isn’t connected to the game world though is playing sounds such as navigating to the game’s menu. Effect sounds are connected to an action happening in the game world, such as firing a pistol and hearing a gunshot being played, the zone category describes sounds which relates to the scene and is diegetic such as the sounds of a waterfall running. Last category in the IEZA model is affect and is used to describe sounds that are related to the setting however is not diegetic, for example emotional music.
An updated model based on the IEZA framework was later created called FGAMC (figure 4), which stands for Flexible Game Audio Moderator for Categorizing. The FGAMC model exchanges diegetic and non-diegetic to game world and game space which were based on a theory by Jorgensen (2010, p.89-91) about diegetic sounds in video games, the term game world is used to describe the playable world whilst game space is the elements that belong to the game though not to the game world such as heads-up displays.

**Figure 3** IEZA framework model

**Figure 4** FGAMC model
3 Problem

How does spatialized audio and soundscapes when used without any visual aspects connected, effect the experience for the user?

3.1 Method

For the artifact different scenarios from Alice’s Adventure in Wonderland (1865) by Lewis Carroll was adapted into linear soundscapes with spatialization of the audio in mind.

The reason to adapt Alice’s Adventure in Wonderland is for the different scenarios that happens during the book that could create interesting scenarios translated into spatialized audio, scenarios such as the opening when Alice falls down the rabbit hole and different objects are passing her. The book is also used to provide a base for the sound design and the decisions revolving around it to create the soundscapes.

The artifact is created with the idea of placing the player in the position of Alice as they listen to the different soundscapes, the artifact is sound based only with no visuals, not having any visuals is to put focus on the sound and how sounds can be used for immersion, having visuals could take away the focus on the sounds and pushing it to the visual aspects.

For the soundscapes the different sounds are created using the digital audio workstation FL studio by Image Line, by using different foley recordings and digital synthesisers in conjunction with digital signal processing effects before being rendered and implemented in the audio engine FMOD.

The artifact was created by using the game engine Unity, this is to get access to the audio engine FMOD for the spatialization of the soundscapes. In the artifact the user can control the camera by using the mouse, depending on the scenario movement can be utilized by the W, A, S and D keys. The different scenarios are chosen at the start of the artifact and to get back into the scenario selection the esc key is used to show a menu.

FMOD provides fuller control of the audio settings than if it were implemented directly into a game engine, in FMOD it provides control over how volume or frequency cutoff can be changed depending on different parameters, such as distance and height, which is useful with spatialization of the audio.

A stereo version of the soundscapes is also created to be used as a comparison with the spatialized version to see if there are any difference when it comes to feeling immersion when listening to sounds.

After listening to the different soundscapes, a survey is filled out with the goal of finding out how the contestant felt about the experience, if they felt some type of immersion or connection to listening to the sounds. Interpretation of the soundscapes is also asked in the survey, this is to gather information on what the contestants though the soundscapes was about or if they were reminded about something whilst listening.
3.2 Method Revision

The artifact where originally planned to have a selection screen for the participants to choose between different soundscape scenarios to listen to and then answer a survey about these different soundscape scenarios. The selection screen wouldn’t tell what specific scenario each selection had, it would just be numbered as scenario one, two, etc. the decision to not inform the participants on what the different scenarios where, had to do with gather data on how what they thought the scenarios were and if the soundscapes made them feel some sort of immersion.

The game artifact revision was made to be a larger area that contains different soundscapes within, this was to let the player themselves explore and find these different soundscapes, instead of choosing each scenario. The decision to have the player freely roam around in a larger area rather than several smaller ones has to do with making the experience feel more like a game of exploration than just a listening experience.

When it comes to the sounds of the artifact themselves a change have been made, moving away from the focus on adapting different scenarios that occurs in the book Alice’s adventure in wonderland. The change has been made to have a vast set of different types of sounds that together creates different soundscapes, these types of sounds vary from nature sounds such as wind, rivers and birds, mechanical sounds such as printers, motors etc., and artificial or synthetic sounds which is based on sounds created through synthesis and not truly based on sounds that could be heard in the real world however could be utilized in genres such as science fiction. The point of moving away from the adaptation idea have to do with that the focus of the study is on spatialized sounds and the experience of the user, and not if the participants can recognise a story scenario based on sounds. Another point of moving direction with that artifact sounds have to do with creating soundscapes based on different textures and rhythms that the sounds have as well as other characteristics that the sounds have such as frequency information and how they can be utilized to shape the soundscapes. Utilizing the frequency spectrum lets two soundscapes that are in the same category as each other have a different feel even if the textures and rhythms of the sounds are similar, if there are two nature soundscapes for the player to explore and listen to, with one of them have a more focus on lower frequencies and the other higher ones, does the participants feel a difference with not only immersion but also generally what they feel listening to the different scenarios even though they are in the same type of category of sounds.

