A Comparison Of Two Commercially Available Alternatives For Spatializing Audio Over Headphones In A Game Setting

Per-Robin Eriksson

Ljudteknik, kandidat
2017

Luleå tekniska universitet
Institutionen för konst, kommunikation och lärande
A comparison of two commercially available alternatives for spatializing audio over headphones in a game setting

Per-Robin Eriksson
Luleå Tekniska Universitet
Epoeri-4@student.ltu.se
Perrobin.eriksson.kontakt@gmail.com
2017
Acknowledgement

I would first and foremost like to thank my supervisor Arne Nykänen for his interest and guidance in the field of spatialized sound. Without his help this study would have concluded the way it did. I would also like to thank the participants of the prestudy which helped creating a more repeatable scenario for the final experiment. Finally, I also want to thank all the participants that participated in the experiment.
Abstract
The video game industry is experiencing tremendous growth and developing games that stand out against the competition requires ever increasing production budgets. Graphics have been a selling point for video games for a long time, resulting in huge technical advancements in visuals. Audio has received more attention during the recent years, both in terms of sound-design and technical developments.

This study compared two alternatives for spatializing audio over headphones in a game setting, while also comparing them against a more traditional stereo output. A commercial game was used to provide the game setting and participants rated the three different outputs based on three attributes and described their experience using Microsoft Product Reaction Cards. No statistically significant differences were found between the three outputs in any of the six rated attributes when doing a paired t-test or a Wilcoxon Signed-Rank Test. Further analysis of frequency content, participants’ comments and Microsoft Product Reaction Cards were done to find out what sets the spatializing alternatives and stereo apart from each other.
# Table of content

1. **INTRODUCTION** .................................................................................................................................................. 5  
   1.1 The video game industry ......................................................................................................................................... 5  
   1.2 The audio in video-games ......................................................................................................................................... 5  
   1.3 How do players listen to video-games? .................................................................................................................. 6  
   1.4 Purpose and research question ................................................................................................................................. 6  

2. **BACKGROUND** ....................................................................................................................................................... 7  
   2.1 Head Related Transfer Functions and Binaural Stereo ............................................................................................. 7  
   2.2 The drawbacks of Head Related Transfer Functions ............................................................................................. 7  
   2.3 Headphones ............................................................................................................................................................... 8  
   2.4 The lack of System awareness .................................................................................................................................... 9  
   2.5 Conclusion of the Background ................................................................................................................................ 10  

3. **METHOD** ................................................................................................................................................................. 11  
   3.1 Technical equipment and setup ................................................................................................................................. 11  
   3.2 Prestudy ....................................................................................................................................................................... 12  
   3.3 Loudness matching ...................................................................................................................................................... 14  
   3.4 Game structure and settings ....................................................................................................................................... 15  
   3.5 Test subjects ................................................................................................................................................................. 16  
   3.6 Test environment ......................................................................................................................................................... 17  
   3.7 Questionnaire .............................................................................................................................................................. 17  
   3.8 Microsoft Product Reaction cards .......................................................................................................................... 18  
   3.9 Procedure ..................................................................................................................................................................... 18  

4. **RESULTS AND ANALYSIS** ..................................................................................................................................... 19  
   4.1 Rated attributes and prior experience ..................................................................................................................... 19  
   4.2 Objective data ............................................................................................................................................................. 23  
      4.2.1 Waveforms ........................................................................................................................................................... 23  
      4.2.2 Frequency Content ................................................................................................................................................ 24  
   4.3 Microsoft Product reaction cards ............................................................................................................................ 26  
      4.3.1 Word Clouds .......................................................................................................................................................... 26  
      4.3.2 Categorization ....................................................................................................................................................... 29  
   4.4 Questionnaire Comments ........................................................................................................................................... 31  

5. **DISCUSSION** ............................................................................................................................................................. 33  
   5.1 The results of the study .............................................................................................................................................. 33  
   5.2 Criticism ....................................................................................................................................................................... 34  
   5.3 Future work ................................................................................................................................................................. 35  
   5.4 Conclusion .................................................................................................................................................................... 35  

6. **REFERENCES** ............................................................................................................................................................ 36  

7. **APPENDIX** ................................................................................................................................................................. 38  
   7.1 Appendix 1 – Frequency Diagrams ........................................................................................................................... 38  
   7.2 Appendix 2 - Histogram Attribute Ratings ............................................................................................................ 41  
   7.3 Appendix 3 – Paired T-Test Two-tailed and Wilcoxon Signed-Rank Test ......................................................... 50  
   7.4 Appendix 4 - Questionnaire ....................................................................................................................................... 52  
   7.5 Appendix 5 – List of used Microsoft Product Reaction Cards .............................................................................. 53
1. Introduction

1.1 The video game industry

The video-game industry today is directly competing with the music and movie industry when it comes to generating sales, having already surpassed the music industry. In 2014 the video-games industry reached sales of almost 65 billion dollars while the music- and movie industry made around 20 and 90 billion dollars respectively. In 2010 average development cost went from around 18 million dollars to around 28 million and the development cycle lasted around 18 to 36 months. Development studios tries to outperform one another, leading to big blockbuster like video-games, requiring bigger and bigger production teams to create better graphics, artificial intelligence and game-engine, just to name a few technical aspects that are implemented into a video game. Take in account that the marketing budget for triple-A titles now are around two to three times bigger than the production budget and one can understand that the publisher needs to ship a lot of copies of the game to make a profit (Egenfeldt-Nielsen, Smith & Tosca, 2015)

1.2 The audio in video-games.

Graphics has been a big selling point in video-games often having big development teams in game-studios compared to the size of the audio-team. However, the sound-design has gained notable attention in recent years, with games in the Battlefield series, developed by the studio DICE, receiving praise for its sound-design.

Audio, just like graphics, can be used as a selling point for a game, whether it is through sound-design, soundtracks or new ways for processing the game audio for different playback alternatives. Intricate and detailed sound-design can be found in a myriad of video-games, ranging from small indie developers to triple-A studios through several different genres.

In the case of audio processing the output of the game audio, there has been solutions offered mainly by third party equipment manufacturers, but in recent years spatializing processes has been implemented into the game engines, such as in the case of Counter-Strike: Global Offensive, which implemented a binaural stereo option the 7th December 2016.
1.3 How do players listen to video-games?

In 2009 the developers at Codemaster conducted a study (Goodwin, 2009) to gather information about players available and preferred setups for listening. The results of the study showed that headphones were a very common way for players using a PC and not very common among console players’. They further mentioned that their study only addressed customers of Codemasters and that several instances in the video-game industry like equipment manufacturers, publishers and other developers could benefit from the information.

Since the study was published, smart-phones have become increasingly popular and with them, headphones have followed suite. This is something that is discussed in a paper by Olive and Welti (2012) where they mention that smaller and more portable devices are likely the cause for headphones increased popularity.

1.4 Purpose and research question

This study aims to find out what sets two spatializing processes, namely Dolby Atmos for Headphones and Spatial Sound Card, as well as stereo apart both in terms of performance, appeal and perceived experience when using them. There is a lack of research looking at commercial solutions for spatializing game audio over headphones. The solutions that do exist are often only described from a biased and marketing standpoint, which does not take real user experience into account. This study is directed towards developers of spatial audio software and consumers that use software to spatialize audio over headphones.

