Visual Substitutes for Audio Cues
Providing situational awareness for players with auditory disabilities

Casper Strand
David Haddad

Main Field: Computer Science
Level: Bachelor 180 hp
Credits: 15 hp
Submitted: Spring of 2019
Supervisor: Jeanette Eriksson
Visual Substitutes for Audio Cues
Providing situational awareness for players with auditory disabilities

Casper Strand
Game Development Program
Computer Science and Media Technology
Malmö University
casperstrand@outlook.com

David Haddad
Game Development Program
Computer Science and Media Technology
Malmö University
davidhaddad93@outlook.com

Abstract—Video Games have turned into a mainstream source of entertainment and have evolved alongside technological improvements. To the point of being able to bring the player immersive gaming experiences that rely heavily on their senses. However, the attention to the accessibility aspects of video games has been lagging behind.

This thesis explores the available solutions that aim to help Deaf or Hard of Hearing (DHH) players, primarily in terms of situational awareness by substituting audio with a visual representation. The currently available solutions are then compared with the findings in related research and a survey that was conducted for this paper. The objective was to find what worked as a visual representation of audio, with the survey responses providing insight into DHH players perspective.

These findings can serve as a reference point for developers who want to implement a visual substitute for audio or a similar feature in their game.

Index Terms—visual assistance, deaf, hard of hearing, hearing impaired, accessibility, visual overlay

------------------  ◀  ------------------
1 INTRODUCTION

The purpose of this paper is to investigate and study the different attempts at replacing or complimenting audio with a visual Heads Up Display (HUD) that is intended to enable situational awareness for players who are Deaf or Hard of Hearing (DHH). Whilst also shedding some light on why this type of accessibility should exist, at least for games that rely heavily on situational awareness where audio is a crucial factor.

An artifact created to alleviate a specific disability will be referred to as an accessibility layer, examples of accessibility layers include the ability to remap input controls, having subtitles where there are voiced dialogues or adding different color filters and descriptive texts to help with color blindness [20, 1].

The attention for accessibility in video games has gained traction in recent years, with an increasing number of options aimed at alleviating a range of disabilities populating the settings menus in video games [6]. Some of this can be attributed to the increase in the available resources that developers have at their disposal when making their games more accessible [6]. Resources such as AbleGamers Accessible Games [1] and a website that has a compiled list of Game Accessibility Guidelines [20]. But as of writing this thesis there are no developer guidelines or resources regarding ways to substitute audio in a way that would help a DHH player with situational awareness.

The currently most popular games require a lot attention from a player as they rely heavily on audiovisual input to provide player’s with situational awareness, this is especially the case with action-based games [7, 22].

But DHH players are not able to appreciate all the information available to them, as most or substantial portions of the audio aspect gets lost for a DHH player, resulting in a difficult time to perform at the same level as a player without an auditory deprivation.

This places players who are unable to hear the game, or otherwise unable to discern sound directionality at a great disadvantage as they are only able to process a limited amount of the information available for making up a player’s situational awareness. Creating a gap within the playerbase between these two groups, as only one of them are able to appreciate the full spectrum of audiovisual input that informs the player of their surroundings. Something which could discourage a DHH player from playing certain types of games, which could even lead to the DHH player to become socially excluded in that player’s social circle depending on the games they enjoy [10].

But before looking at various methods that tackle this issue and the data behind it, it’s important to note that there are many variations in terms of a individuals hearing situation. The hearing loss can be symmetrical or even asymmetrical as in that a person has better hearing on one side over the other, or even no hearing on one side and perfect hearing on the other. The level of hearing can also differ in regards to certain frequencies, where certain frequency ranges can be difficult to hear compared to other frequencies [27].

Fig. 1. An example of a Cookie Bite Hearing Loss audiogram. The dashed line represents unimpaired hearing [26, 25].

An easy to comprehend example of a type of hearing loss is the less common “Cookie Bite Hearing Loss”, that is named after the way the audiogram looks, as illustrated in Figure 1. From what normally should be a smooth line for someone with unimpaired hearing, there is a significant dent, usually in the mid-range frequencies, as illustrated by its fitting name [25]. Meaning the individual usually registers the low and high frequencies with relative ease, but misses out on the middle frequencies. Which is where some of the most crucial sounds for speech come from [4], as illustrated in Figure 2.
By studying the related research, the available solutions and the available technologies as potential alternatives or improvements to these solutions. This data form the basis for all the questions of the online survey that targeted DHH individuals, data was then collected, analyzed and processed, all in an effort to answer the following questions:

**RQ 1:** What are the different visual substitutes for sound that help with situational awareness for DHH players currently available?

**RQ 2:** What are the drawbacks or limitations, if present, with visual substitutes that help with situational awareness?

**RQ 3:** What is the research that support or contradict the visual substitutes for sound that help with situational awareness?

**RQ 4:** What are the different opinions of DHH players regarding these solutions, is there a general consensus?

As Yannakis Georgios N. and Julian Togelius described, player experience is an important issue in most game development processes and in order to create games that keeps the players curiosity level high, an analysis of the player interaction with every aspect of the game, even sound must be made. Note, this statement was made in the context of all types of sound in a game. This statement is important because it illustrates the requirements that a compliment or replacement for sound needs to adhere to, or at least get as close as possible to.

The outline for the remainder of this thesis is as follows.

Following this introduction is a description of the methodology and process that was followed through out this paper and the survey, as well as how the survey data was prepared for the result section. This is then followed by a review section for existing solutions in software and video games that replace or compliment audio with a visual interface, continued with related research section that mentions research relevant to the topics discussed in this paper, there after it moves on to the related technologies section that examine the potential alternatives and/or improvements to the solutions mentioned in prior sections. Lastly, it demonstrates and discusses the results which is then followed with a conclusion where everything gets tied together.

## 2 Methods & Work Process

### 2.1 Work Process

The initial steps of this thesis consisted of finding games that had an accessibility option aimed at DHH individuals that utilized a visual interface, and related research that either supports the notion that a visual interface is an appropriate solution or contradicts it.

Most of the work done in the initial stages of this thesis involved procuring research papers related to the topic at hand, the research was screened thoroughly and highlighted potential segments of a paper that could be used, but no research was found that directly supported or contradicted the notion that visual substitutes or compliments are an appropriate medium to replace sounds with.

The research papers related to accessibility were mainly focused on physical disabilities where physical movement was somehow affected, as well as intellectual or cognitive disabilities or degradation.

However, two papers stood out early on during this thesis paper, that showed that deaf individuals who had experienced an early onset of hearing deprivation, had benefited from cross-modal plasticity. Which is when neurons...
adapt by integrating two or more sensory inputs, more so when one or more of the sensory inputs are deprived. But there were no direct answers that could answer the research question of this thesis, which is why these findings became the basis for the questions formulated in the survey.

After concluding that the available research related to the topic was lacking, the decision was made to delve deeper into exploring what games were available that included some level of accessibility option that helped DHH players. However, the games that used visual overlays to indicate sound directionality as a way to help with situational awareness, were implemented in many different ways. All the games found are elaborated on in the 3.3 Existing Solutions section. Other titles that used the accessibility option called Closed Captions were more common, which is why only Half-Life 2 is included in the single example since they are all essentially the same. All of the games that in this thesis are modern and well-known game titles, with seemingly well functioning visual overlays, others more so than others.

These solutions are all elaborated on and compared to each other to see how they differ from each other. The Closed Caption option does not contribute to situational awareness and is therefore off-topic to the thesis. But it is still included as a baseline and as an option for developers who are not making games that rely on situational awareness, but still want to make their game accessible to DHH players.

Due to the lack of conclusive answers from the research papers, the need to get a large sum of answers that could answer the final research question became more important since none of the research papers answered it directly. Which is why the survey was compiled based on the two papers that showed results that seemed to suggest that a visual overlay is an appropriate medium for replacing sound cues that help with situational awareness. With the goal of trying to find a general consensus regarding visual substitutes for sound cues, that attempt to provide situational awareness for DHH players. The amount of games was limited to two game titles that were deemed the best solutions available, the full reasoning can be found in the 3.3 Existing Solutions in Video Games section. All “outside-in” approaches were excluded because they did not compare to the game-specific solutions.