The sounds in the artifact and their function based on the sound modules are ambiguous such as with Chion’s updated model, there are sounds that would clearly fit into the category of ambient such as the different sounds that are based with nature in mind, whilst some could be interpreted to fit into either offscreen or onscreen though where to place these becomes a challenge as in the artifact there isn’t a screen to reference whether the sounds belong on or offscreen.
The IEZA and the FGAMC models was in mind during the creation of the different soundscapes, such as the sounds being categorized as either effect or zone as the different sounds were designed with the idea of them being diegetic in a normal game setting. Most of the different sounds in the nature scenario fits into the zone category as they are created with the envision of being there physically in the game world for the player to hear though not being interactable with. The mechanical sounds fit into the effect category as these types of sounds are created with an object being interacted with in mind such as a keyboard or a motorcycle, these sounds can be also in the zone category as they can be interpreted as sounds being created not by a players interaction but with the idea of someone else such as a non-playable character are the one that are interacting with an object which makes the sound.

The third scenario and the abstract nature of these sounds makes it difficult to place in either zone or effect as there isn’t any concrete objects or animals for the sounds to be connected to, the one category that scenario three could fit into are affect, which related to a setting though not being diegetic. When it comes to the FGAMC model the three scenarios can be placed somewhere between emphasized interface and integrated interface as the sounds could be part of a setting or a result of an activity.

When starting the artifact, the participants are introduced to the controls of the game, and what they will do once inside, the game world of the artifact is pitch black and contains no visual information other than the introduction text at the start, the reason for this is to put focus on the sounds themselves and not be influenced by the visuals. An example of this would be if when in the nature soundscape the participants could not only hear a river but also see that there is a river close by, this would remove some of the ambiguousness that sounds have and making the connection between audio and visuals be clearer and could change how the sounds are interpretated.

For the participants to play the sounds of the artifact the left mouse button had to be pressed down, this then casts invisible lines from the audio source to the audio player which is connected to the camera, and if the player are within a close proximity and nothing obscuring the path then the sounds will be heard and if walking towards the source will then make the audio louder and clearer for them to hear. What also effects how loud or clear the audio is for the player has to do with the direction regarding where the source is and where the camera is pointing, if the source is behind and the camera is pointing in the opposite direction the audio will be quieter and have some low pass filtering effecting the sounds, with the functions of spatialized audio implementation it would also make the sounds seem to appear behind the player. Another example on how the position of the camera effects the sounds have to do with left or right position of the player, if the audio source is more on the left side of the player, most of the sounds received will be with the left ear whilst the right ear wouldn’t receive as much sound information and the same goes if the sound in on the right side of the player, the left ear would receive less audio information than the right ear.
Another function of the lines that are being cast from the source to the player, have to do with if there are anything obscuring the path, such as walls, if the path of the lines are obscured it will filter the sounds making them appear more muffled, this is to further mimic how sounds work in the real world as if something is played in another room the sounds received would be muffled. The reason to let the player themselves activate the sounds with a button press and letting the players distance from the source matter as it has to do with not to overwhelm the user with lots of different sounds constantly being played, letting the player themselves choose when they want to the audio to be played.

3.2.1 Study Survey

The survey for the study was created with Google forms for ease of use and as a convenient tool to send out the survey to the participants online. As the study focus is to gather qualitative data about audio and the experience of the user, the survey has been created with that in mind, to gather different thoughts and feelings about what the participants had experience when playing through the artifact.

Before the participants start with the study, they are informed that if they participate the information will be compiled with others answers and no personal information will be connected to them, information such as age and gender identification will be used to identify the demographic of the study and that the answers provided in the study will only be used for this study. In the survey form the participants will be informed that if they consent to be a part of this study, they are free to stop at any point and are also informed that if they decide to stop that they should not submit any answers so that there will be no mix up with the other answers. If the participants agree to this, they are informed to start the study by downloading the artifact which is provided in the survey form, once the artifact have been downloaded and opened there is information on how to play and what to do once inside the game, the starting screen also informs that once the participants feel ready to continue with the study they can close the application and return to fill out the survey.

The survey (Appendix A) starts with determining the demographic of the study, to gather this data questions about the age of the participants and what gender they identify as is being asked, this is to get a base on who is participating in the study. Another thing being asked in the beginning of the survey to further understand the demographic is what their relationship with video games are, how regular they play and what they play video games on if they do. As the study not only focuses on video games but also audio, a question about what headphones is being used by the participants will be asked, this question is to gather data on the frequencies that the headphones have, the reason is that not every headphone may have a full register when it comes to frequencies and could alter how the sounds are being perceived.