The questions of the study are:

- Are there any perceivable differences between Dolby Atmos for Headphones, Spatial Sound Card and stereo in a game setting and are they statistically significant?
- Can objective measurements be correlated to the players experience and ratings of Dolby Atmos for Headphones, Spatial Sound Card and stereo?
2. Background

2.1 Head Related Transfer Functions and Binaural Stereo

Head Related Transfer Functions can roughly be described as how we humans can localize sounds. Consisting of several inter-aural cues that uses physical differences of a sound which the brain interprets to determine for example if a sound comes from the left or the right.

The three cues that are the considered to be the most important are the inter-aural level difference (ILD), the inter-aural time difference (ITD) and the filtering and refraction of the head, shoulders and upper torso. Depending on a sound source’s location there will be a varying time difference between both ears, creating a difference in phase. With varying arrival times for both eardrums, a level difference will also manifest itself as the sound will be filtered by the head and possible the upper portion of the torso. Lastly, the sound will be filtered, both by the earlier mentioned body parts, but also by the shape of the pinnae and the eardrum. This is all based on a listener that has two ear, but Rumsey and McCormick (2014) mentions in their book Sound and Recording, that mono-aural cues have been proved to exist.

These functions can be measured and used to filter sounds, giving the impression that the sounds appear to be played outside of the listener’s head instead of inside the head like traditional stereo played over headphones. Depending on the amount of measurements, a sound can appear to be panned around, over and under the head. Systems using the transfer functions can be used to filter sounds as they are panned around, like for example in a game to give the impression that the action is taking place around the player. Systems using HRTF models to filter audio are known as binaural stereo, but are also commonly called 3D audio or surround sound for headphones.

2.2 The Drawbacks of Head Related Transfer Functions

The previously mentioned aural cues are different for each individual person since the shape of the ears, head and torso varies in shape and size. The brain has learned how to interpret the inter-aural cues and if presented with a different set of HRTFs the cues can be misinterpreted and locating a virtual sound source over headphones might suddenly be challenging (Rumsey & McCormick, 2014).
These differences make it hard to create one set of HRTFs that works for a big user base. HRTFs have been generalized before in both digital filters that are implemented in binaural stereo algorithms or as ears mounted on a dummy head, one such case is the Neumann KU 100 microphone.

An experiment conducted by Roginska, Santoro and Wakefeld (2010) tried to find out how user selected HRTFs from a database compared to the subject’s individual HRTFs based on three attributes, externalization (Perceiving a sound source to be outside the head), elevation discrimination (Perceiving a change in height of a sound source.), and front/back discrimination (Perceiving a sound source to be in front or at the back of the head). This experiment showed one set of HRTFs from the database was picked almost the same amount of times as the participants individual HRTF models. The discussion by Agnieszka et al (2010) suggests that more research is needed to find out which HRTF details are more important than others.

Capturing a set of HRTFs to create a HRTF model is also a costly and time-consuming task and cannot be conducted by a home user as it involves expensive equipment and anechoic rooms where the measurements are conducted.

Binaural stereo is not considered for use in loudspeakers since each setup and listening room is not the same among consumers and the stereo field might change dramatically if a listener moves their head. Headphones are considered the most optimal reproduction format for binaural stereo, since the room can be completely negated. (Rumsey & McCormick, 2014).

2.3 Headphones

Despite being such a popular listening format for players, according to the survey conducted by Codemasters (Goodwin, 2009), there are no published recommendations or standards on how controlled listening tests measuring perceived sound-quality should be conducted and the international standards that do exist have never been validated or proven with listening tests, this according to Olive and Welti (2012).
The problem lies in how to properly measure in both a subjective and an objective way as Olive and Welti brings up a lack of advancement in scientific understanding of headphones perceived sound quality relative to the rising popularity of headphones.

By conducting an experiment, Olive and Welti discovered out that the most preferred headphone pair in terms of sound quality among the participants, were the headphones that had deviated the least from a flat frequency response.

In a different experiment conducted by Boren and Roginska (2011), HRTF models taken from the IRCAM database where tested over three different headphones. The participants rated the performance of the headphones in externalization, elevation discrimination and front/back discrimination. The results from the study showed that the pair of headphones with the flattest frequency response performed better than the other two headphones.

The results from both studies shows that out of the headphones tested, the pairs with the smallest deviation from a flat frequency response were both preferred more by the participants in Olive and Welti’s experiment and performed better at binaural stereo in Boren and Roginska’s experiment. These results of these two studies are further supported by what Rumsey and McCormick (2014) have written about frequency response and its effect on sound quality, where deviations from a flat frequency response in a signal chain might over emphasize certain aspects of the sound and thus determine its quality.

2.4 The Lack of System Awareness

With varying headphone models and varying listening setups in each home, there is a possibility to ruin an already spatialized headphone mix on the way from the game to the listener’s ears. This has been brought up by Digenis in his paper Challenges of the Headphone Mix in Games (2015), Digenis brings up several possible problems which a spatialized audio option in a game might face, the lack of system awareness being one of them.

Considering how a setup might change between two individuals, where one might include a receiver with several built-in spatializing algorithms for headphones and the other one might have their headphones connected directly to the sound-card of a computer. The listener using a receiver might apply some form of algorithm that alters the sound in some way, if the game
audio is already spatialized, this algorithm might completely negate the desired effect of the spatializing. Another listener might use a software sound-card that maybe applies a room effect before converting the signal to an analogue signal.

With most third-party products acting as black-boxes, it is almost impossible to find out what goes on when and where in the signal-chain. Possible solutions to this lack of system-awareness and communication between devices could be the MPEG-H standard which has the capability of scaling to the output format used. Digenis mentions in his paper (2015) that the MPEG-H standard at the time of writing, didn’t mention any binaural stereo support.

With devices becoming smarter for every generation, headphones due to their analogue nature can be perceived as a “dumb device” in an eco-system where devices can recognize each other and their capabilities. Digenis claims that if headphones would make transposition to smart-headphones, the interface used for connecting these smart-headphones to a system could include information about the capabilities of the headphones along with data to avoid the problems mentioned earlier in his paper (Digenis, 2015).

2.5 Conclusion of the Background

There is not a lack of solutions for spatializing audio, whether it is for loudspeakers or headphones, but the lack of communication between software and hardware can ruin an already spatialized signal. An already ruined signal that reaches a pair of headphones does not have an optimal frequency response for spatialized audio, might completely negate the intended function of a spatializing process and distort the sound-design of the game (Digenis, 2015).

Making it easy for the user-base to setup a spatialized playback should be a priority among third-party developers. Another option is to directly implement them into the game-engines and making the option easy to switch on and off, such as in the case of Counter Strike: Global Offensive (2012).
3. Method

To test the two different techniques for spatializing audio, an experiment had to be designed and conducted. Overwatch (Blizzard Entertainment, 2017) was picked to provide the game setting for its custom game mode. The custom game-mode made it possible to create a repeatable scenario for use in the experiment. Overwatch was also chosen for its intricate sound design which utilizes several systems to process to sound, one of these is an occlusion system that filters sounds depending on how the sound from a source travels to the player character, this is explained in detail at the Game Developers Conference by Lawlor and Neumann (2016).

Dolby Atmos for Headphones is a spatializing process developed by Dolby Laboratories which uses metadata to position sounds in a virtual sphere. It is at the time of writing only implemented in the game Overwatch, developed by Blizzard Entertainment (2017).

Spatial Sound Card is a third-party software developed by New Audio Technology. The software uses a multi-channel output from other programs to create a binaural representation of a surround-speaker setup. The software includes several impulse responses for different listening rooms which can be selected by user.

3.1 Technical Equipment and setup

The following equipment was used for both the prestudy and the final experiment. The setup of the used equipment can be seen in figure 1.