After the results from the responses had been analyzed, the survey did not end up with the clear-cut answer that was expected. Which meant other previously unexplored avenues of related research had to be explored in an attempt to answer this. Which led to the discovery of an additional research paper previously unexplored that compared non-video game players with habitual video game players. The main point of interest from this paper was the final experiment where the non-video game players played a action-based game daily over 10 days, a game which was similar to those that the habitual video game players played. At the end of it all non-video game players showed significant performance increases.

This paper became an important point of reference when discussed in conjunction with the responses from the two games that were demonstrated in the survey. Because the results of the survey showed similar numbers between the participants who thought a particular solution helped or somewhat helped, and the participants who thought a solution did not help or was too distracting.

With the inclusion of this research paper the results are arguably not as split, since it can be argued that the participants might not fully appreciate a particular solution as a substantial part of the participants had not played one or the other game.

Instead it allowed the opportunity to argue that a participants response is actually more likely to change if they regularly played one of these games over an extended period, since they would most likely improve over time and get adjusted to the visual interface.

The following search terms were used when looking for information about the various software solutions and game-specific solutions, as well as finding the research papers that are discussed through-out this thesis:

Disability, Deaf, Deafness, Visual field, Sensory Deprivation, Gaming, Accessibility, Visual Aid, Hard of Hearing, Hearing Impairment, Visual Indicators, Visual Feedback, Games for Deaf or Hard of Hearing, Sound Radar, Action Video Game Players,

These are just the main search terms in no particular order, several variations of each term were used when applicable (e.g. roman and numeral integers were both used in the case of Final Fantasy XIV [40]) of these as well as combinations of multiple terms were used when searching for relevant information. The abbreviations and the full names were both searched when present for some of the terms (e.g. Counter-Strike: Global Offensive and CSGO [47]). Some of words were combined or had hyphens added them. (e.g. Thirdparty, Third-party or Third Party).

The reason that this thesis was approached by first examining the related research was because it felt important to learn what aspects had not been explored by already established research. This also made sense because it contributed directly to how the survey was formulated, since the related research provided a framework to work from. Which then provided the responses to the surveys early on which was needed because it was expected to receive a sizable number of responses, albeit still relatively small in the grand scheme of things. The responses were then manually screened and compiled into a single survey, since they were originally in two different languages. This allowed the opportunity to intercept early on if there was an unexpected outcome from the surveys, which actually turned out to be the case as there was no real black or white answer. Thereafter, since the survey was already done, time was spent trying to reach a conclusion by bringing in new research that was previously unexplored since they were unknowns until all the responses were compiled.

2.2 Participants
The survey was distributed on two different online deaf communities, the first one of the biggest Facebook groups in Sweden for deaf people [8] and the second one being the largest forum for deaf people on Reddit [9]. Although the survey was mainly directed towards avid gamers who are deaf, anyone with some level of auditory deprivation was encouraged to take part in the study. To make sure that as many participants as could realistically be managed took part in the survey, the survey was made available in two languages, Swedish and English to cover international audiences.

Due to the nature of the surveys some participants with multiple disabilities had to be excluded from the results, as their combination of disabilities resulted in a response that is unfair to the implementations demonstrated. An example could be an individual with a significant visual impairment in combination with hearing loss, which makes a visual substitute for situational awareness unsuitable. This was done because including these would complicate the scope of this thesis paper, instead the decision was made to focus on how these implementations fare under optimal conditions, where the participant is only affected by means of hearing deprivation.

2.3 Compiling and Screening Process
First step after gathering all the responses was to compile the two surveys into one, as they are identical in content with the only difference being their language. There was a total of 12 responses from the Swedish survey and 43 responses from the English survey, making it a total of 55 participants that took part in the surveys.

The second step was to screen all the participants responses, to ensure that all of the information was relevant to the research and to look for any other factors that could affect the participants answers. One participant was omitted due to having a combination of disabilities which meant they were an unsuitable candidate for this survey, according to the reasoning described in the Participants section prior.

Some exclusions were made of specific answers from a participants response if they failed to see what a demon-
stration tried to highlight. Examples of this included participants who were unable to see or notice the visual overlay, misunderstanding or not fully appreciating aspects of a visual overlay (e.g. not noticing the directional arrows adjacent to the text in Minecraft’s Subtitles).

This could partially be attributed to Google Forms small embedding of the demonstration videos, additionally the ability to fullscreen is not enabled on Google Forms. Clear instructions were in place as to how to view the video on YouTube so the participant could watch it in fullscreen, but despite a valid effort some people missed these instructions.

In the case of Fortnite, some participants had troubles with the colors of the visual overlay being hard to distinguish from the rest of the game. However, these ones are still included in the results. Five individuals pointed out that it was difficult to see the visual overlay. Six individuals could not see the overlay at all.

In the end 11 participants were either fully or partially omitted from the participants responses. 3 of which were fully omitted, the rest were only partially omitted. Partially omitted means that their response is not included in both games. Making it a total of 52 participants after the screening process, where 8 of them have certain answers that will not influence the results.

2.4 Ethics and Legal Statement

Permission was given by the moderators for each forum prior to publishing the surveys. Both versions of the survey comply with GDPR, based on a set of guidelines aimed specifically at writing surveys. Finally, the participants of the surveys are unable to proceed without first giving their explicit consent that their response can be collected and processed for this thesis paper. They are also made aware that the personal information shared will not be made publicly available, it is solely used to verify that a real individual submitted it. The data has then anonymized prior to being presented in the result section. The full Terms of Agreement that participants have agreed to can be found in the surveys located in the Attachment section.

2.5 Credibility

The data presented in the result section consists of the sum of answers from the two surveys, the one written in English was posted on a subforum for members on Reddit called r/Deaf. The second one written in Swedish was posted in a closed group for members on Facebook called “Deafhood behövs i Sverige”. This limits the exposure of the surveys to only include the type of participants whose data we are interested in, but also allowed the survey to reach a relatively large group of people.

The participants of the surveys were asked to disclose personal information including their name, gender, hearing situation and gaming habits in order to ensure that the data was unique. Age was not relevant for the study since this thesis focused on the accessibility aspect and without focusing on any particular target audience.

The possibility that the information is false is always a risk when performing an online survey, but given the questions asked and the communities these were posted in, it is unlikely that the data has been fabricated. However, before the results were presented they were subject to a manual screening process, ensuring that each participants data was unique and did not indicate attempts of fabrication or otherwise raising concerns as to the submissions validity.

2.6 Procedure

The online surveys were distributed via social media in the appropriate forums. The surveys were created and hosted on Google Forms. PDF versions of these are available in the Attachment section. It was split into three parts, that was prefaced with a Terms of Agreement section that all the participants had to agree to if they wished to participate in the study.

The first part of the survey covers information about the participant, all entries in this section are required with the exception of gender. The reason for this was to be able to screen or reduce the risk for duplication, fabrication or otherwise suspect data. This section asks the participants to give information regarding their name and gender. Followed by the participants hearing situation, including when
in life it was found out, to what degree and what type of hearing aid they use (if any). The reason for this was to see if there were any trends in the answers of the survey based on certain hearing situations.

This is then followed by a section dedicated to their gaming habits where the participants were asked how much time per week they spend playing video games, if they have played games that have heavy emphasis on situational awareness and if they feel that they are at a disadvantage due to their hearing situation.

The second part of the survey covers the participants knowledge and experience when it comes to gaming, as well as a short explanation of the terms used in the survey that the participants might be unfamiliar with. Followed by a section dedicated to asking the participant if they have played games that use a First Person Shooter (FPS) or Third Person Shooter (TPS) perspective, if they play games online with others, if they feel at a disadvantage when playing online due to their hearing situation and if the games provide enough accessibility layers that adequately substitute or compliment game audio. The last two questions are also open-ended since they are followed with an opportunity for the participant to elaborate.