Once done with the demographic and general part of the survey the participants continue with the part about the artifact itself, starting with if the instructions on what to do was clear and understandable, followed up had to do with what the participants thought the study was about, this is to understand the participants perspective going into the study.
Questions was asked about how many different areas of sounds, they could locate, if there were any complications founding the areas and if they could categorize the different sounds, the part after was on how they felt when listening to the soundscapes they found. Questions about immersion is asked after this, questions such as their definition on immersion, what creates the feeling of immersion for them and lastly on how immersed they felt playing the artifact and if there were any specific area which they found that made them feel immersed in and why they felt that.

Once the question been answered the next page of the survey provides the explanation of what the study was about, the reason on why the explanation is at the end of the study have to do with not trying to lead the participants in any direction that could affect the results. In this part the participants also could add anything if they wanted, if there is something new that they might have thought about after knowing what the study was about or any other thoughts relevant to the survey or the artifact.

Another reason why using Google forms to create the survey is the tools that are provided with compiling the data into graphs such as pie charts, making the data easier to read and analyse, whilst with the answers that are done through free-flowing text are organized into the question they belong to, the option to go through each individual answer is also available making it easier to analyse each participants by themselves and together.
4 Study Executions

4.1 Survey Alterations

Before the study commenced changes were made to the survey (Appendix A), the alteration had to do with some of the questions asked during the artifact section, the reason for the alterations were to focus on the artifact itself and ask questions that is tailored to it.

With this alteration of the survey some of the questions where either rephrased to further expand on the topic that was asked, while some question got removed and exchanged with new ones.

The first two questions remained the same which was about the clarity of the instructions for the artifact and what the participants thought the study was about. Question three changed from “How many different areas with sounds did you find during your playthrough?” to “How many different areas with sounds did you find in each scenario during your playthrough?” this had to do with the change of the artifact itself as the first idea iteration had multiple areas in one scenario whilst the artifact that was used for testing had three different scenarios to choose from with each scenario having different types of sound categories from each other. A question about if there where any complications locating any of the sound areas were added before an unaltered question about describing or categorizing each sound area the participant located was asked.

In the first version of the survey which was asked after the questions about the sound areas, had to do with describing and what creates the feeling of immersion, having these questions about immersion and if the participants felt more or less immersed within certain areas of the artifact were removed completely from the survey. The removal of these questions had to do with the change of the artifact itself as the first idea iteration had multiple areas in one scenario whilst the artifact that was used for testing had three different scenarios to choose from with each scenario having different types of sound categories from each other. A question about if there where any complications locating any of the sound areas were added before an unaltered question about describing or categorizing each sound area the participant located was asked.

the questions that were replacing the questions about the generalized topic of immersion was changed into asking about if the different sound areas made the participants feel anything while listening and if so if they could describe what they felt. questions about which scenario the enjoyed the most and which one they least enjoyed and why they felt that way where asked, question about which scenario they felt like revisiting and which they would not revisit was also added. The last question about the scenarios were about which if any scenario felt like the participants spent more time being in and why they may have felt like that.

The last section of questions in the artifact part of the survey has to do with playing in complete darkness, asking about the experience of playing in darkness, if it added anything to the experience itself. Lastly if there where any visuals would that affect the experience in any way, and to further add on the visuals a question about what they would add as a visual to the different scenarios if they could and describing why they thought of that.

The final change of the survey had to do with the final message for the participants after filling out the survey, telling them what the study was about and what the three different scenarios of sounds was. (Appendix B)
5 Analysis

5.1 Demographic analysis

In the study there was a total of nine individuals that participated in the study, with an age range that was between twenty-one up to twenty-eight, among these participants there two individuals that identified themselves as female, while the remaining seven individuals identified as male.

![Figure 5: Age group of participants](image)

When asked about the relationship they had with video games, six participants answered that they played every day, while two play games regularly and one answered that they play sometimes. Out of the options on the primary device for playing games, playing on pc was the option that was picked the most, there were one answer that stuck out from the rest when it comes to the platform of choice and that was mobile, the one participant with mobile as their primary device for video games was also one of the individuals that identified as female and answered that they played video games daily.

![Figure 6: Gender identification](image)
When looking into if there was any correlation between the answers provided with the relationship of video games and the participants age or gender identity, there was not a clear connection that was recognized with the limited sample size of participants. The two individuals that said they play video games regularly both identified as male, they were also both the youngest participant as well as the oldest one of the nine, the individual that said they played games sometimes was one of the individuals that was twenty-four year of age and they also identified as male.

