Improving the Player Experience of Collaborative Multiplayer Games for Visually Impaired Children

Marc Ferriggi
1003879

August 2023

Master Thesis

Utrecht University
MSc. Game and Media Technology
Graduate School of Natural Sciences

Supervisors:
Prof. Anja Volk (first)
Dr. Hanna Hauptmann (second)

Keywords: visually impaired children, inclusive games, accessibility, music information retrieval, beat tracking, rhythm games, audio feedback, dynamic difficulty adjustment

Abstract

Visually Impaired (VI) children face barriers when it comes to social engagement and integration, and engage in less cooperative play compared to sighted children. Collaborative multiplayer computer games can be used as a tool to encourage children to develop their social skills. However, there is a lack of research on developing collaborative multiplayer games for the VI. This thesis provides a deeper insight into developing such games by building on the work done in a previous study which developed the game *Bongo Beats: Tap With Me*. The game was improved by implementing an algorithm to easily add more songs to the game using beat tracking techniques, improving the quantity and quality of the feedback given to the players, and implementing an adaptive difficulty algorithm to ensure the game is engaging for players of any skill level. Pilot studies were conducted to verify the quality of the implemented features, and the final experiment was done at Bartiméus school for VI children where insights were gathered on how the VI children perform, collaborate, and experience the new version of the game. The results of the experiments have shown that the children enjoyed the game and collaborated well in their teams. The insights provided in this thesis further show that it is possible to create collaborative multiplayer games with accessibility features for the VI that are also enjoyed by a sighted audience.

Acknowledgements

I would like to start by thanking my supervisors Anja Volk and Hanna Hauptmann for always being available to provide feedback and guidance when needed. This thesis would not have been possible without their support.

Furthermore, my warmest thanks and gratitude goes out to Myrthe van der Heijden whose help and patience with the children at Bartiméus was invaluable and who selflessly offered to help all the children fill out the questionnaires on the day of the experiment. I would also like to thank Geertje Peters Rit for continuously helping me with ideas and feedback during the process of working on this thesis and for translating the information sheet and consent form into Dutch.

I would also like to thank everyone involved at Bartiméus for their participation in the experiments. A special thanks goes to Dick Lunenborg who helped schedule the experiments and who provided his office for the experiments to take place. Furthermore, I want to thank all the students who participated in the pilot studies, and everyone involved in the expert interviews whose feedback proved invaluable.

Finally, I would like to thank my close friends and family for supporting me during the writing of my thesis.

Table of Contents

1 Introduction and Background **6**
1.1 Literature Research Protocol 8

2 Related Work **9**
2.1 Music Information Retrieval 9
2.1.1 Beat Tracking 9
2.1.2 Vamp Plugins and Software 12
2.2 Games for the Visually Impaired 13
2.2.1 Designing Games for the Visually Impaired 14
2.2.2 Game Feedback 17
2.3 Collaborative Play 18
2.4 Adaptive Difficulty 20
2.4.1 Difficulty in Rhythm Games 24

3 Research Questions **25**

4 Methodology **26**
4.1 The Game 26
4.2 Developing the Game 27
4.2.1 Developing an algorithm to generate a level given a particular song 28
4.2.2 Improving the quality of feedback given to the player during gameplay 31
4.2.3 Implementing realtime DDA 33
4.3 Experiments 35
4.3.1 Pilot Studies 35
4.3.2 Final Experiments 38
4.3.3 Expert Interviews 38

5 Results **39**
5.1 Pilot Studies 39
5.1.1 Research Question 2 39
5.1.2 Research Question 3 41
5.1.3 Research Questions 4.1 and 4.2 45
5.2 Final Study 52
5.2.1 Research Question 1.1 52
5.2.2 Research Question 1.2 54
5.2.3 Research Question 1.3 56

5.3 Expert Interviews 59
5.3.1 Collaborative Game 60
5.3.2 Game for the VI 62
5.3.3 Rhythm Game 63

6 Discussion **65**
6.1 Results - Implemented Improvements 65
6.2 Results - Final Experiment 67
6.3 Results - Expert Interviews 68
6.4 Collaboration vs. Rhythm 68
6.5 Target Age Group 69
6.6 Limitations 70
6.7 Future Research 71

7 Conclusion **72**

References **75**

Appendix **81**

1 Introduction and Background

In the Netherlands, three-quarters of visually impaired (VI) children follow education programs at traditional schools (Smeets and Boer, 2019). Despite this, VI children are still facing barriers when it comes to integration, participation, and social engagement (Metatla et al., 2020). In fact, VI children in mainstream schools engage in less cooperative play with their peers when compared to sighted children, and thus tend to play in a more solitary fashion (Hestenes and Carroll, 2000). There are plenty of examples of technologies that assist and help VI children develop their social skills and participate in group play (De Schipper, Lieberman, and Moody, 2017; Verver, Vervloed, and Steenbergen, 2020), and a typical example of this are computer games. However, there is a lack of studies on collaborative multiplayer games for VI children to play with sighted children (Potamianos, 2022).

This is highlighted further in a focus group study with qualified teachers of children with visual impairments (QTVI) and teaching assistants (TAs) of students with visual impairments in mainstream schools run by (Metatla et al., 2020). The authors found that students with particularly severe visual impairments require assistance to develop and maintain meaningful social engagements with their peers and they go on to state that "there was not enough support for inclusive play between children with mixed visual abilities. That is, once VI children do find sighted friends to play with or vice versa, there are not enough options engendering engaging play experiences available to them from that point forward."