The data from the second part is mainly used to verify that they player is experienced with the types of games where situational awareness is important, and whether or not they are pleased with the existing accessibility layers (if any). The answers from these questions also provides some insight as to a participant’s experience prior to being shown the various solutions aimed at providing situational awareness found in commercial video games in the final part of the survey.

The final part of the survey present the different solutions used by modern games to help with situational awareness for DHH players. The games presented are the same ones that were brought up earlier in the section for 3.3 Existing Solutions in Video Games, with the exception of Final Fantasy XIV and Half-Life 2. These two games were omitted from the survey because they fell short on providing sufficient situational awareness. The full analysis for these two games can be found in the section 3.3 Existing Solutions in Video Games, but the main reason they were omitted from the survey is as follows. Half-Life made use of Closed Captions which tells the player all the sound that happens and all the dialogue, but this method does not tell the player where these sounds happen. Final Fantasy XIV did provide situational awareness, but on a very simple level, since the player can only see where sound is coming from but not what it is or how many targets are making sound. Since all sounds overlap which causes noise in the visual overlay. Requiring that the player look towards the sound source to discern that information.

The solution of each game was presented with a detailed description and a video demonstrating it in use. After which the participant get to answer whether or not they have played the demonstrated game before. Thereafter the participants are asked to answer whether they found it useful and to what degree, followed by an opportunity to elaborate. This data is what is used to answer RQ 4.

The questions present in the survey are phrased in such a way so that the participants get to voice their opinion on whether they think that the accessibility options in the examined games are helpful, as well as allowing the participants to elaborate on how they think the issues could be remedied. The reason for formulating the questions in such an open-ended matter is due to the lack of available research around this subject. The information provided by the participants in these open-ended questions also influence the answer for RQ 4.

3 RELATED WORK

3.1 Related Research

The World Health Organization has compiled data on the Global Estimates on Prevalence of Hearing Loss, which estimates that there’s approximately 466 million individuals with disabling hearing loss, which is roughly 6.1% of the world’s population.

The World Health Organization classify an individual to have disabling hearing loss if the individual has a hearing loss greater than 40 decibels (dB) in the better hearing ear in adults (15 years or older), and a hearing loss
greater than 30 dB in the better hearing ear in children (14 years or younger). The World Health Organization’s projections indicate that these number will continue to grow, estimated to reach 640 million individuals by 2030 and over 900 million by 2050.

An article that compared video game players and deaf observers Goldmann visual fields written by David Buckley et al. [7], found that deaf observers had significantly larger visual fields than video game players. The Goldmann Kinetic Perimetry [7] is a standard clinical test used when assessing a patient’s visual field sensitivity in UK eye clinics, a spot of light is projected onto a half dome where it is moved along a radius from the outer periphery towards a fixation point, where a patient’s gaze is directed to the central fixation point and is told to press a button when they detect the spot of light. The participants for the Goldmann Kinetic Perimetry tests consisted of 13 deaf participants and 10 habitual action video game players, none of the video game players were deaf or users of sign language. David Buckley et al. [7] measured the ability for each group to detect small moving lights at various locations in the central (around 30° from fixation) and peripheral (around 60°) visual fields. David Buckley et al. [7] did not find any conclusive difference in the central Goldmann field between the deaf and video game players, which might suggest that both groups have undergone the same changes in attentional mechanisms in this portion of the visual field.

However, the distribution were significantly larger in deaf participants when it came to the peripheral Goldmann field, with the largest difference compared to the video game players occurring in the lower regions of the visual field. Something which could be attributed to prolonged stimulation from sign language viewing, as the lower regions of the peripheral visual field is responsible for processing the hand movements involved in sign language. Although it is worth pointing out that David Buckley et al. [7] found no data to support this, but their sample size was not large enough to conclusively rule out the possibility either.

A similar article written by Charlotte J. Codina et al. [11] measured the visual reaction times between deaf, British sign language (BSL) interpreters and hearing adults. The results from this survey concluded that deaf have faster reaction times than the other two groups. Some of the authors from the Goldmann visual fields study [7] were also the authors for this study.

The testing was performed using a hemispherical dome with 96 LEDs placed in steps of 5° starting from 30° and ending at 85°, placed along eight meridians, with an LED in the center as the fixation point. The schematic for this can be seen in Figure 3. The test was done by illuminating a random LED for 200ms, and the result was only recorded if the participant’s gaze remained fixated on the central target.

The results showed that deaf adults demonstrated significantly faster reaction times than both hearing non-signers and hearing BSL interpreters. BSL interpreters displayed faster reaction times than hearing non-signing adults across the entire visual field, but not as significant as the deaf adults. Early onset deafness leads to visual compensation with much faster reaction times in the peripheral vision, something that is consistent with the cross-modal plasticity benefits (neurons re-organize to integrate two or more sensory systems in an effort to adapt) to vision of auditory deprivation and use of visual spatial language. Furthermore, the far peripheral vision is important when monitoring the environment due to sensory deprivation. [11]

The article called “Action video game modifies visual selective attention” [24] compares a group of video game players against a group of non-video game players in an array of experiments. The first test uses the flanker compatibility effect, a standard experimental paradigm in attentional studies designed to see if video game playing
extends the individuals attentional capacity. The task is measured by having the test subject focus on a target task whilst avoiding a to-be-ignored distractor of different sizes. The results of this test showed that video game players possessed an enhanced attentional capacity, outlasting the non-video game players.

The second experiment of interest was made using an enumeration task, where the participants are asked to say how many squares are represented in a briefly flashed display. The video game players could subitize (identify objects) more items than non-video game players. Which proved that video game players have an enhanced ability to unerringly apprehend visual items in a rapid manner.

The third and last experiment of interest tested for attentional blink, which looks for the ability to avoid “bottlenecking” an individual’s attention during temporal processing. Attentional blink refers to a phenomena where subjects have a difficult time processing a secondary target that is introduced very shortly after an initial target. This test was made with five different levels of lag for the secondary target, ranging from 200-500ms. The results of this test revealed that video game players outperformed the non-video game players with a higher correct detection rate, indicating less attentional blink.

Lastly, to rule out that it could be a fluke that the non-video game players performed worse than the video game players, a group within the non-video game players were asked to play a game similar to those that the video game players played as a form of training, for 1h per day for 10 consecutive days. This training resulted in tangible improvements in the useful field of view and faster recovery from the attentional blink. Thus ruling out that it was a mere fluke that this group underperformed compared to the habitual video game players, instead it was because the habitual video game players had already honed their attentional skills previously. [24]

3.2 Related Technologies

This section is dedicated to exploring various technologies that are of interest to visual indicators for sound effects, including established and alternative technologies that could serve as the base for a visual indicator, as well as potential ways to extend the functionality of the visual indicator. Additionally, alternative technologies for enabling situational awareness for DHH players are also explored.

With the recent growth of popularity in Virtual Reality (VR) gaming peripherals [41] many companies have started improving existing gaming peripherals, such as headphones, computer mice, chairs, vests [28] or full body-suits [42] that incorporate haptic feedback. One of the companies developing immersive accessories, Razer, has recently released a line of PC peripherals that all have haptics built in [35], allowing the user to feel the events and sounds happening in games. As these tools let you feel the sounds, it may be possible to adjust them to used as a physical aid to help DHH players not only see where the sounds are coming from but also feel the direction of the sound. This technology could yield promising results given that it becomes widely supported by developers.

Tobii Eye Tracking [44] is a piece of hardware that allows a user interact with the computer using their gaze, by using the eyes as a pointer on the screen. Tobii [44] is already widely used in various assistance-type applications that enable people with certain types of disabilities a way to communicate, by utilizing augmentative and alternative communication [45]. Which works so long as there is nothing affecting the users ability to look around with their eyes. Tobii Eye Tracking [44] could be interesting to implement with a visual indicator to extend its functionality, which could be a research topic worth exploring on gaze interaction with the various elements composing a visual indicator for sound effects.