Before the artifact part of the survey there was an optional question about the headphones that the participants used during the playthrough of the game, this was to get an overall grasp of the frequency spectrum that the participants headphones had in range, however as modern technology have improved during the years, the majority of headphones used had a range of 20 Hz to 20 kHz which is the full range of human hearing, while some had a much larger range of frequencies. This question was asked to see if there was a user with headphones that didn’t have a full frequency range as that could affect how the sounds was perceived, a thing that was overlooked is how headphones with a larger range than the human hearing could affect how the sounds is perceived.
5.2 Artifact

When it comes to the part of the survey that was about the artifact itself, the start was about if the instructions on what to do in the game was clear or not, eight of the nine participants though that the instructions where understandable, and the last one answered that the instructions weren’t clear. Something that was overlooked with this question was that there wasn´t a follow up question if someone though the instructions weren’t clear enough, with this overlook there is no understanding on what exactly wasn’t clear in the instruction part.

![Pie chart showing the responses to the question about the clarity of the instructions.]

**Figure 9** Answers for understanding the instructions.

The first official question after the question about instructions had to do with what the participants though the study was about which resulted in a few different ideas on what the study was about. There were a few answers that overlapped in in the general idea of what was written and could be put in two types of categories or groups, these categories can be summarized as the following, spatial audio perception and audio environment, however some of the answers could be in either of those groups as how they are phrased leave it up to interpretation.

There were four answers that would fit in the spatial audio perception group, all having different phrasing to describe something similar, which went from directional audio, spatial audio perception to describing the thought of spatial audio perception with different words. One of the describing answers said, “the importance of spatial sound to navigate in a 3D environment” while the other said “how sound is heard from different angles”. These two answers are similar in a sense on what they could be interpreted as, with “the importance of spatial sound to navigate in a 3D environment” is talking about using audio to navigate through the game, whilst with the sentence “how sound is heard from different angles” could be interpreted if talking about the x, y and z position and angles that an object or camera has in a 3D based game and the perceived audio that corresponds to those angles.

When it comes to the answers that best would fit in the audio environment group the thing that most answers have in common is the keywords used, such as environment or soundscape. There were a few answers that wrote about identifying or recognizing what the different scenarios environment was, such as “identifying different environments through the use of audio cues only” and “I think the study is something related to being able to recognize what kind of environments I am exploring simply with sound and moving around”. There were some answers that provided different ideas on what they might think the study was about, such as it being a study of gaming for the visually impaired, and one answer mentioning different ages of human technology on earth.

15
The next question asked how many different areas with sounds the participants found in each scenario, this provided a nuance of answers as some provided just a number, whilst some gave a number for each scenario and some even provided a list of what sounds they could hear. If taking the answers that just said a number and going through with that the question asked about the number of sound areas in each scenario the range would be from three to seven, one answer said they found about eight to ten different sounds in scenario one and two whilst in the third scenario they found six or seven sound.

There was one participant answer that provided a list of the different sounds they found in each scenario, in the first scenario they said that they found a total of nine sounds withing four areas, those being rain, thunder-rumbling, thunderstrike, flowing stream of water, bushes rustling, strong wind, step in snow, birds chirping and birds flapping their wings.

In scenario two they found a total of ten sound within four areas, these were motorbike engine, scissors cutting or being dropped, paper printer, typing on a keyboard, some machinery, a bellow, buzzing of lamps, a generator, birds chirping and faint wind or machinery. In the last scenario they found five sounds within four areas, these was described as sci-fi laser, a different sci-fi laser, some wind, sci-fi machinery, and something like a hovering sound. This response is interesting as with scenario one which is the scenario with the nature theme of sounds, the participant managed to identify all of the sounds correctly, whilst for the second scenario being themed around machinery and mechanical sounds, they managed to identify some sounds correctly such as the motorbike engine and the keyboard sound, and with the third scenario the sounds they guessed were not that far from what the sounds could have been though with the abstract nature of the synthesised sounds it can be difficult to fully describe those correctly.

The question about if there where any complications locating any sound areas, there was some that felt that navigating around was a bit complicated as some described them getting lost, either when walking around not knowing if they were stuck on a wall or getting lost when moving the mouse around and not knowing how fast the mouse moved made it difficult to navigate. One participant though it was hard to tell if they were walking somewhere or just walking into a wall, they also mentioned that some sound clips were harder to tell which area they belonged to. There were three individuals that answered with “not really” then providing their thoughts on why, such as one felt like it was divided into about three-by-three or four-by-four pieces in a square and then walking to the edges and moving between the noisy areas. Another described that it took some time to figure out a path for the character based on where the sounds are but after that it was straightforward, the last person that started the sentence with “not really” said that they tried to imagine that they were in an empty room and then went from corner to corner. There were two individuals that mentioned potential problems with the audio, one mentioning that they thought that there was only one sound source per scenario but other than that they had no problems, the other individual mentioning having the game muted which made them though the game was broken and after that mentioned that they first though there would only be one area and got confused when hearing multiple sound sources at first.
Asking about how the participants would describe or categorize each sound area they located provided in some interesting results, the most common word which most individuals used was forest which were used by five different answers, however forest was not the word that was used the most, that word was sci-fi though it was only said by four individuals, one of the answers used it three times in their answer as they described different sounds. With scenario one the thing a lot of the answers had in common with each other was mentioning or describing nature, the responses that mentioned forest also mentioned bird song and river or flowing water of sorts, and how the word forest frequently was written in five different responses showed that some of the sounds in scenario one corresponded to an idea of what they imagine a forest sounding like.