**Hardware**

- **Headphones:** Beyerdynamic DT770 Pro
- **Audio Interface:** Universal Audio Apollo Twin Duo MK I
- **Computer Screen:** BenQ 2420HD
- **Computers:**
  - Macbook Pro Retina late 2013: used for recording the participants’ sessions and making volume adjustments.
  - PC: AMD FX 6100 CPU, nVidia GT 660, 8Gb RAM: used for running the game Overwatch (Blizzard Entertainment, 2017).
- **Gamepad:** Microsoft Xbox One Controller
- **Keyboard:** Cooler Master Quick Fire TK
- **Mouse:** Microsoft IntelliMouse 3.0
Software
Video game: Overwatch (Blizzard Entertainment, 2017)
DAW: Cockos Reaper v5.311/64
Frequency analysis: Izotope RX 5 Audio Editor v5.02.431
Loudness analysis: Izotope insight v1.05.481.OSX64.Intel.VST
SWS/S&M Extension, a Reaper plugin. v2.8.8
Other Universal Audio Console

Figure 1 - Setup of technical equipment used during the prestudy and the real experiment. The Mac was used to record the participants' sessions. The PC was used to run the game Overwatch (Blizzard Entertainment, 2017).

3.2 Prestudy
A prestudy was conducted to determine which listening room to be used in Spatial Sound Card, determine a playback level and to gather information on the settings in the game. Three people participated in the prestudy, all studying the last year of audio engineering at Luleå University of Technology. The rooms Shanghai and New York in Spatial Sound Card were chosen by how they were perceived by the author in terms of sound-quality. This was to limit the amount of rooms to be tested in the prestudy so that it would not take too much time to complete.

Procedure of the prestudy:
- Participants were given instructions to escort a payload and to listen to the sound of the game.
- They were told to set a comfortable listening volume. The participants could change this volume until they felt it was comfortable.
The participants played one match with either New York or Shanghai as a room in Spatial Sound Card.

When the match was finished, they filled in a questionnaire while the room was changed.

The participants then played another match with the other room being used in Spatial Sound Card.

When the second match was finished, they filled in another questionnaire.

When finished, participants were interviewed about the settings used in the game and their general experience during the two matches.

Each participant rated the sound of the two matches based on three attributes, these being frequency response, naturalness and appeal. These attributes were used to see how the frequency response of the rooms were perceived in comparison to each other and if they introduced any phase irregularities and if they generally appealed to the participants.

The number of participants were too few to conduct a paired t-test.

The room Shanghai received higher ratings on the questionnaire than the room New York and when interviewed participants described Shanghai as the more natural of the two rooms.

**Spatial Sound Card Settings.**

The following settings were used in Spatial Sound Card for the real experiment. The built-in equalizer that Spatial Sound Card uses to compensate for the different frequency responses among headphones was not used. A screenshot of the settings can be seen in figure 2.

Volume settings were kept to factory default for the version used during the experiment:

**Version**

Core: 17050152 64-bit (Oct 26 2016 12:27:12)

SoundCard: 17050152 64-bit (Oct 26 2016 12:29:00)

**General settings**

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker setup</td>
<td>5.1</td>
</tr>
<tr>
<td>Latency target</td>
<td>5.3ms</td>
</tr>
<tr>
<td>Headphone EQ</td>
<td>No</td>
</tr>
<tr>
<td>Room envelope</td>
<td>Short</td>
</tr>
<tr>
<td>Headphone Surround</td>
<td>Yes</td>
</tr>
<tr>
<td>Room</td>
<td>Shanghai</td>
</tr>
</tbody>
</table>
Volume settings

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Master volume:</td>
<td>100</td>
</tr>
<tr>
<td>Left Speaker:</td>
<td>-10dB</td>
</tr>
<tr>
<td>Right Speaker:</td>
<td>-10dB</td>
</tr>
<tr>
<td>Center Speaker:</td>
<td>-10dB</td>
</tr>
<tr>
<td>LFE:</td>
<td>0dB</td>
</tr>
<tr>
<td>Left Surround Speaker:</td>
<td>-10dB</td>
</tr>
<tr>
<td>Right Surround Speaker:</td>
<td>-10dB</td>
</tr>
</tbody>
</table>

Figure 2 - Settings used in Spatial Sound Card during the experiment.

3.3 Loudness matching

To reduce the impact of volume differences between stereo, Dolby Atmos for Headphones and Spatial Sound Card, they all had to be measured in a similar scenario to make the measured material as similar as possible.

The program material used for the loudness analysis and matching consisted of the character Bastion in sentry mode emptying a full magazine of his automatic weapon at the same target, at the same location for stereo, Dolby Atmos for Headphones and Spatial Sound Card. The audio was recorded in the DAW Reaper and edited to start and stop at the same time for all three audio clips.

Each audio clip was then analysed using the SWS extension for reaper. Izotope Insight was used to control that the program loudness value based on the standard EBU R128 was equal to the value given by the SWS extension. Stereo and Dolby Atmos for Headphones were
normalized to the loudness value of Spatial Sound Card. The normalization was done with the SWS extension. The attenuation values for stereo and Dolby Atmos were then implemented on the fader for the S/PDIF input in Universal Audio Console. The results of the loudness measurement can be seen in table 1.

The SWS extension is a collection of functions and features that integrates into the DAW Reaper. It is an open-source project and some of the features included is the loudness analyser and normaliser previously mentioned.

Table 1 – Results of the loudness measurement in Reaper using the SWS extension. The table shows the measured program loudness, target program loudness and the corresponding volume adjustments from the SWS extension in Reaper.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial Sound Card</td>
<td>-23.1</td>
<td>-23.1</td>
<td>0</td>
</tr>
<tr>
<td>Dolby Atmos for Headphones</td>
<td>-19.6</td>
<td>-23.1</td>
<td>-3.5</td>
</tr>
<tr>
<td>Stereo</td>
<td>-17.6</td>
<td>-23.1</td>
<td>-5.5</td>
</tr>
</tbody>
</table>

3.4 Game structure and settings

The structure, objective and settings in Overwatch were finalized after the feedback collected from the presudy.

**Overwatch – Settings**

- Map: Watchpoint: Gibraltar
- Objective: Escort
- Player Character: Soldier 76
- AI Characters: Roadhog (Player Team), Lucio (Player Team), McCree (AI Team), Reaper (AI Team), Sombra (AI Team)
- AI Difficulty: Hard (Player Team), Easy (AI Team)
Watchpoint: Gibraltar was chosen as the map and setting in which the participants would complete their objective due to its linearity which made it hard for players to get lost. On Watchpoint: Gibraltar the participants team must escort a payload to its end-destination while defending it from the opposing team, reaching checkpoints along the way.

Soldier 76 was picked to be the only playable character in the experiment due to the character’s similarity to characters in other first-person-shooter games. The characters that were controlled by the computer added support to player in the form of Roadhog and Lucio and opposition in the form of Reaper, McCree and Sombra. All the computer controlled characters use distinct sound-cues for their abilities, for example, Sombra that can turn invisible but through her foley-sounds still be located.

### 3.5 Test subjects

The subjects which participated in the experiment were aged between 19-28 years, with 17 participants being male. Participants listening experience varied, with most participants studying audio engineering at Luleå University of Technology.

The condition for conducting the experiment was that the participants had to have some form of experience playing video games. The reasoning for this was to see how the spatializing processes alternatives would perform for a wide range of people that play video games, as “3D-sound” is not necessarily only directed at dedicated gamers.