Surround sound technologies are an interesting technology to explore since a “outside-in” approach could benefit greatly from the latest surround sound technologies. CanetisRadar [39] relied on Razer Surround [36] which is virtual 7.1 surround, which has a couple of limitations. The main limitation being that 7.1 surround sound does not feature elevation, which means that all the sound is on the same height with no sound from above or below.

That is because of the limitations of 7.1 or 5.1 surround, the first number represents the number of speakers
surrounding the user. The second number represents the number of subwoofers. But all of these speakers, virtually simulated speakers, are all placed on the same horizontal plane in height. Therefore unable to replicate elevation of sound [33].

However, this has changed with the introduction of speakers that are responsible for elevation. An example of a speaker setup like this could be 7.1.4, where the last number that have been added to the end reveals that there is four speakers for the height channels. Technologies such as Dolby Atmos [15] will then approximate the 3D positions to a given speaker setup, both horizontally and vertically [33].

These technologies that approximate 3D positions of sounds are called 3D or Object-based Surround Sound, one of these being Dolby Atmos [15]. Dolby Atmos [15] receives audio as individual objects instead of different channels, and supports up to 128 distinct sound objects, which then will be processed and approximated to a speaker setup that supports Object-based Surround Sound [33].

Another alternative that could benefit “outside-in” approaches is Spatial Sound, which also enables elevation of sound as well as surround, with one of the implementations being Microsoft’s Windows Sonic. A feature that is available on Windows 10 (version 1703 or newer) and Xbox One. This alternative can be used with any speaker setup or pair of headphones, it will fallback to virtual spatial sound if it is not supported natively [50].

However, the issue with Windows Sonic [50] or Dolby Atmos [15] is that they are locked down, there is no way to access the 3D positions or information on the sounds themselves. But if it was possible to access this data, it could benefit developers as it would open up the possibility of creating a universal solution for enabling situational awareness using visual indicators. Which would work as long as a game utilized the technology, without requiring the developers of the game to create their own visual indicator if they wanted to implement this accessibility feature.

There is plenty of third-party realistic audio engines that can be used for a game, that makes sound effect behave in a realistic manner through the virtual environment. An example being VERA [29], which stands for Voxel Engine for Real-Time Acoustics, which converts 3D voxel geometry into voxels that has density that becomes the environment for the sound to reverberate, reflect, portal and get occluded or obstructed in. Enabling realistic audio propogation that helps with immersion for hearing players.

Most game engines also have competent audio engines that are built-in that could be extended on for creating a visual indicator for sound cues, like Unreal Engine 4’s Audio System [18], which supports realistic modelling of sound like reverb effects where you can adjust individual sound properties such as echo density, air absorption and various sound reflection values.

Third-party or in-game solutions like these are interesting as they could serve as the back-end for a visual indicator, where the sound could remain realistic without showing up as noise or incorrect directions on a visual indicator. Thus avoiding the issues that arose with “outside-in” approaches CanetisRadar [39] or Awded [30].

There is software in the works that aims to alleviate the issues inherent to “outside-in” approaches. Ideally, by using Machine Learning, this could allow for separating multiple audio sources in an audio mix whilst also filtering out the noisy sound effects. Which could result in a “outside-in” based visual indicator that is functional, even in a virtual environment that has realistic sound effects and background sounds, both of which should not affect the visual indicator if the machine learning is working. This could be a research topic worth exploring.

Awded [30] has a complimentary software utility called AwdedTrainer [31] that uses Machine Learning to train the visual indicator.

### 3.3 Existing Solutions in Games

This section is dedicated to exploring the various solutions in video games that attempt to create visual alternatives for sound effects.

#### 3.3.1 Half-Life 2

The developer’s behind Half-Life 2 and the Source Engine [46] have implemented a solution they call “Closed Cap-
tions” [48]. Which not only display all the subtitles for the game, but also displays text descriptions for all the sound effects as they happen in the game. However, an important distinction with this solution compared to the following solutions is the lack of information related to where the sounds originate from. An example can be seen in Figure 4, where the Closed Captions [48] have been enlarged for illustration purposes in the upper section of the figure. This example includes a subtitle at the top followed by two sound effects, one for the gunfire and one for the water splash as a result of said gunfire. All sound effects are encapsulated with brackets and color coded to match their purpose. White is for environmental sound effects or cues while orange is related to sound effects that deal or is associated with damage and item pick ups within the game. Sound effects with a red text is used whenever an enemy or the player dies.

3.3.2 Minecraft

The game studio behind Minecraft, Mojang, have implemented a solution they call “Subtitles” [32]. Which when enabled appears in the bottom right corner of the screen, as portrayed in Figure 5, displaying a string of text with arrow indicators on each side. The text informs the player what type of sound it is and the arrows point to which direction it originates from. These arrow vary in their level of opacity, depending on the player’s view and the source of the sound. The text will eventually fade away or disappear, as the sound source goes out of range or otherwise ceases to emit sounds.

3.3.3 Final Fantasy XIV (14)

Square Enix offers a different solution in their game Final Fantasy XIV [40], where sound is depicted as a semi-transparent visual equalizer along the edges of the screen, as demonstrated in Figure 6. Offsetting the visual cues to the player’s peripheral vision, whilst also avoiding obfuscating the UI elements since it is semi-transparent.

The audio visualizer changes positions to account for the direction of the incoming sound and the heights of the peaks depicts the volume level, giving the player an estimation of
the proximity and direction of a single or multiple sound sources.

Similar to Minecraft’s solution, this method does not offer much in the way of displaying the verticality of the surround sounds, instead the game just shows the horizontal direction of the sound source on each side of the screen, making it difficult to discern whether a sound originates from above or below the player. However, this method introduces visual noise which Minecraft does not. As the player is only informed that there is sound coming from a certain direction, but is unable to tell what type of sound it is. For example, the sound could be that of a crackling sound effect from a fireplace or a monster charging at you from behind. Whichever it is, the player will not know without having it in line of sight. Which could end up hindering or distracting the player.

3.3.4 Fortnite

The game studio behind Fortnite, Epic Games, have implemented an overlay they call “Visualize Sound Effects” that can be enabled in the settings for PC players. On mobile devices this option is enabled by default. However, there is a key difference between the mobile version and the PC version. When this option is enabled on PC, all game audio is disabled. This is not the case on mobile, instead both options are enabled simultaneously. Note, this has since been updated as of version 8.50, allowing the audio and visualizer to work simultaneously.

Fortnite’s visualizer, as demonstrated in Figure 7, consists of a radial circle that is centered around the crosshair reticle, and has multiple ring elements for all the different kinds of sound effects. The sounds are specified with a distinct color and icon for each type. With additional effects and animations added to specific types of sound effects, which are designed to highlight nearby dangers, such as gunfire or the destruction of structures.

The scenario in Figure 7 shows the visualizer depicting three nearby sound effects, in this case every symbol is represented with a different icon for each sound effect. The symbols have been enlarged in the top left corner of the figure and starting from the left the symbols represent footsteps, gunfire and a parachute. The footsteps are coming from the opponent that is in front of the pink bunny player and the parachute is coming from a separate opponent behind the player. The last symbol is on a separate radial circle outside the one with the footsteps and the parachute. This outer circle is dedicated to nearby dangers, which in this case is populated by the gunfire symbol that is coming from the opponent in front of the player, this symbol has been emphasized with a red tint and a “wavey” animation. A layer of emphasis that is similar to Half-Life 2’s Closed Captions.

Compared to the previous games listed, this solution is able to show verticality to the sound effects with its radial circle pointing towards sound sources. Since the direction indicators are relative to the player’s camera perspective, and are able to point downwards and upwards unlike the previous games. However, since the visualizer is displayed in a radial circle offset from the center of the screen, it becomes difficult to tell if a sound is right in front of the player since the visualizer can only visualize the sound direction around the radial circle outside the center of the screen. Thus giving a visual indicator that could potentially be misleading in that particular scenario.