The second scenario didn’t seem to have a direct correlation with a type of setting such as scenario one had, though something that was a bit common were the idea of scenario two being an office, an industrial area or even a workshop. The thirds scenario had sci-fi as a connection with the responses, from either the four individuals that mentioned the words, or that futuristic technology and even alien spaceships was being mentioned, which can be grouped together into being science-fiction as those are ideas that show up frequently in the genre of sci-fi. There were a couple of responses that provided different insight of describing or categorizing the sounds, such as one response mentioned that the scenarios reminded them on different types of genres, such as the first scenario reminded them of a psychological game of sort, the second scenario felling either industrial or robotic whilst the third one was more towards sci-fi. One individual provided a short although interesting answer as they described each scenario with a single word, those words where calming for the first scenario, interesting for the second and scary for the third scenario.

The next question was about when listening to each area located, did the sounds make the participants feel anything and describe the feeling the felt if they did, this provided a lot of different responses. When it comes to the first scenario a feeling that got mentioned frequently was calming or relaxing, with these responses most of the time they also mentioned it in connection with the “forest” sounds. however calm and relaxing wasn’t the only feelings that got mentioned, another common feeling describes was creepy or scary, this related to the thunderstorm sounds and one area that was describes as the cave area. There was one response about the first scenario that stood out in a way, and it said “Yes, the first scenario. Creepy soundscapes are uneasy to me. If this was a real game and the room was pitch black as in the test, I would be really scared if the audio was the only lead.” This was interesting as when the following questions about which scenario they enjoyed the most out of the three and why the individual answered that the first scenario was the one they enjoyed the most and the reason why had to do with the horror elements and thought that was the purpose of the scenario.

There were a few responses that provided some input on the feeling that scenario two gave them, and two out of the four which mentioned scenario two in their answer mentioned that it made them feel as they either was working or at a workplace, there was also one that though that the second one was the most interesting because the like mechanical sounds. One of the participants mentioned that they where mostly focusing on where the sounds were going to be and not necessarily associating them with feelings, however out of the three scenarios they did mention the second one as daydreaming and how the sounds of a typewriter made them feel like an author.
The third scenario had the least amount of being mentioned when it comes to feelings that the participants felt, however there was some connectivity between some of the answers and that had to do with mentioning aliens either experimenting or simply just interacting with them, there was a couple of emotions that showed up as well, these were odd or weird but also lonely and uncomfortable.

When asked about which scenario option out of the three did the participants enjoy the most and why, five individuals answered that they enjoyed scenario one the most, whilst the remaining four said that scenario two was the one they enjoyed the most. Out of the five that answered scenario one as the one they enjoyed the most, two of them mentioned the reason had to do with the sound of nature, some mentioned that it had the most calming and relaxing areas and therefore scenario one was the most enjoyable. A response mentioned that the reason on why they enjoyed scenario one the most had to do with that they felt it was the one that had close to a clear theme and that it was also easier to describe the sounds they were hearing. The last response that enjoyed scenario one the most of the three was enjoying it for the reason of that it is containing horror elements to them, and that the area scared them made them enjoy it more.

When it comes to the answers on why the four individuals that enjoyed the second scenario the most, one answer was about the sense of familiarity and it had sounds that they could hear daily, another said that it made them think of ASMR. The other two had similar thoughts on why they enjoyed the second scenario the most, one wrote that the typewriter felt satisfying and having the printer on the side made them feel as if they were a writer and was publishing a book. Whilst the last of the individuals had a story in mind when listening to the sounds of scenario two, the story they mentioned was about sewing machines, handmade processing and that a tailor is working at that place, they also mentioned that they prefer older and historical things and that it reminded them about their grandma which made it feel warm and memorable to them.

![Figure 10](image1.png)

**Figure 10** answers for which scenario was enjoyed the most.
When asked which scenario the participants enjoyed the least and to describe why so, five out of the nine participants said that scenario three was the least enjoyable to them, whilst three said scenario two and one person said scenario one. The individual which said that they enjoyed scenario one the least describe it to be because they didn’t associate the sounds with a lot of stuff, and that it didn´t feel as memorable as the others. The three that said scenario two was the least enjoyable all had different mindset on why they didn’t enjoy that scenario as much, one though that the different sounds didn’t go well together and that it also reminded them about something else they have seen in games or films, one though that the scenario in their own words “gave away spooky vibes”, whilst the last of them explained that it had to do with the industrial sounds as industry locales are dirty and dusty and therefore did not appeal to them as much, however they did mentioned that some industrial themes such as steampunk could be interesting.