![Figure 3 – Pie chart displaying the participants prior experience with Overwatch](image-url)
3.6 Test environment

The experiment was conducted over two occasions and took place in G105, a quiet room at Luleå University of Technology in Piteå School of Music in 2017. The first occasion lasted from the 28th of February to the 2nd of March. The second session took place on the 10th of March.

3.7 Questionnaire

To gather quantitative data, a questionnaire was designed and written in Swedish. Six attributes were chosen for evaluation and translated to Swedish., these were externalization (Externalisering), presence (Närvaro), representation (Representation), sense of space (Känsla av rymd), localization (Lokalisering) and appeal (Tilltalande). The attribute externalization is based on a prior study conducted by Roginska et al (2010), where user selected HRTF models from the IRCAM database were compared to the measured models of the test participants.

Presence, localization and sense of space are based on three attributes used for describing music in 5.1 and originates from a study written by Jan Berg and Francis Rumsey (2003). These were picked for their fitting descriptions and the suggestion from Berg and Rumsey to be used outside of music in 5.1 surround as in the following quote: “To find the applicability of the method for product evaluation tasks, the comparison of different system components can be made, e.g. loudspeakers, spatial enhancers, etc.” (Berg & Rumsey, 2003, p.13).

Sense of space is also based on the attribute feeling of space from the study Spatial Sound Attributes – Development of a Common Lexicon (Zacharov & Pedersen, 2015) where the authors have gathered definitions and attributes from several studies to create common definitions for spatial attributes.

Representation was picked as an attribute to see if participants experienced the filtering of the spatializing processes to be destructive on the sounds in the game. The last attribute appeal was picked to see how appealing or enjoyable participants experienced the sound of their three sessions. The focus of the ratings are not how well the different outputs performed against each other, but rather how participants experienced them. The ratings were done retrospectively after each session.
The participants answered the questionnaire anonymously and the only information they had to give out were their experience playing video games and if they had prior experience of Overwatch.

3.8 Microsoft Product Reaction Cards

Microsoft Product Reaction Cards is a list of words provided by Microsoft Corporation and is used to evaluate user experience of products. The words can be used to generate word clouds that can easily be visually analysed.

Microsoft Product Reaction Cards were used in the experiment to see how participants experienced Dolby Atmos for Headphones, Spatial Sound Card and stereo with same words available. The use of Microsoft Product Reaction Cards in this study is largely inspired by the bachelor thesis written by Christoffer Björnram (2015). In his bachelor thesis, Björnram limited the number of cards that participants could pick, resulting in each participant choosing the five most critical cards they felt described their experience. The full list of words provided by Microsoft was not used and the words not deemed relevant by the author were sorted out. The full list of words used can be found in appendix 5.

3.9 Procedure

Participants given instructions on the structure of the experiment, while also given time for questions. Instructions were then given about their objective and participants could chose the controls they were most comfortable with. The sessions were planned using a counterbalanced Latin square design to assign the playback order for each participant.

The procedure of the main experiment was as follows

- Participants were given instructions that they were going to play three matches in the game Overwatch and that they would fill in a questionnaire rating the sound of the game.
- The questionnaire was shown to the participant and the use of Microsoft Product Reaction Cards was explained to the participants.
- Participants were asked what form of control they preferred and that they could not change any settings in the game during the match. Control options could be changed if they mentioned it to the test conductor after a match.
• Participants were given information about their objective in the game. Written instructions were available in case the subjects needed clarification.

• Participants played through a match and moved to a different table to fill out the questionnaire and select their Microsoft Product Reaction Cards, meanwhile settings for the next session was adjusted by the test conductor.

• The procedure repeated itself until the participants had played and rated all their three sessions.

The results of participant 1 and 11 had to be declared invalid due to incorrect settings set by the test conductor during one or more of their sessions and was not used at all. To maintain the same number of playback scenarios, participants 19 and 20 performed the test with the same playback order as participants 1 and 11. The results of participant 19 replaced the results of participant 1 and participant 20 replaced the results of participant 11.

4. Results and analysis

4.1 Rated Attributes and Prior Experience

This sections deals with the results and the analysis of the quantitative data collected from the questionnaire.

![Figure 4 – Box plot for the attribute externalization](image1)

![Figure 5 – Box plot for the attribute presence](image2)
Figure 6 – Box plot for the attribute representation

Figure 7 – Box plot for the attribute sense of space

Figure 8 – Box plot for the attribute localization

Figure 9 – Box plot for the attribute appeal
When looking at the boxplots, namely figure 4, 5, 6, 7, 8 and 9, for the questionnaire results, there are some minor differences in the data spread, which were to some degree expected due to the nature of binaural stereo and HRTF models. When paired t-tests were conducted, no statistically significant differences were found between stereo, spatial soundcard and Dolby Atmos for Headphones in any of the six attributes. A table of the calculated t-values and the comparison against a critical t-value can be found in appendix 3. A Wilcoxon Signed-Rank Test was also done to see if the differences that do exist are significant despite the lack of normal distribution in the data, this table can also be found in appendix 3. No statistically significant differences were found when using the Wilcoxon Signed-Rank Test.

Analysing the data spread found in the histograms of the participants’ ratings, which can be found in appendix 2, reveals a lack of normal distribution of the data. There are signs of normal distribution in the rating of representation for stereo and the attribute appeal for Dolby Atmos for Headphones. The number of outliers visually represented in the boxplots at page 20 might be what prevents the differences from being statistically significant. However, the outliers together with the spread in the data might be caused by the participants’ preferences in game audio and HRTF models that does not work as intended for everyone.

It is however notable that Spatial Sound Card have a big spread in the ratings for all the attributes, often using most the scale, while Dolby Atmos for Headphones and stereo varies more in their spread, with stereo having smaller differences in its ratings than both Spatial Sound Card and Dolby Atmos.

If more participants took part in the experiment some of the differences could possibly be statistically significant and in the case of the data spread, it would probably display more of a normal distribution than the data collected in this experiment.

A two-sample independent T-Test was conducted to see if there if a statistically significant difference could be found between the participants that had prior experience with Overwatch and the participants that did not. The t-test tested each group and its ratings for each of the outputs in all six of the attributes and the outcome can be seen in table 2.
Table 2 – Results of the two-sample independent t-test, testing participants with prior experience with Overwatch against the participants that did not have any prior experience.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Output</th>
<th>T-value</th>
<th>P-value</th>
<th>Alpha</th>
<th>Statistically Significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Externalization</td>
<td>Atmos</td>
<td>-1.807</td>
<td>0.090</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Externalization</td>
<td>Stereo</td>
<td>-2.353</td>
<td>0.032</td>
<td>0.05</td>
<td>Yes</td>
</tr>
<tr>
<td>Externalization</td>
<td>Spatial Sound Card</td>
<td>-0.289</td>
<td>0.776</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Presence</td>
<td>Atmos</td>
<td>-1.038</td>
<td>0.315</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Presence</td>
<td>Stereo</td>
<td>-0.439</td>
<td>0.667</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Presence</td>
<td>Spatial Sound Card</td>
<td>-0.642</td>
<td>0.530</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Sense of Space</td>
<td>Atmos</td>
<td>-0.808</td>
<td>0.315</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Sense of Space</td>
<td>Stereo</td>
<td>-0.240</td>
<td>0.814</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Sense of Space</td>
<td>Spatial Sound Card</td>
<td>-1.048</td>
<td>-0.310</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Representation</td>
<td>Atmos</td>
<td>-1.237</td>
<td>0.234</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Representation</td>
<td>Stereo</td>
<td>0.317</td>
<td>0.756</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Representation</td>
<td>Spatial Sound Card</td>
<td>-1.478</td>
<td>0.159</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Localization</td>
<td>Atmos</td>
<td>-0.481</td>
<td>0.637</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Localization</td>
<td>Stereo</td>
<td>-1.064</td>
<td>0.303</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Localization</td>
<td>Spatial Sound Card</td>
<td>0.725</td>
<td>0.479</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Appeal</td>
<td>Atmos</td>
<td>0.023</td>
<td>0.982</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Appeal</td>
<td>Stereo</td>
<td>0.098</td>
<td>0.923</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Appeal</td>
<td>Spatial Sound Card</td>
<td>-0.887</td>
<td>0.388</td>
<td>0.05</td>
<td>No</td>
</tr>
</tbody>
</table>

As seen in table 2, one statistically significant difference can be found in the attribute externalization. This result shows that the participants without any prior experience with Overwatch (Blizzard Entertainment, 2017) experienced the stereo output as coming more outside of their head than the participants that did have prior experience.