3.3.5 Apex Legends

A survey participant mentioned Apex Legends praising the accessibility options available for DHH players in the game. Apex Legends include an option to enable speech to text, which will translate voice communications into text, there is also an option for the reverse that is text to speech.
Fig. 8. Apex Legend's pinging system

By enabling these two a deaf player can get all the voice communications translated into text, and be able to answer in text which is then read aloud to the hearing players. The game also includes a pinging system as one of its main game mechanics, where the player can ping objectives such as where they intend to go and where enemy targets have been seen.

The pinging system as seen in Figure 8 has 8 different target signals, starting from the top and going clockwise there is “Go”, “Enemy”, “Looting This Area”, “Attacking Here”, “Going Here”, “Defending This Area”, “Watching Here” and lastly “Someone’s Been Here”. The description for each target signal is seen when hovering them in the game.

3.4 Existing Solutions in Software

The attempts to find any commercially available third-party software designed to alleviate the accessibility concerns discussed in this paper were mostly unsuccessful. However, there were several open-source solutions that attempt to achieve this to varying degrees.

There are several open-source software solutions that attempt to create visual indicators based on the audio of a game. But they’re in experimental stages at the time of writing this paper and utilize an outside-in approach, which means they have no access to the actual audio data from within a game but instead rely on decoding the audio output.

However, in its current state outside-in approaches become problematic due to the lack of data on the audio from the games themselves. Since there’s no algorithm to differentiate the Critical Audio Cues from Noisy Sound Effects.

Critical Audio Cues are those sounds that provide information pertaining to someone’s situational awareness, these could be the footsteps of an enemy opponent or gunfire in the distance. Critical information that can help a player decide on a strategy based on what they know about their environment.

Noisy Sound Effects are any form of sounds that could obscure the critical audio cues that provide situational awareness. Sound effects such as fire, wind and various background noises. Noise could also be introduced when the games recreates realistic audio propagation which allows the audio to reflect, bounce, reverberate and or otherwise get manipulated by the virtual environment the same way it would behave in a real environment.

Simulation effects could also create noise, such as adding echoes. Noisy sound effects become an additional layer of audio input that an outside-in approach cannot differentiate from the critical audio cues. A hearing player can easily distinguish the critical sounds from the noisy ones in an audio mix. Whereas a “outside-in”-based visual overlay fails to replicate this ability of separating different sounds in an audio mix, resulting in noise that render the overlay unusable or to some degree distracting. Lastly, “outside-in” approaches have an additional concern compared to in-game solutions, which is latency. Since they need to be able to process the audio and update a visual indicator without introducing any noticeable lag which would distract the player.

3.4.1 Awded

Awded [30], is an open source solution that aims to visualize audio by overlaying the game with audio staples on the left and right side of the monitor as seen in Figure 9, similar to Final Fantasy XIV’s implementation that is demonstrated in Figure 6. Except Awded’s implementation has no support for transparency, thus obstructing the view of the game if there’s sound playing.

The testing of Awded was met with issues, getting it to work optimally proved challenging. The main issue was
that the visualizer kept indicating that there was sounds playing, even when there was not any sound output from the game. The visualizer also seemed overly sensitive to the sound output of a game, which made the visualizer overlay erratic and distracting.

3.4.2 Canetis Radar

Canetis Radar [39], may pose a challenge for the average user to set up, as it relies on third-party software such as Razer Surround [36] or VB-Audio Cable [49] and includes a 14-step guide for setting it all up.

As demonstrated in Figure 10, Canetis Radar [39] renders a transparent grey box in the lower half of the screen that incorporates a single red dot that is based in the center of the grey box, this dot then moves outwards towards the estimated direction of a sound source. Using the surround sound rendered by VB-Audio Cable [49] or Razer Surround [36]. Note, the Canetis Radar [39] overlay in Figure 10 has been scaled up for illustration purposes.

However, since it only has a single dot to point out sounds with, it’s unusable if there’s multiple sounds effects playing simultaneously. Something that is made worse since it is unable to filter out background noise, which results in even more interference with the directional feedback.

3.4.3 Sonic Radar

The one third-party application commercially available is called Sonic Radar [3]. However, it is exclusively available to those who have a motherboard equipped with a SupremeFX audio solution [2]. As demonstrated in Figure 11 there is a directional arrow that points in the direction of the sound, as well as a radar that also points to the direction of the sound except in a top-down fashion where the center point is the player. Note, in Figure 3 the sound is coming from a ticking bomb behind the player and the overlay has been scaled up on the right side for illustration purposes.

Sonic Radar [3] faces similar issues or limitations as Canetis Radar [39], where the indicator can only point to a single sound source at a time. However, it adds a feature that makes the visual overlay react solely to sounds within certain frequencies (e.g. only mid-range frequencies which makes it react to footsteps but nothing else) which partially alleviates this limitation.

Therefore, these “outside-in approaches” only work optimally within very specific scenarios, where there’s only one target for the visualizer to listen to with no background noise that interferes with the audio output from the game. In the case of Canetis Radar or Sonic Radar [39, 3], it is
only able to point in a single direction. Unlike Awded [30], which is able to show the sound from multiple sources, but easily becomes unusable when there’s multiple audio sources nearby that end up overlapping each other in the visualizer, causing visual noise that will not help a DHH player with situational awareness.

Conclusively, these outside-in approach fail to provide any useful situational awareness for DHH players, as the overlay only works in very limited scenarios. In all other scenarios they become obfuscated due to audio noise or its inherent limitation of only being able to point in a single direction. Without even considering the consequences of adding various audio modulation to the game, such as echoes and distance, that could end up as misinformation in the visual feedback with these kinds of approaches.

Finally, there’s a potential risk that third party anti-cheat software such as EasyAntiCheat [16] or BattlEye [5] could detect this type of software and deem them as methods for cheating, which could end up with the player being unable to use the overlay for the game or even receive a ban from the game.

4 RESULTS

The participants who felt that their hearing situation did not fit with the available choices had the option to elaborate in more detail by picking “Other”, which 12 participants did. 8 of these have one-sided deafness, with no hearing loss on the other ear. 2 of these had asymmetrical hearing loss and used hearing aids to compensate. One had a ski-slope hearing loss, where the hearing loss became increasingly severe as the sounds get closer to the high frequencies. The last one has a cognitive attention disability which causes a delay in the processing of sounds. But the participant’s hearing loss is not severe to require the use of any hearing apparatuses as seen in Figure 12.

The hearing situations of the participants and when it developed is interesting to examine. The intent behind collecting this data was to see if there were any trends between different hearing situations and the feedback to a particular question or solution. However, trends did not seem to prevalent in the data. Which could be attributed to the small sample size of games or the relatively small number of participants.

We did encounter an issue with formulating this with a single question meant to address both ears, which means that the results here reflect the ear with the best hearing (the same way WHO presented it with their statistics [51]). Some granularity is lost since there is no question regarding the participants full frequency response. However, some of these were captured if the participant elaborated on it.

The types of hearing technology in Figure 13 was important to include since individuals with Cochlea Implants are technically considered deaf, because without the apparatus they are unable to hear. By including this option there were not any participants that could be misinterpreted even if they picked “Profound Hearing Loss (90 dB and greater)”, as long as they also answered what hearing technology they used. Since a participant who uses hearing apparatuses in the form of Cochlea Implants are able to hear to a certain extent, their response should be distinguishable from the participants who are deaf and do not have any hearing apparatuses.

Figure 14 demonstrates how much time per week a participant spent playing video games. Of these 55.8% played FPS and/or TPS titles consistently, whereas 34.6% only played them occasionally and the last 9.4% did not play these types of titles. The next question was if the participants played online with others and 51.9% answered that they did consistently, while 30.8% do it occasionally and last 17.3% do not play video games online at all. This
was then followed by a question regarding if they felt that they had a difficult time keeping up due to their hearing situation, of which a staggering 71.2% said yes. The reason for the large percentage that do not or only occasionally play online can be attributed to the fact that they feel like it is difficult to keep up due to their hearing situation. Only 9.6% answered no, where one had one-sided deafness and the rest had either severe-to-profound or profound deafness, which was unexpected. This was further investigated to find out whether this might be because they do not play online or avoided playing FPS and/or TPS games but that is not the case. By all accounts these 9.6% are all casual gamers (most of these reported playing 5-15h per week) who play online to some extent. The remaining 19.2% reported that they could only keep up sometimes, which most participants reported to be game-dependent. Saying that it was depending on whether or not the game relied heavily on sound for situational awareness, many also pointed out that communication is difficult when the game includes a voice chat which becomes the default for hearing players. The ones with one-sided deafness could not keep up because they had a limited directional hearing on one side, whilst all the sounds from the deaf ear are lost.