As for why five of the participants felt why they enjoyed scenario three the least, most of the reason they gave was about that the sounds felt unfamiliar, strange, and not natural, such as one wrote “sounds that I can´t hear on a daily basis, unfamiliar or never heard before” while another wrote “it was strange and scary. Non natural I would say”. An answer that was provided described how it was hard to tell if there was a new sound or if they had already heard it before, mentioning how some sci-fi sounds sounded very similar to each other and that it was hard to describe the sounds with more words rather than “sci-fi”. With the topic of the sounds being different from the other scenarios, one participant described that they thought that the other two before was “cute” while this one was different and made them confused, they also mentioned that the sounds were loud and came out suddenly which could have been an overlook when implementing the sounds. the final response on why they enjoyed scenario three the least had to do with them feeling that it had too many high pitch sounds which were becoming annoying to listen to after a while.

![Pie chart showing the responses to the question: Which scenario option out of the three did you enjoy the least?](image)

**Figure 11** Answers for which scenario least enjoyed.
The next set of questions had to do with which out of the three scenarios the participants thought they would revisit and which one they wouldn´t revisit, six of the participants answered that scenario one was a scenario they thought that they would revisit while the remaining three said scenario two. With the question about which scenario the participants thought that they wouldn´t revisit there was an error made which was that the question was not set as mandatory which resulted in only eight out of the nine participants answered the question, four out of the eight participants that answered the question said that scenario three was the one the wouldn´t revisit, three said scenario two and scenario one had a single person saying that they wouldn´t revisit that scenario.

![Figure 12 Answers on revisiting a scenario.](image)

![Figure 13 Answers on not revisiting a scenario.](image)
The last question that was about the different scenarios had to do with if there was any scenario which the participants felt as if they spent more time in over the others and if they could describe on why that might have been, there was three individuals that said that they felt that they spent more time in scenario one, three for scenario two, and two for the third scenario, whilst there was a single individual that thought that they didn’t spent more time in one of the scenarios from the rest of them.

With scenario one there was a reason for spending more time in that scenario given by two individuals, which had to do with trying to understand what is going on at first and that one got lost and felt confused at the beginning. The third response said they spent more time in scenario one because it felt relaxing and that it made it easier to visualize the environment that they were in from the sounds. The thought of scenario one being relaxing wasn’t just a single persons thought, there was one individual that said that even though they think they spent more time in scenario three as they were trying to figure out what the different noises could mean or what they do, they felt that otherwise they spent more time in scenario one as it felt peaceful. The other individual that said they might have spent more time in scenario three wrote that they spent more time because of the “laughter sounding thing” and were trying to find the source as they thought that the sound was amusing to them.

With scenario two, there was responses about comfort and relatability, one said they were trying to figure out more sounds since it was an environment, they felt comfortable in, while another said that the sounds were the most relatable to them. Then there was the final individual which felt like they needed to hear the sounds multiple times and, in their words, wrote “Scenario 2 because I was questioning whether or not I was going crazy, so I needed to hear the sounds a lot.”

The final part of the study had to do with the experience for the participants of playing a game in complete darkness, and in general the responses on that was positive as the participants though that it was interesting or unique to play in darkness. One participant wrote “it was fun due to it not being a common trope in video games, which makes it feel new and exciting”, there was two responses which both said that they closed their eyes when playing to either be able to focus more or that they closed their eyes for imagination. The topic of imagination comes up a few times from different responses and how with it being dark made them think on what the environments would look like, with one saying that it made them sort of creating their own game in their mind as they imagined where they are going based on the sound. One participants mentioning that that it was different though they liked the idea that they let their mind create the visuals for them, however they also mentioned that they would have liked to see what was going on especially for the third scenario as they wanted to see what was happening in that area, they also said that playing in darkness during the first scenario felt the most natural to them as it could have been something being played in the background at their home. One individual mentioned how with the 3D sounds made them feel like they were moving around and that it would be “creepy” in a real game.
When asked if they though that the darkness added to their experience of playing through the different scenarios, seven answered yes and two answered maybe. One of the persons that answered maybe also had described for their experience that it was interesting not seeing though they also wanted to be able to see what was going on, the other person that said maybe wrote “it provides an opportunity for player’s imagination to create different levels of experience”. The individuals that answered yes to the question about it they though the darkness added to their experience, had a similar response on why they though that the darkness added to their experience and that was focus. They explained that with their being complete darkness they were able to put more focus on the sounds that they heard and made them more aware of them and with that made the sounds feel more impactful, there was some mentioning that if there were visuals, they wouldn’t have put as much focus on the sounds but rather the visuals.