Outside of the statistically significant difference that have been found, there is one case that comes close to being statistically significant, this one being Dolby Atmos for Headphones in the attribute externalization. If more participants had participated both with and without experience with Overwatch, this difference might prove statistically significant.
4.2 Objective data

This section will analyse the waveforms and the frequency content of stereo, Dolby Atmos for Headphones and Spatial Sound Card as to find data that could be used as evidence for participants’ comments, which are analysed later in this study.

4.2.1 Waveforms

![Waveform Comparison](image)

Figure 10 – The waveforms recorded for a participant’s three sessions. The audio clips have been edited with a common beginning and end, which showcases the difference in time between the sessions.

In figure 10, the waveforms are screenshots taken in Reaper, the sessions length is preserved and no gain alteration have been made.

When looking at figure 10, the waveform of stereo the strongest portions are when the participant was engaged in a firefight or close to one. If compared to the others, it seems that stereo appears to be compressed in some form. Dolby Atmos for Headphones and Spatial Sound Card has a lot more varying levels when the participant either fired their gun or was close to a gun firing. Even when the participant did not experience any action, the levels in the stereo session were higher in amplitude.
The compression of the stereo output can be the cause to why participant 19 experienced it as if sounds were “in your face” in the comment on page 31.

As a conclusion of the waveform analysis the three different outputs showcases a varying degree of dynamic range with stereo appearing to be more compressed than Dolby Atmos and Spatial Sound Card.

4.2.2 Frequency Content

When analysing the frequency content of each processing technique, each session had to be divided with set start and stop markers. Considering that each session varied in length and scenario, differences were expected. Each session was volume adjusted with the same values used during the experiment and exported as WAV files. These were then imported into Izotope RX5 to generate a spectrogram and a frequency diagram of the sessions frequency content.

Figure 11 – FFT frequency diagram showcasing the frequency response of the three sessions of participant 16. The diagram also shows the frequency response for the left and right channel for each output. The grey curve that is deviating from the other curves is the left channel for Spatial Sound Card.

Figure 11 displays the frequency content of participant 16’s sessions. The attenuation values used for loudness matching were also applied to the audio files before generating the frequency diagrams in Izotope RX 5. A bigger frequency diagram can be found in appendix 1 at the page 38.
There are small variations between the frequency response of Dolby Atmos for Headphones to that of stereo. Other than the variance in amplitude in the frequencies from 30 Hz to 600 Hz another difference between Dolby Atmos and Stereo can be found between 5 kHz to 15 kHz, where Dolby Atmos has some small attenuations that deviates from the Stereo signal.

Spatial Sound Card shows similar frequency curves as stereo and Dolby Atmos but with much greater amplitude differences like in the area between 100 Hz and 300 Hz. But beyond that, the frequency response of Spatial Sound Card is vastly different than that of the other two, also showcasing big differences between the left and right output channels. Most notable is the big bump between 4 kHz and 10 kHz and the earlier cut-off around 30 Hz and 20 kHz compared to that of Stereo and Dolby Atmos.

As seen in the frequency diagram, there are both minor and big differences between the sessions frequency responses. The time differences between the sessions might exaggerate or diminish these differences, but the diagram show a treble difference that participant 16 might have experienced when using Spatial Sound Card.

The conclusion of the analysis of the objective data shows that Dolby Atmos for Headphones has a frequency response very close to that the frequency response of stereo. Spatial Sound Card has a very varying frequency response that differs a lot from that of Dolby Atmos for Headphones and stereo, the cause of this might be the room impulse that is used. Both Dolby Atmos for Headphones and Spatial Sound Card display less compression than stereo in their waveforms.
4.3 Microsoft Product Reaction Cards

The results of the Microsoft Product Reaction Cards have been presented in two ways. One with Word Clouds and the other through categorization of the words used in the experiment.

4.3.1 Word Clouds

The word clouds are created using all the words picked by the participants to describe their experience with the sound of the sessions.

Font size and contrast is determined by the frequency of the words picked by the participants. The frequency of the words is included so that the visual difference between Word-Clouds can be interpreted in a more representative way. The word clouds are meant to give a visual representation of how the participants as a group experienced Dolby Atmos for Headphones, Spatial Sound Card and the stereo output of the game-audio from Overwatch.

Figure 12 – Word Cloud displaying the words describing Stereo.
Figure 13 – Word Cloud displaying the words describing Dolby Atmos for Headphones.
When visually analysing the word-clouds it can be noted that a few words stand out in the word-clouds for Dolby Atmos for Headphones in figure 13 and stereo in figure 12. For stereo the words energetic, exciting, clear, high-quality and entertaining stand out the most. For Dolby Atmos for Headphones the words energetic, exciting, clear and stimulating stand out the most while more words have the same frequency in the word-cloud of Spatial Sound Card in figure 14. It easier to distinguish words in the word-clouds of Dolby Atmos for Headphones and stereo due to difference in frequency of the words picked to describe them. The word-cloud of Spatial Sound Card have words like fun, high-quality and entertaining have the same frequency as the words simplistic and unattractive. This may be correlated to the big spread in the data for Spatial Sound Card as shown in figures 4, 5, 6, 7 and 8 on page 19 and 20.
4.3.2 Categorization

To further analyse the Microsoft product reaction cards, the words were categorized based on what they described and in the context of the participants’ ratings in the questionnaires. This is to more easily clarify what words roughly describes different parts of the participants’ experiences. The result of the categorization can be seen in table 3.

Table 3 – Categories for the Microsoft Product Reaction Cards. The words are categorized based on the words definitions and how they may relate to each other in describing an experience.