For the last question in this section the participants were asked whether they felt like the currently available solutions were adequately evening out the gap between DHH players and hearing players, of which only 3.8% answered yes. 38.5% answered that they were not aware of a game that presented an adequate solution, and the remaining 57.7% stated that only some games offer a solution that is helpful. The participants were given the opportunity to elaborate in a open-ended fashion here which provided some interesting insights. Most of which are not relevant to the research but will be presented regardless as they highlight some important issues, despite being out-of-scope for this thesis.

Some of the players who answered “Yes” or “Some do”, explain that this is mostly due to the fact that the games they play on a regular basis do not rely on audio cues, such as MMORPGs. Since these games do not usually include voice communication and do not heavily rely on situational awareness. The combat indicators in these types of games are more commonly a combination of audio, visual and text cues.

A couple of the participants stated that they would like to know whether a game is reliant on sound cues or if it provides accessibility options before making a purchase, suggesting that the developers provide this or similar information on the game’s store page, cover or website.

The participants with one-sided deafness stated that they would like to see a mono sound option, this would remove the directional sound for the non-deaf ear but they would get all the sound from the deaf ear as well, the directionality can instead be replaced with a visual substitute for audio. Other participants who did not have a severe hearing loss but were hard of hearing stated that they would like to see individual sound level sliders for different sound tracks (e.g. one for voice lines, one for ambient sound, one for
music etc.), since their main issue was concerned with sounds getting muddled together as they had a hard time distinguishing multiple sounds in the audio mix of a game. One participant also raised the concern whether a visual representation for sound will lose some aspects that sound provides, such as setting the mood or feeling of a game.

Lastly, a participant brought up Apex Legends [38] which includes plenty of accessibility options for DHH players and has been elaborated upon in the 3.3 Existing Solutions section.

Most participants agreed that the solution presented in Minecraft [32] was helpful but also lacking in certain aspects, as seen in Figure 15. The two most common complaints were that the offset position of the text box caused the player to constantly look away from the center of the screen (which is the point of focus in the game), some also complained that it was too small. The second complaint was that the visual aid was cluttered with unimportant information, such as the main players footsteps. Another complaint that was brought up was that all subtitles were white, where the ones representing threats could be a different color such as red. The participant who was omitted from this question failed to understand the full functionality, and failed to notice the directional indicators and ended up thinking it was essentially the same as Closed Captions [48].

Fortnite’s [17] solution, as seen in figure 16 was pretty evenly split among those who found it helpful and those that thought it could be improved, regardless of whether or not they had played the game before.

The main distinction being that the participants who picked “Somewhat, it could be improved though” found that the HUD was difficult to distinguish from the rest of the game at times due to the colors of the HUD being semi-transparent. One participant suggested that they should make it more opaque or otherwise easier to distinguish. Another participant raised the concern that this implementation is arguably more accurate than hearing and thus could become necessary for all players to use. The six participants that were omitted from this question failed to see the visual overlay of Fortnite [17], which could be attributed to the embedded video on Google Forms [23] being too small or the overlay being too transparent as mentioned above.

Finally, to see how the feedback compared based on a particular hearing situation, and if that had any correlation with their response a stacked column chart was compiled. As seen in figure 17 and figure 18. The one interesting trend found was how the “Profound Hearing Loss” group answered mostly “Yes, it helps with situational awareness” with Fortnite, but said mostly “Somewhat, it could be improved though” with Minecraft. Which was also the only group whose change in answer was large enough to draw conclusions from. However, due to the small sample size no conclusions can be made with the remaining hearing groups.

Despite the small sample size of the survey, the results strongly indicate that there is a lack of accessibility options for the DHH demographic of gamers, as they do not feel like they are able to keep up with the average player due to the lack of situational awareness. Brought on by the
Fig. 17. Responses based on hearing situation with Fortnite. Note, the responses have been abbreviated for clarity reasons.

Fig. 18. Responses based on hearing situation with Minecraft. Note, the responses have been abbreviated for clarity reasons.

fact that they are unable to hear the game, or otherwise unable to process directional audio (e.g. one-sided deafness) or specific frequencies of audio (e.g. “Cookie Bite Hearing Loss” [26]).

With these results RQ4 is believed to have been answered, both in terms of considering all the individual opinions but also in regards to the attempt of finding a general consensus concerning visual alternatives that help with situational awareness according to DHH players.

5 DISCUSSION

This section discusses the content and seeks to conclude whether it answers the research questions. This section is split up in the same order as the research questions.

Taking into consideration that there are many variations in terms of a player’s hearing situation, one should come to the conclusion that the individual would either turn on the accessibility option that enables situational awareness, or is still able to hear sufficiently that they don not feel the need to enable it. However, the accessibility option should still have a positive impact on the game experience regardless of hearing situation. Whilst still retaining the mood/feeling of a game even without the sound, at least as much as possible.

5.1 Visual Substitutes of Sounds

The first research question, RQ 1, which explores the different visual substitutes for sound that help with situational awareness for DHH players currently available.

However, there is no established standard for this type of accessibility option in video games, instead game developers have had to come up with their own creative solutions for it. Something which is apparent when studying and comparing the various implementations available.

Closed Captions [48] like the one used in Half-Life 2 [46] were included as a baseline, it does not provide any situational awareness but it does portray all the sounds that are happening in the game. This could be a viable option if a game does not require situational awareness but has a rich soundscape to portray. Allowing DHH players to still be able to appreciate the tone of the game the way that the developers had intended, at least as close as possible.

The most viable and competent solutions currently available are the game-specific solutions included in certain games, such as the ones analyzed in the 3.3 Existing Solutions in Video Games section. Making them a good starting point for other developers to iterate on. For developers that want to create a visual substitute for audio intended to help DHH players with situational awareness, whilst still utilizing realistic sound such as echoing and sound occlusion. The visual overlay should be integrated with an audio engine such as VERA [29] or a built-in option like the one included with Unreal Engine 4 [18], to avoid the problem of showing inaccurate indicators for sounds that have been manipulated by the virtual environment (e.g. echoes).

An "outside-in" approach would be the most beneficial if it were not for its inherent issues with noise that obstructs
or presents incorrect information. This is partially alleviated with Sonic Radar [3] and how it is able to separate certain frequencies. However, a more promising approach could be to use machine learning algorithms similar to what AwdedTrainer [31] sets out to achieve, which could possibly result in a visual indicator that could be functional after extensive machine learning training. Where the algorithm is able to filter out the unnecessary sounds and only represent crucial sound cues on the visual indicator, maybe even understand the sound modelling effects such as echoes and still represent the sound cues in a visual indicator correctly regardless of the sound modulation that has been added. However, it is important that this is achieved without introducing any noticeable lag. “Outside-in” approaches could greatly benefit from recent technology innovations such as spatial sound (e.g. Windows Sonic [50]) or object-based surround sound (e.g. Dolby Atmos [15]). Which would provide greater detail as to the 3D positions of audio from the games, but since these options are proprietary and locked down it is currently not an option. The reason that a “outside-in” approach would be the most beneficial is because it could become a universal solution that would be compatible with most games. That way game developers would not have to worry about this accessibility concern. But all of the “outside-in” approaches would need major overhauls on the visual overlay component, since most of the currently available ones face major limitations, as described in the 3.4 Existing Solutions in Software section.

5.2 Drawbacks and Possibilities

The second research question, RQ 2, focuses on drawbacks and/or the limitations of the various solutions elaborated on in this paper, if present.