![Figure 14](image1.png) Answers for if playing in darkness added to the experience.

When asked if the participants though that if there where any visual to the different scenarios would their experience be different, seven said yes to that it would their experience would have been different, on the other hand there was one that answered maybe and one simply that they didn’t know.

![Figure 15](image2.png) answers on if visuals would change the experience.
With the final question they participants got asked if they could assign any visuals to the different scenarios what those would be. This question could have been formulated in a better way or removed completely as it did seem to cause some confusion on what is exactly being asked. However even though the question could have been rephrased there was a couple of interesting views gathered on what types of visuals that could be assigned together with the scenarios. There were a few recurring depictions such as with scenario one would have visuals that are depicting a forest and that scenario twos visuals could be either depicting a factory or and office, whilst the third one got mentioned that the visuals would be depicting alien technology or something else that is connected to science fiction. Whilst there were some answers providing a setting or even objects there was one answer stating that they would assign some visual aid to help them navigate the surroundings to help them to tell if they were currently walking into a wall, there was also one that said to add something with great shock value in the darkness or fading into something visually astonishing.
6 Conclusions

6.1 Summary

Using spatialized audio and HRTF to create sound scenarios within complete darkness showed that depending on the types of sounds being used it can be utilized to let the players put more focus on what they are hearing, and what types of environmental setting they could be in. By simply focusing on the sounds a person’s imagination can take over and envision what the sounds being heard are, and with sounds being as abstract as they are, it is worth mentioning that it matters more on what is being heard and what someone envisions that sound being than what the actual sound was before implementation in a game setting.

How well someone enjoy or connects to a certain setting or sound can create the feeling of relaxation, it is also worth mentioning that sounds that are more familiar for a person makes it easier to then envision a setting or environment that makes the person feel comfortable in.

By testing and comparing different sounds and sound scenarios, showcased that the feelings that these sounds create differ vastly from person to person and that they enjoy and would potentially revisit scenarios that are made from sounds which they can imagine clearly what is making those sounds. and when it comes to feeling if they spent more time in one scenario than the others had to do with either the feeling of relaxation and comfort or interest in the sounds being heard.

Something that is worth mentioning is that the implementation of the sounds wasn’t perfect and that some of the sounds may have been either louder or quieter than other sounds.

6.2 Discussion

With game genres such as horror relies on sound to create atmosphere and by utilizing 3D based audio and HRTF it could be used to further create environments for the player to be immersed in when it comes to the feelings that horror games want the players to feel when playing those games. Though horror games isn’t the only game genre that could benefit using spatialized audio when it comes to wanting their players to feel immersed, developers can use spatialized audio to create experiences with relaxation in mind and having audio to focus on could help with that.

With more understanding on how audio and spatialized audio can be used to as tools to expand the players experience and how they feel when playing games is something that could benefit developers to create better experiences for the player.

There were some limitations with this research, which is worth mentioning, one of them being the smaller sample group of participants that partook with this study, the nine individuals that participated did provide interesting and useful information when it comes to the experience of spatialized audio in complete darkness however having a larger sample group could further expand on the results provided during the study. By having a larger sample group could diversify the participants demographic further and include groups of individuals that could provide useful and interesting insight with this study, such as individuals that are older or don’t have seemingly strong relationship with playing video games such as the nine participants had during this study.
Another limitation that is worth mentioning have to do with the questions asked to create a demographic for the study, as the questions asked about the participants age, gender identify and their relationship with playing video games, this does provide a view of the studies demographic though it could have been expanded to further understand who is partaking. By further expanding the demographic section of the survey to include questions such as about occupation which could be used in correlation with the question about the relationship to video games and see if there could be a connection with what the participant does and how much time they spend on video games. When it comes to the relationship with playing video games not only could the occupation of the participants affect how much they play but also in what way the play such as pc, console or mobile, another question related to video games that could have been explored is about the games or genres that the participants are usually playing, to further understand the connection to games but also the types of sounds that they would probably be familiar with when playing games.

6.3 Future Work

To further expand on this study, changes would be made to make sure that all the sounds are at similar volume level, so that there isn’t a sound that feel overpowering, when listening. With this study focusing on three different categories and comparing them to each other, an expansion on this study could be more focused on a specific scenario rather than three different ones or other types of scenarios. Instead of putting four different areas with sounds in each scenario, it could be beneficial to have just one sound areas and then comparing that one with another sound area in a different scenario within the same type of category, such as comparing two different areas from the nature scenario with each other and see how the results differ.