<table>
<thead>
<tr>
<th>Category 1</th>
<th>Category 2</th>
<th>Category 3</th>
<th>Category 4</th>
<th>Category 5</th>
<th>Category 6</th>
<th>Category 7</th>
<th>Category 8</th>
<th>Category 9</th>
<th>Category 10</th>
<th>Category 11</th>
<th>Category 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convenient</td>
<td>Simpatic</td>
<td>Simplistic</td>
<td>Connected</td>
<td>Ineffective</td>
<td>High-Quality</td>
<td>Creative</td>
<td>Compelling</td>
<td>Annoying</td>
<td>Calm</td>
<td>Appealing</td>
<td>Boring</td>
</tr>
<tr>
<td>Essential</td>
<td>Sterile</td>
<td>Consistent</td>
<td>Poor-Quality</td>
<td>Impressive</td>
<td>Fresh</td>
<td>Empowering</td>
<td>Confusing</td>
<td>Clean</td>
<td>Attractive</td>
<td>Dull</td>
<td>Inviting</td>
</tr>
<tr>
<td>Helpful</td>
<td>Unresolved</td>
<td>Effective</td>
<td>Unattractive</td>
<td>Professional</td>
<td>Inspiring</td>
<td>Energetic</td>
<td>Difficult</td>
<td>Clear</td>
<td>Entertaining</td>
<td>Old</td>
<td>Ordinary</td>
</tr>
<tr>
<td>Relevant</td>
<td>Efficient</td>
<td>Undesirable</td>
<td>Intuitive</td>
<td>Engaging</td>
<td>Disconnected</td>
<td>Comfortable</td>
<td>Ran</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reliable</td>
<td>Integrated</td>
<td>Motivating</td>
<td>Exciting</td>
<td>Disruptive</td>
<td>Comprehensive</td>
<td>Satisfying</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time-saving</td>
<td>Responsive</td>
<td>Unconventional</td>
<td>Powerful</td>
<td>Intimidating</td>
<td>Organized</td>
<td>Stimulating</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trustworthy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Useful</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valuable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 15 – Histogram displaying the frequency of words in categories.
Before the categorized words are analysed, a few concerns must be addressed. The frequency of the words in the categories are directly affected by the amount of words that the participants decided to pick. This number varied both for individual participants when they described the spatializing processes and the stereo output and it varied between participants all together. Some participants used a scarce number of words and some used more words to describe their experience. This can both be interpreted as invalid results or a direct result of how the participants experienced the sound of each session.

In figure 15, the histogram displays an interesting distribution of words in the categories. In Category 1 stereo is clearly standing out compared to the two spatializing processes, this can also be seen in category 10. In short, stereo can be described as more reliable and energetic than Dolby Atmos for Headphones and Spatial Sound Card. The compression of the stereo output as seen in figure 10 might be the cause why stereo was experienced as more energetic and powerful, just to name a few of the words in category 10. As of category 1 a possible cause might be that the participants were more used to traditional stereo.

Dolby Atmos for Headphones have the most notable frequency of words in categories 6 and 8, where it is described both as the most creative and fresh out of the three, but is at the same time the most annoying, confusing and difficult. The reason for this might simply be that the HRTF model that is presumably being used in Dolby Atmos for Headphones, did not work for all the participants or that the participants simply were not used to the spatial processing.

The reason why Spatial Sound Card stands out in category 4 which includes the words poor-quality, ineffective, unattractive and undesirable, might be because of the frequency response shown in figure 11. Participants might have reacted to the large level increase for frequencies between 4 kHz and 10 kHz which can be correlated to the how the frequency response affect sound-quality (Rumsey & McCormick, 2014).
4.4 Questionnaire comments

The comments written down by the participants in the questionnaire have been translated from Swedish to English. Some comments could not be directly translated to English and have therefore been translated as closely as possible and arranged in table 4.

Table 4 – Participants comments. Writing a comment was entirely up to the participant.

<table>
<thead>
<tr>
<th>Doly Atmos for Headphones</th>
<th>Stereo</th>
<th>Spatial Sound Card</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very reminiscent of the first alternative.</td>
<td>Participant 9</td>
<td>It sounded more distanced this time.</td>
</tr>
<tr>
<td>Participant 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some sounds were perceived as coming from the back and other from the front (A little bit).</td>
<td>Participant 18</td>
<td>Felt boring. &quot;Mono&quot;-feeling even though it wasn't. Lots of high frequencies? (Lacking in bass)?</td>
</tr>
<tr>
<td>Participant 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most compelling of all.</td>
<td>Participant 19</td>
<td>Good directional precision</td>
</tr>
<tr>
<td></td>
<td>All the sounds felt very here and in your face.</td>
<td></td>
</tr>
<tr>
<td>Participant 14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The music did a lot. The characters were a bit irritating.</td>
<td>Participant 18</td>
<td>Gets a little exciting without the music. Can be a contrast thing.</td>
</tr>
<tr>
<td>Participant 17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adventurous and dramatic.</td>
<td>Participant 19</td>
<td>Good directional precision</td>
</tr>
<tr>
<td></td>
<td>A lot of treble.</td>
<td></td>
</tr>
<tr>
<td>Participant 19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Especially hard to differentiate between sounds coming from the front or the back. A lot of the sounds had the same volume, was perceived as having more bass.</td>
<td></td>
<td>Participant 14</td>
</tr>
<tr>
<td></td>
<td>Found more details that made the environment feel more natural. For example, birds towards the water in the beginning.</td>
<td></td>
</tr>
</tbody>
</table>

Some of the comments in table 4 seems to be directed towards the sound-design of the game rather than the overall experience with the sound, as in the case of participants 17, 14 and 4. The cause of why participant 14 found more details in the sound design might be that audio
cues are played at random or that the participant decided to go to a new place on the map not visited during the two other sessions.

Participant 4 who described stereo as enveloping, also rated stereo higher than Dolby Atmos and Spatial Sound Card in the attribute of presence. The same applies for the attribute localization for participant 4. This participant had prior experience with game and might be used to playing with stereo set as audio output which might explain the rating of the attributes.

There is evidence for the comments mentioning the frequency response of Dolby Atmos for Headphones and Spatial Sound Card. Participant 9 and 16 mention in their comments that Spatial Sound Card was perceived to have a lot of treble and a possible lack of bass, both of which can be seen in the frequency curve seen in figure 11. In the case of Dolby Atmos for Headphones, it was described by participant 19 to have more bass which also can be seen in figure 11. Participant 2’s comment about “Good punch” might also correlate with the frequency curve of Dolby Atmos for Headphones. The frequency diagrams for participants 2, 9, 16 and 19 can be found in appendix 1.

Another interesting part of participant 9’s comment is the description of Spatial Sound Card being perceived as a bit “mono”. This might simply be a description on the effects of the HRTF model used by Spatial Sound Card.
5. Discussion

5.1 The Results of the Study

The implementation of Dolby Atmos for Headphones in a game engine marks a form of paradigm shift, where spatializing processes like Spatial Sound Card is no longer needed for spatial audio in a game. However, it’s use is extremely limited as it can only be used for that one game that has implemented it. If other games implement their own spatializing processes, there can be huge differences between them. If a player has spent time with one process and become comfortable with it, a new one might be perceived as confusing or not appear as appealing. Spatial Sound Card can however be used for other applications, such as listening to the 5.1 audio of a movie instead of the stereo track. It also gives users more options to customize to their taste. Despite maybe not providing binaural stereo that works a 100 %, it can give users an experience that stereo might not be able to provide.

Dolby Atmos for Headphones also marks a shift from the use of the multichannel output from a game to spatialize a virtual surround setup, to using metadata for positioning sounds in both a horizontal and vertical axis. This is favourable in competitive first-person-shooter games since enemy players or AI might be positioned above the player.

The results of the experiment were somewhat expected since the drawbacks of HRTF models is that they don’t always work for all individuals, as explained in the background section. The results do not speak for a consensus that one is better than the other, but showcases how physical differences and subjective preferences can differ when it comes to spatialized audio or traditional stereo over headphones.

Analysing and comparing each comment written by the participants with their ratings on the questionnaire while also looking at their choice of Microsoft Product Reaction Cards was not done in this study. The reason for this was that the time it would take to do the analysis was not available during this study. Looking at individual results in detail would be a suggestion for future study where fewer participants can be analysed to gather as much data as possible from each individual participant.