As has become apparent throughout the duration of this research, most of the currently available solutions have some kind of problem, limitation or drawback. “Outside-in” approaches more so than game-specific solutions.

The major issue found with the radar-like solutions, Canetis Radar [39] and Sonic Radar [3] are the fact that they are only able to point towards a single sound source at a time, making them erratic and unreliable when there are multiple sound sources around the player.

The major issue with equalizer-like solutions, those that have audio equalizers for left and right corresponding to the same side of the monitor, such as Final Fantasy XIV [40] and Awded [30], is that the player is unable to distinguish between all the sounds that are being represented. Because they all overlap, the player is unable to distinguish the number of targets. Awded [30] also has the issue of overlapping parts of a game, obstructing the view which could block important UI elements of a game since the overlay is not transparent.

Furthermore, all of the solutions [30, 40, 3, 39] above lack the ability to separate the Noisy Sound Effects from the Critical Audio Cues. Either because multiple sounds overlap as is the case with Final Fantasy XIV [40] and Awded [30], or because the indicator is only able to point in a single direction as is the case with Sonic Radar [3] and Canetis Radar [39]. However, this is slightly improved with Sonic Radar [3] due to its ability to limit which frequencies to focus on, but still not adequate.

The “outside-in” approaches [3, 39, 30] are also prone to displaying misinformation if the game simulates realistic sound behaviour (e.g. echoes) [18, 29].

Finally, the “outside-in” approaches [3, 39, 30] also run the risk of being detected by third-party anti-cheat software [15, 16]. Any application that overlays a game while collecting data from it faces the possibility of being seen as malicious software and can cause the user to receive a ban from the game.

After reviewing the results of the surveys, it became apparent that there were issues and limitations among the two games presented as well, the most common complaint for Fortnite [17] was that because of the color scheme it was hard to distinguish the [HUD] elements of the visual overlay from the rest of the game. Some participants also found that it sometimes difficult to immediately tell where a sound comes from if it is near the center of the screen, since the overlay is better at showing directionality outside the circle. In the case of Minecraft [32], most participants said that the offset position of the Subtitle text box was
unfavorable. Participants also pointed out that the issue of all text items being white, when the threat-related ones could be of a different color such as red, similar to Half-Life 2 [46]. Lastly, participants pointed out that certain sounds were included in Minecraft’s subtitles [32] despite not being helpful (e.g. the players own footsteps).

Tobii eye-tracking [44] is a combination of hardware and software that allows a player to interact with their computer using their gaze. This hardware is widely used for research [43] and when making software and hardware for people with disabilities [45]. By using something like Tobii eye-tracking [44] it would possible to extend the functionality of the visual overlay, e.g. a player could be looking at a visual indicator that is pointing towards a sound and be provided with more detailed information. This could be a possible solution to Fortnite’s [17] issue with displaying sound sources that are near the center of the screen.

Several participants of the survey voiced concerns about whether a visual substitute for audio that helps with situational awareness could end up being more accurate than hearing, thus becoming a necessity for all players to enable the accessibility option.

To summarize, the most common issues faced by the "outside-in" audio substitutes include the isolation of important sound queues from sounds that are not vital to improving situational awareness. For example sounds that contribute to the ambience of the game but does not necessarily inform the player of what is in the environment.

It is also important to have a non-obstructive UI overlay on top of a game that clearly informs the player of the directions of the sounds around them, without obscuring the UI elements of the game itself.

A potential solution to the issues of filtering out the crucial sounds in the audio mix could be solved with machine learning.

Eye-tracking could be promising as a way to extend the functionality of the visual overlay, for example by focusing eyesight on a particular UI element that represents a sound source from the environment, that UI element could then expand to display more details.

5.3 Viability of Visual Substitute

The third research question, RQ 3, focuses on research that support the various solutions elaborated on in this paper.

There were a total of three research papers [11, 7, 24] elaborated on in the Related Research section of this thesis, two of which were related to peripheral visions [11, 7] and the last one comparing video game players with non-video game players [24].

The first paper that explored deaf individuals peripheral vision focused on the capacity of the visual field [7], and found that the deaf observers had significantly larger visual fields than video game players [7].

The second paper that explored deaf individuals peripheral vision focused on the reaction times in the periphery portion of the visual field [11], this paper did comparisons between deaf, BSL interpreters and hearing non-signer adults. The results of this paper showed that deaf adults demonstrated significantly faster reaction times than both hearing non-signers and hearing BSL interpreters [11].

The results from these two studies [11, 7] indicate that a visual overlay seems to be a viable option as the medium for substituting audio for situational awareness.

The last piece of research showed that habitual video game players performed better compared to non-video game players [24] when it came to attentional blink, which looks for the ability to avoid "bottlenecking" an individuals attention during temporal processing. The video game players also performed better when it came to the enumeration of objects that were flashed only for a brief moment, as well as having extended attentional capacity compared to non-video game players. There was also a test that only the non-video game players participated in, where they played a game routinely for 10 consecutive days and showed a significant performance difference at the end of it in comparison to where they started [24].

The results from the study with habitual video game players [24] can be placed in contrast with the negative responses from the survey, since a prolonged exposure to using a visual overlay might make it easier to use as a player becomes adjusted to it.
5.4 Players View

The last research question, RQ 4, concerning what the general consensus of DHH players is regarding various solutions that were answered in the results section.

After analyzing the surveys thoroughly a conclusion was reached that there is currently no general consensus on what type of accessibility solution should be the industry standard, this could be due to the relatively small sample size, or because the survey was conducted online as opposed to having an interview were followup questions can be made to clarify answers.

The general opinion about Minecraft’s solution, named “Subtitles” is that although helpful with the players situational awareness, it has a few issues. The most common complaint being the offset position and the size of the text box being to small, which distracted the player by making them have to look away from the center of the screen. The second complaint was that the text box became filled with unimportant information such as the players own footsteps.

As previously stated in results, the solution presented by Fortnite had an even split among those who thought that it was helpful in providing the DHH players with situational awareness and those that thought it was only somewhat helpful but had room to improve. The most common complaint was the overlay being too transparent and the choice of colors that made the HUD elements difficult to see over bright areas in the game.

Although RQ 4 is examined based on a limited set of games, some conclusions from the responses can still be extracted and applied to a different implementation. For example, Fortnite’s had an even split among those who thought that it was helpful in providing the DHH players with situational awareness and those that thought it was only somewhat helpful but had room to improve. The most common complaint was the overlay being too transparent and the choice of colors that made the HUD elements difficult to see over bright areas in the game.

This is a complicated accessibility concern, where situational awareness is only one aspect to consider when making a game accessible for DHH players. Where other aspects such as team communication also need to considered, or how to set the intended mood in a way that DHH players can appreciate.

Apex Legends includes a ping system, that allow a player to ping a certain objective or target for their teammates to see, all essential communications can be done using the ping system. The game also includes a speech to text and the reverse of text to speech so that deaf players can type to hearing players who will have the message dictated and the voice communications get translated into text for the deaf player to read.

Fortnite include different animations in the visual overlay that could be considered attempts at recreating a certain mood, e.g. gunshots are given a red tint together with a rapid wave animation.

6 Conclusion

The research supports the notion that a visual overlay is a viable option for representing audio that intend to provide situational awareness.

“Outside-in” approaches fail to sufficiently help with situational awareness in their current state, because they only work in very limited scenarios.

With the game-specific solutions found in Fortnite and Minecraft being the most competent implementations. With Fortnite showing some simple attempts at trying to translate the intended mood of the sound equivalent into animations and colors on the indicator.

Despite the game-specific solutions being fairly competent, there are aspects to these that could be further explored. Like adding eye-tracking to extend the functionality of a visual overlay.

For “outside-in” approaches it is worth exploring the use of machine learning to potentially counter the issues inherent to this type of implementation, the quality of the sound information could also be improved if there was a way to access or decode the data from spatial or object based surround sound technologies.

It could also be worth exploring the viability of haptic feedback as an alternative to help DHH players with situational awareness.