Spatialized audio has a lot of potential to be further studied, with more understating on how it can be used to create feelings and how it can be used to let a person’s imagination create the environment when playing a game can be expanded on creating new ways of experiencing media. Expanding on this would be to explore how far spatialized audio can be used to create experiences that are more in the shape of a full game or world.

Not only would it be worth studying how spatialized sound makes someone feel or experience a game, but it is also worth to further research the use when it comes to an option of accessibility, letting more individuals be able to enjoy games. By researching spatialized audio and it’s use as an accessibility tool and how developers can use the tool in their games no matter what type of game they are developing, with programs such as FMOD can lead into opening the possibility to include visually impaired individuals to be able to play and enjoy interactive media and video games.
References


Appendix A - First iteration of the Survey
**Section 3 of 4**

| Artifact |  
| --- | --- |
| Description (optional) |  

**Were the instructions of what to do in the game clear?**
- [ ] Yes
- [x] No

**What do you think the study is about?**
Long answer text: ____________________________

**How many different areas with sounds did you find during your playthrough?**
Long answer text: ____________________________

**Where were any complications locating any sound areas?**
Long answer text: ____________________________

**How would you describe or categorize each sound area you located?**
Long answer text: ____________________________

**How would you describe immersion?**
Long answer text: ____________________________

**To you, what creates the feeling of immersion?**
Long answer text: ____________________________

**How immersed did you feel during the playthrough?**
Long answer text: ____________________________

**Was there any particular area which made you feel more immersed than others? If so please describe the sounds which made you feel so?**
Long answer text: ____________________________

**Was there any particular area which made you feel less immersed than others? If so please describe the sounds which made you feel so?**
Long answer text: ____________________________

---

**Thank you for participating!**

Thank you for taking your time to participate in this study. The study is about spatialized sound and feelings of immersion.

In this research the focus has been to understand the connection that spatialized audio could have with the feeling of being immersed, and if there are any particular types of sounds that impact that more or less than others.

**Any last thoughts or comments that you want to add?**
Long answer text: ____________________________

---

Section 4 of 4

Continue to next section →
### Appendix B - New changes of the survey

<table>
<thead>
<tr>
<th>Question</th>
<th>Long answer text</th>
</tr>
</thead>
<tbody>
<tr>
<td>How many different areas with sounds did you find in each scenario during your playthrough?</td>
<td></td>
</tr>
<tr>
<td>Where were any complications locating any sound areas?</td>
<td></td>
</tr>
<tr>
<td>How would you describe or categorize each sound area you located?</td>
<td></td>
</tr>
<tr>
<td>When listening to each area located, did the sounds make you feel anything if so describe the feelings?</td>
<td></td>
</tr>
<tr>
<td>Which scenario option out of the three did you enjoy the most?</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Describe what you enjoyed in that scenario</td>
<td></td>
</tr>
<tr>
<td>Which scenario option out of the three did you enjoy the least?</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Describe what made you not enjoy that scenario</td>
<td></td>
</tr>
<tr>
<td>Out of the three scenarios which do you think you would revisit?</td>
<td></td>
</tr>
<tr>
<td>Scenario 1</td>
<td></td>
</tr>
<tr>
<td>Scenario 2</td>
<td></td>
</tr>
<tr>
<td>Scenario 3</td>
<td></td>
</tr>
<tr>
<td>Out of the three scenarios which do you think you would not revisit?</td>
<td></td>
</tr>
<tr>
<td>Scenario 1</td>
<td></td>
</tr>
<tr>
<td>Scenario 2</td>
<td></td>
</tr>
<tr>
<td>Scenario 3</td>
<td></td>
</tr>
<tr>
<td>Was there any scenario that felt like you spent more time in over the others? If so please write which one and try to describe why you felt that way.</td>
<td></td>
</tr>
</tbody>
</table>
How was the experience of playing in complete darkness?

Long answer text

Do you think the darkness added to your experience of playing through the different scenarios?

- Yes
- No
- Maybe

Describe your thought on if either the darkness added to your experiences or if it didn’t.

Long answer text

Do you think if there where any visuals to the different scenarios your experience would be different?

- Yes
- No
- Maybe
- I don’t know

If you could assign any visuals to the different scenarios what would those visuals be and why?

Long answer text

Thank you for participating!

Thank you for taking your time to participate in this study, the study is about spatialized sound and feelings of immersion.

In this research the focus has been to understand the connection that spatialized audio could have with the feeling of being immersed, and if there are any particular types of sounds that impact that more or less than others.

The three scenarios in the artefact had different categories in mind, for the first scenario the category in mind was nature. The second scenarios sound category was machinery. And finally the third scenario was artificial sounds, made using synthesis and not recordings like the other two scenarios.

Any last thoughts or comments that you want to add?

Long answer text