Also, worth mentioning is the invalid results of participants 1 and 11, which were caused by the very problem that was mentioned in the background section, the lack of system awareness.
and the human factor of the author resulted in a ruined spatialized signal during two sessions. The problem is very real and despite doing everything to counteract it, it showed up during the experiment. Even if only one notable statistically significant results was found in this study, a repeatable and controlled experiment was constructed and executed, evaluating third-party products and collecting the participants’ ratings and experiences.

5.2 Criticism
The scope and aim of this study ended up a bit too big than what was originally considered and the amount of results and data speaks for that, not to mention the time it took to do the experiment. A bigger number of participants should have been considered to find any statistically significant differences and the number of attributes in the questionnaire could have been reduced.

Conducting an experiment evaluating third-party products proved itself to be a difficult matter, where every part before the PC output was a black box. Only a few parameters were known, which made measurements time-consuming. Despite trying to observe the several volume stages, the measurements done for this experiment can probably differ in another setup, as the monitoring situation will vary greatly between setups.

The use of Microsoft Product Reaction Cards should have been executed in a different way, one way would be by limiting the amount of words that participants could pick to describe the sound and interview them afterwards on why they picked their specific words. This however would have made the experiment unbearable long for the participants, not to mention the test-conductor who had to be present during each test.

The results of the Microsoft product reaction cards can be criticised since the amount of words vary for each playback alternative which can be seen in figure 14. The amount of words for each session can also be interpreted because of how the participants experienced the playback alternatives, but the difference in the amount of words should be considered when looking at the results of the Microsoft Product Reaction Cards.

Finally, more time could have been spent on determining the attributes which were used in the questionnaire, mainly due to how they could be interpreted. Participants could ask for a more
detailed explanation on the different attributes when they were filling in the questionnaire. A reduced number of attributes could also cut down the time it would take to do this experiment.

5.3 Future Work

If a similar study will be conducted in the future, it might be wise to consider the scope of the study and only focus on a very small part of the spatializing processes. One part could be just about preference, another one could be performance and a final suggestion would be to just make measurements in preparation for a study. This study should be a starting off point where every aspect can be improved and developed to test different aspects of spatial audio in games. How prior experience with a game might affect the perception and experience of a sound output might be interesting to look at in the future.

Microsoft Product Reaction Cards can be presented in so many ways that it would be a good idea to create some form of set standards for different kinds of areas. In this study the cards were presented in the form of word clouds and as a histogram with the words categorized. It would certainly make it easier to compare the use of Microsoft Product Reaction Cards between studies.

5.4 Conclusion

Several conclusions can be made after analysing the results of the experiment and these are:

- That no statistical significance could be found between stereo, Dolby Atmos for Headphones and Spatial Sound Card when rating six different attributes.
- Individual results vary greatly both in terms of HRTF performance and overall appeal, based on the questionnaire ratings of individual participants.
- There appears to be a dynamic difference between the stereo output and the spatial outputs when visually analysing the recorded waveforms.
- The frequency response is different between the three different outputs.
- Objective measurements can be correlated to the experience and ratings of the players.
- There is one statistically significant difference between participants with prior experience with Overwatch and the participants without prior experience.

There may still be a long way to go before spatial processing of game audio works for everyone and the results of this study seems to be that it does not matter if binaural stereo or traditional stereo is used in a game setting.
6. References


7. Appendix

7.1 Appendix 1 – Frequency Diagrams

Figure 16 – Combined frequency diagram showcasing the frequency response of Dolby Atmos for Headphones, Spatial Sound Card and stereo for participant 16.
Figure 17 – Frequency response for participant 2 when using Dolby Atmos for Headphones.

Figure 18 – Frequency response for participant 19 when using Dolby Atmos for Headphones.
Figure 19 – Frequency response for participant 9 when using Spatial Sound Card.

Figure 20 – Frequency response for participant 16 when using Spatial Sound Card.
7.2 Appendix 2 - Histogram Attribute Ratings

**Externalization**
Dolby Atmos for Headphones

**Externalization**
Stereo
**Externalization**

Spatial Sound Card

**Presence**

Dolby Atmos for Headphones
Localization
Atmos

Localization
Stereo
### 7.3 Appendix 3 – Paired T-Test Two-tailed and Wilcoxon Signed-Rank Test

Table 5 – Table showing the results of the paired T-test. A significance level of 0.05 was used.

<table>
<thead>
<tr>
<th>Pair ( (x_1 - x_2 = d) )</th>
<th>Attribute</th>
<th>Calculated T-value</th>
<th>Critical T-value</th>
<th>Statistically Significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dolby Atmos for Headphones</td>
<td>Stereo</td>
<td>Externalization</td>
<td>1,073</td>
<td>2,110</td>
</tr>
<tr>
<td>Stereo</td>
<td>Spatial Sound Card</td>
<td>Externalization</td>
<td>-1,898</td>
<td>2,110</td>
</tr>
<tr>
<td>Spatial Sound Card</td>
<td>Dolby Atmos for Headphones</td>
<td>Externalization</td>
<td>0,701</td>
<td>2,110</td>
</tr>
<tr>
<td>Dolby Atmos for Headphones</td>
<td>Stereo</td>
<td>Presence</td>
<td>0,638</td>
<td>2,110</td>
</tr>
<tr>
<td>Stereo</td>
<td>Spatial Sound Card</td>
<td>Presence</td>
<td>0,713</td>
<td>2,110</td>
</tr>
<tr>
<td>Spatial Sound Card</td>
<td>Dolby Atmos for Headphones</td>
<td>Presence</td>
<td>-0,829</td>
<td>2,110</td>
</tr>
<tr>
<td>Dolby Atmos for Headphones</td>
<td>Stereo</td>
<td>Sense of Space</td>
<td>-0,218</td>
<td>2,110</td>
</tr>
<tr>
<td>Stereo</td>
<td>Spatial Sound Card</td>
<td>Sense of Space</td>
<td>1,516</td>
<td>2,110</td>
</tr>
<tr>
<td>Spatial Sound Card</td>
<td>Dolby Atmos for Headphones</td>
<td>Sense of Space</td>
<td>-1,511</td>
<td>2,110</td>
</tr>
<tr>
<td>Dolby Atmos for Headphones</td>
<td>Stereo</td>
<td>Representation</td>
<td>-0,695</td>
<td>2,110</td>
</tr>
<tr>
<td>Stereo</td>
<td>Spatial Sound Card</td>
<td>Representation</td>
<td>1,330</td>
<td>2,110</td>
</tr>
<tr>
<td>Spatial Sound Card</td>
<td>Dolby Atmos for Headphones</td>
<td>Representation</td>
<td>-0,630</td>
<td>2,110</td>
</tr>
<tr>
<td>Dolby Atmos for Headphones</td>
<td>Stereo</td>
<td>Localization</td>
<td>-0,574</td>
<td>2,110</td>
</tr>
<tr>
<td>Stereo</td>
<td>Spatial Sound Card</td>
<td>Localization</td>
<td>1,175</td>
<td>2,110</td>
</tr>
<tr>
<td>Spatial Sound Card</td>
<td>Dolby Atmos for Headphones</td>
<td>Localization</td>
<td>-0,444</td>
<td>2,110</td>
</tr>
<tr>
<td>Dolby Atmos for Headphones</td>
<td>Stereo</td>
<td>Appeal</td>
<td>1,606</td>
<td>2,110</td>
</tr>
<tr>
<td>Stereo</td>
<td>Spatial Sound Card</td>
<td>Appeal</td>
<td>1,245</td>
<td>2,110</td>
</tr>
<tr>
<td>Spatial Sound Card</td>
<td>Dolby Atmos for Headphones</td>
<td>Appeal</td>
<td>-1,606</td>
<td>2,110</td>
</tr>
</tbody>
</table>
Table 6 – Table displaying the results of the Wilcoxon Signed-Rank Test used to see if there was a significant difference despite the lack of normal distribution in the data.