But for games that do not rely on situational awareness it could still leave a DHH not being able to fully appreciate
the game, especially if the game has a rich soundscape or is heavy on dialogue, something that could be remedied with Closed Captions [48].

World Health Organization estimates that 6.1% of the world population has a hearing disability [51], with projections indicating that these numbers will continue to rise. This is not an insignificant number and it is why we believe this accessibility concern should be seriously considered when making a video game.

As mentioned in the Introduction there are resources available for developers that can help with making their games more accessible. Such as AbleGamers Accessible Games [1] or the Game Accessibility Guidelines website [20].

The easiest way for a developer to start when it comes to making their game accessible to DHH players is to simply mute all the audio from a game, and test out their game that way, whilst considering the experience and mood of the game and whether it remains the same or similar enough with the various substitutes for the audio aspect.

Everyone should be able to appreciate the implemented solution no matter the hearing status of a player, the solution should be a positive inclusion to the game experience.

However, it is important to know that situational awareness is only one aspect that needs to be considered when making a game accessible for DHH players.

6.1 Future Work

This thesis covered a wide range of games and will hopefully help developers who want to make their games more accessible. However, this field could certainly use more research and what follows are some suggestions.

It could be worth exploring how eye-tracking similar to Tobii [44] could enhance a visual overlay for situational awareness in the future.

It could also be interesting to create a game that can compare the reaction times between audio-based sound cues and visual cues. Furthermore, multiple variations of visual substitutes could also be compared within the same game (perhaps substitutes identical to the ones found in the games mentioned in this paper), both in terms of reaction times but also other factors such as usability, readability and more.

7 Acknowledgements

We would like to extend a special thanks to our supervisor, Jeanette Eriksson, for all the supervision and guidance she has provided. A special kudos for all the help and feedback she provided during the final week leading up to the deadline.

We had also like to take the opportunity to thank everyone who participated in our survey, with a special kudos to those who spent the extra time elaborating their thoughts and feedback. This helped us plenty as some of the responses included software and games that did not show up during our own searching.
DEFINITIONS

Accessibility Layer
A layer added on top of a game in an effort to alleviate a particular handicap or disability.

Audiogram
A graph that shows the audible threshold for a standardized frequency range as measured by an audiometer.

Ban
To be forbidden access to something, in this paper referring to the act of being blocked access from a game.

Bottleneck
A choke point, a specific point that becomes a limiting factor that severely affects the maximum potential performance.

British Sign Language (BSL)
A sign language used mainly in the United Kingdom.

Critical Audio Cues
The critical sound effects in a game that are valuable in terms of informing a player's situational awareness, an example being enemy footsteps.

Deaf or Hard of Hearing (DHH)
Deaf or Hard of Hearing, an individual with a light to severe hearing loss.

Decibel
A type of measurement, in this paper talking about sound levels. The measurement is logarithmic which means that an increase with 10dB results in sound that’s 10x louder.

First Person Shooter (FPS)
First Person Shooter, a type of game that involves shooting, seen through the eyes of the player controlled character. Example: Counter Strike: Global Offensive [47].

Haptic Feedback
Haptic feedback refers to technology which can create an experience of touch by applying forces, vibrations, or motions to the user.

Head-Up Display (HUD)
An area of the screen where the player can retrieve important information. Information can for example be the player character’s health status and resources.

Massively Multiplayer Online Role-Playing Game
A multiplayer online role-playing game with a persistent world where a large number of players play simultaneously.

Noisy Sound Effects
The sound effects that can obfuscate Critical Audio Cues when visualizing the sound effects, for example background noise like rain.

Open-source
This is the term used when the original source material has been made freely available and may be redistributed and modified.

Outside-in Approach
Where a program works as a separate entity, in this paper referring to when a separate program attempts to create visual indicators based on the audio output from a game.

Playerbase
The active community of players that play a particular game.

Sound Directionality
A specific direction from which a sound originates.

Third Person Shooter (TPS)
Third Person Shooter, a type of game that involves shooting, seen from behind the player controlled character. Example: Fortnite [17].

Third-party
An unrelated party, in this paper referring to a piece of software that can be applied to a game but is created by a developer that isn’t associated with the game.

Virtual Reality
Virtual Reality is the use of computer technology to create a simulated and immersive environment in a seemingly real or physical way.
REFERENCES


[22] GitHyp. Current Video Game Statistics for the Current-
PC [Game] (ver. 1.36.9.4), 2012.


ATTACHMENTS


RESPONSE LETTER

7.1 Pre-presentation

The feedback received from the pre-seminar version have all been addressed, either by minor adjustments or completely rewritten sections, since the thesis changed focus.

- First matter to address is the change in scope, we realized that doing a case study of all available solutions for situational awareness as well as creating a custom artifact for that was too big of a scope. Instead we decided that it would be best if these two were separate research topics. Where we decided to pursue the case study aspect which could serve as a base for future research that aims to implement an artifact.

- All feedback pertaining to the artifact in our original scope is therefore no longer relevant, this includes design research, testing and design criteria for the artifact.

- Research Questions and motivation for the thesis has been added in the Introduction as requested.

- The games that were presented to the participants of the survey has been clarified as requested, it is specifically Minecraft and Fortnite from the Existing Solutions in Video Games.

- The methodology has been expanded as requested, it now includes a description of the test phase as well as all other actions and decisions, and the motivation behind these.

- The areas that needed clarifications have been clarified and made clear.

- Both surveys have been added to the Attachment section, the questions in the survey are elaborated on in the Procedure section in Methods & Work Process. Result section has been added in full.

- The references that were missing in the Introduction have been added.

7.2 Post-presentation

The feedback received from the examiner, supervisor and opponents of the presentation version have mostly been
addressed, some of the comments have been omitted with an explanation attached below.

- The feedback on having too few people in the survey, which allegedly makes it difficult for us to draw conclusions and should have included a margin of error calculation, was omitted. Allegedly this would give more credibility, but we disagree and think that knowing how many participated is good enough for a reader to judge the precision of our responses.
- The document has added titles and sub-sections where it was suggested, as well as renaming some titles and readjusting certain parts to improve the flow of the text.
- The comment about presenting the data in a different format than pie-chart’s was omitted due to us judging it to be the most effective way of portraying the data.
- The comment about not presenting accurate numbers and instead writing ”some participants” have been rectified.
- Some parts of the text were removed, clarified and/or cleaned up as commented due to being deemed unnecessary, vague, incoherent or repetitive.
- The fonts and colors in the figures were adjusted based on feedback.
- Due to the nature of this research document it’s unavoidable to not have long sentences, thus that comment was disregarded.
- Minor spelling and formatting errors were rectified as commented.
- The comment about adding the question relevant to a certain figure in the caption was disregarded, we deemed that it would create repetition and make the caption cluttered.
- The overuse of commas in certain sentences have been rectified.
- The comments about certain ideas for Future Work have been added to a dedicated Future Work section found in the Conclusion, as well as moving other suggestions for future work we mentioned in other parts into a dedicated section.
- Comments about RQ4 were omitted because we do include in the text which specific games were surveyed.
- The second paragraph and onward in the Abstract has been completely rewritten as suggested.
- The comment about being interested to see the difference between the players who had played the games before and those who hadn’t was omitted, instead we clarified in the text that there was no tangible difference between the two groups.
- Charts and descriptions were included to display any trends within certain hearing situations as suggested.
- The Main Field text on the title page has been corrected.
- Misplaced references have been fixed.
- Added a dashed line displaying unimpaired hearing in figure[1] as commented.
- Text portions that referenced other sections now have the correct number reference attached as well, to make it easier on the reader to find it.
- Added a summary to the drawbacks and possibilities section.
- Moved the section of how many responses were included in the result data to an appropriate section, instead of having it at the start of the results.
- Sections with personal pronouns have been rewritten as suggested.
- A section of new information to the reader found in the conclusion has been moved to the introduction, since it isn’t relevant to our work but is related (websites where developers can find tools and help with making their games more accessible).