<table>
<thead>
<tr>
<th>Pair</th>
<th>Attribute</th>
<th>W - Value</th>
<th>Z - Value</th>
<th>Statistically Significant</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dolby Atmos for Headphones</td>
<td>Stereo</td>
<td>49</td>
<td>-0.9825</td>
<td>No</td>
<td>0.32708</td>
</tr>
<tr>
<td>Stereo</td>
<td>Spatial Sound Card</td>
<td>26</td>
<td>-1.6636</td>
<td>No</td>
<td>0.09692</td>
</tr>
<tr>
<td>Spatial Sound Card</td>
<td>Dolby Atmos for Headphones</td>
<td>56</td>
<td>-0.6205</td>
<td>No</td>
<td>0.53526</td>
</tr>
<tr>
<td>Dolby Atmos for Headphones</td>
<td>Presence</td>
<td>41.5</td>
<td>0.6905</td>
<td>No</td>
<td>0.4902</td>
</tr>
<tr>
<td>Stereo</td>
<td>Spatial Sound Card</td>
<td>Presence</td>
<td>50.5</td>
<td>-0.5396</td>
<td>No</td>
</tr>
<tr>
<td>Spatial Sound Card</td>
<td>Dolby Atmos for Headphones</td>
<td>Presence</td>
<td>31</td>
<td>-1.0133</td>
<td>No</td>
</tr>
<tr>
<td>Dolby Atmos for Headphones</td>
<td>Sense of Space</td>
<td>45.5</td>
<td>-0.4394</td>
<td>No</td>
<td>0.65994</td>
</tr>
<tr>
<td>Stereo</td>
<td>Spatial Sound Card</td>
<td>Sense of Space</td>
<td>45</td>
<td>-1.4912</td>
<td>No</td>
</tr>
<tr>
<td>Spatial Sound Card</td>
<td>Dolby Atmos for Headphones</td>
<td>Sense of Space</td>
<td>20</td>
<td>-1.4905</td>
<td>No</td>
</tr>
<tr>
<td>Dolby Atmos for Headphones</td>
<td>Representation</td>
<td>37.5</td>
<td>-0.5591</td>
<td>No</td>
<td>0.57548</td>
</tr>
<tr>
<td>Stereo</td>
<td>Spatial Sound Card</td>
<td>Representation</td>
<td>44</td>
<td>-1.241</td>
<td>No</td>
</tr>
<tr>
<td>Spatial Sound Card</td>
<td>Dolby Atmos for Headphones</td>
<td>Representation</td>
<td>37</td>
<td>-0.973</td>
<td>No</td>
</tr>
<tr>
<td>Dolby Atmos for Headphones</td>
<td>Localization</td>
<td>45.5</td>
<td>-0.8235</td>
<td>No</td>
<td>0.41222</td>
</tr>
<tr>
<td>Stereo</td>
<td>Spatial Sound Card</td>
<td>Localization</td>
<td>27.5</td>
<td>-1.2579</td>
<td>No</td>
</tr>
<tr>
<td>Spatial Sound Card</td>
<td>Dolby Atmos for Headphones</td>
<td>Localization</td>
<td>35.5</td>
<td>-0.6989</td>
<td>No</td>
</tr>
<tr>
<td>Dolby Atmos for Headphones</td>
<td>Appeal</td>
<td>39.5</td>
<td>-1.4737</td>
<td>No</td>
<td>0.14156</td>
</tr>
<tr>
<td>Stereo</td>
<td>Spatial Sound Card</td>
<td>Appeal</td>
<td>50</td>
<td>-0.9308</td>
<td>No</td>
</tr>
<tr>
<td>Spatial Sound Card</td>
<td>Dolby Atmos for Headphones</td>
<td>Appeal</td>
<td>24</td>
<td>-1.5025</td>
<td>No</td>
</tr>
</tbody>
</table>
7.4 Appendix 4 - Questionnaire

Ljud X

Lyssnare #

Varifrån upplevde du att ljuden kom?

Helt inifrån huvudet
Helt utanför huvudet

Hur närvarande i spelet blev du av ljudet?

Inte Alls Närvarande
Extremt Närvarande

Gav ljuden dig känslan av rymd?

Inte Alls
Extremt Mycket

Upplevde du ljuden som representativa för miljöerna?

Inte alls representativa
Extremt Representativa

Kunde du höra vart ljuden kom ifrån?

Inte Alls
Extremt Väl

Hur tilltalande var ljudet?

Inte Tilltalande Alls
Extremt Tilltalande

Övriga kommentarer om ljudet skrivs ned i rutan nedanför
### 7.5 Appendix 5 – List of used Microsoft Product Reaction Cards

<table>
<thead>
<tr>
<th>Annoying</th>
<th>Appealing</th>
<th>Attractive</th>
<th>Boring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calm</td>
<td>Clean</td>
<td>Clear</td>
<td>Comfortable</td>
</tr>
<tr>
<td>Compelling</td>
<td>Comprehensive</td>
<td>Confusing</td>
<td>Connected</td>
</tr>
<tr>
<td>Consistent</td>
<td>Convenient</td>
<td>Creative</td>
<td>Difficult</td>
</tr>
<tr>
<td>Disconnected</td>
<td>Disruptive</td>
<td>Distracting</td>
<td>Dull</td>
</tr>
<tr>
<td>Effective</td>
<td>Efficient</td>
<td>Empowering</td>
<td>Energetic</td>
</tr>
<tr>
<td>Engaging</td>
<td>Entertaining</td>
<td>Essential</td>
<td>Exceptional</td>
</tr>
<tr>
<td>Exciting</td>
<td>Familiar</td>
<td>Fresh</td>
<td>Frustrating</td>
</tr>
<tr>
<td>Fun</td>
<td>Helpful</td>
<td>High Quality</td>
<td>Impressive</td>
</tr>
<tr>
<td>Incomprehensible</td>
<td>Inconsistent</td>
<td>Ineffective</td>
<td>Inspiring</td>
</tr>
<tr>
<td>Integrated</td>
<td>Intimidating</td>
<td>Intuitive</td>
<td>Inviting</td>
</tr>
<tr>
<td>------------</td>
<td>--------------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>Motivating</td>
<td>Old</td>
<td>Ordinary</td>
<td>Organized</td>
</tr>
<tr>
<td>Overbearing</td>
<td>Overwhelming</td>
<td>Poor Quality</td>
<td>Powerful</td>
</tr>
<tr>
<td>Predictable</td>
<td>Professional</td>
<td>Relevant</td>
<td>Reliable</td>
</tr>
<tr>
<td>Responsive</td>
<td>Rigid</td>
<td>Satisfying</td>
<td>Simplistic</td>
</tr>
<tr>
<td>Sophisticated</td>
<td>Stable</td>
<td>Sterile</td>
<td>Stimulating</td>
</tr>
<tr>
<td>Stressful</td>
<td>Time-Consuming</td>
<td>Time-Saving</td>
<td>Trustworthy</td>
</tr>
<tr>
<td>Unapproachable</td>
<td>Unattractive</td>
<td>Unconventional</td>
<td>Undesirable</td>
</tr>
<tr>
<td>Unpredictable</td>
<td>Unrefined</td>
<td>Usable</td>
<td>Useful</td>
</tr>
<tr>
<td>Valuable</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>