In a preliminary study on developing inclusive collaborative multiplayer games for the VI, (Potamianos, 2022) has developed the game *Bongo Beats: Tap With Me*. This was done in collaboration with Bartiméus school for VI children. The game is a VR rhythm game inspired by *Beat Saber*[1], where one player is the conductor who receives information on what notes to play, and the other player is the musician who plays the notes. The initial study shows some promising results; the players collaborated well with their teammates and generally enjoyed the game. Furthermore, the visually impaired children in the study showed a willingness to keep playing the game and stated that it positively affected their mood. However, there were several limitations to Potamianos' study which this thesis seeks to improve on.

The most common themes in the interviews with the participants of the study in (Potamianos, 2022) were reduced enjoyment of the game as they did not like the song that played throughout the level and the lack of feedback given to players in the game. Potamianos suggests using beat tracking to

[1] https://beatsaber.com/

create an algorithm that can easily add many different songs to the game in order to address the first issue. In this study, *Bongo Beats: Tap With Me* was altered to address these issues, and the aforementioned issue mentioned by (Metatla et al., 2020) with regards to games for the VI not being engaging for sighted children was also addressed. This was tackled by implementing dynamic difficulty adjustment into the game to alter the difficulty of the game based on the players' performance and to ensure that the game will remain at the right level of engagement and challenge.

Thus, the main focus of this study is to improve the game based on the feedback and results found in (Potamianos, 2022) and compare the results to the previous study to gain a further understanding of developing collaborative multiplayer games with accessibility features for the VI. This was achieved by improving the experience of the game based on the feedback given by the children, observers, and experts that were present for the tests in the study in (Potamianos, 2022), and based on the future work section of the same study. In particular;

1. Beat tracking techniques will be used to develop an algorithm to simplify the process of adding new songs to the game as this was previously done manually where each note was placed by hand. This will thus make it easy to increase the number of songs in the game to ensure that every player can find a song that they enjoy.

2. The quantity and quality of the feedback given to the players while they play the game will be improved, directly addressing the comments made by the participants in (Potamianos, 2022) regarding the lack of feedback.

3. An adaptive difficulty algorithm will be implemented in the game to make the game more challenging for participants finding it too easy, and easier for participants finding the level too challenging thus ensuring that the game is more engaging for all players, regardless of their abilities.

This master thesis provides deeper insights into designing collaborative multiplayer games for the VI. With this information, I hope to help bridge the quality gap between games for the VI and mainstream games, as well as provide a collaborative multiplayer game for the VI that can be accessed and enjoyed by anyone. It should be noted that due to the scale of this master thesis, it is not possible to create a game of the same quality as a current mainstream game on the market, however, the insights provided in this study will help developers make better decisions with regards to accessibility.

The remainder of this section will cover the literature research protocol used for this study. Following this, section 2 will cover the literature review. In particular, the field of beat tracking in MIR is first covered in section 2.1 as this will be used to add more songs to the game as suggested by (Potamianos, 2022), following that a detailed study into games for the VI is presented in section 2.2, and a study of adaptive difficulty in games is presented in section 2.4. The research questions for this master's thesis can be found in section 3, and the methodology is then presented in section 4 where the technical specifications of the changes made to the game are explained in detail in section 4.2, and details of the experiments can be found in section 4.3. The results of the experiments are presented in section 5, and this is followed by a discussion of these results in section 6. Finally, a conclusion summarizing the findings of this thesis is given in section 7.

1.1 Literature Research Protocol

There are three main research areas where literature needed to be gathered, mainly in the areas of music information retrieval (MIR) focusing on beat tracking, games for the VI and accessibility, and adaptive difficulty in games. The main way research literature was collected for this project was through Google Scholar. For the research on beat tracking, the following keywords were used in the search; "music information retrieval", "beat tracking", "beat tracking AND deep learning", "beat tracking AND PLP", "beat tracking AND online", "beat tracking AND real-time", and "music based procedural content generation for games". Regarding games for the VI, the following keywords were used in the search; "games for the visually impaired", "visually impaired children AND collaborative games/play", "accessibility in games", "audio feedback in games", "audio games", and "accessibility guidelines for games". Regarding adaptive difficulty, the following keywords were used in the search; "adaptive difficulty", "dynamic difficulty adjustment", "adaptive difficulty AND rhythm games", and "player modeling". Note that when looking for the latest developments in the field, a filter was sometimes set to only include recent years (from 2017). The relevant papers were found and the abstract, introduction, and conclusion were read before deciding if the paper is relevant to the study.

Another way of collecting information was through papers recommended by my supervisor Anja Volk, and from the previous study on this topic in (Potamianos, 2022). All these papers usually led to other interesting papers from their literature review. Thus plenty of sources were also gathered from the references of papers that were deemed relevant to this study.

2 Related Work

In this section, related work is presented. Each section will cover research related to the different aspects of the game that was improved in order to contribute to an increased player experience. In subsection 2.1, the field of Music Information Retrieval (MIR) is introduced with a specific focus on Beat Tracking as this was used to add more songs to the game as recommended by (Potamianos, 2022). Subsection 2.2 will cover the current state-of-the-art on research on games for the VI, including research on designing games for the VI, audio feedback, and collaborative play, as this was used to improve the quality of feedback given to the players as they are playing the game. Subsection 2.4 will detail related work in the field of adaptive difficulty for games, focusing on identifying the framework that was implemented into the development process of *Bongo Beats: Tap With Me*.

2.1 Music Information Retrieval

The term MIR first appeared in literature in 1966 from an academic paper by Michael Kassler (Kassler, 1966). Although presented as a programming language for the IBM-7094 computer, MIR has since grown into a multidisciplinary field of research and is behind a vast array of technologies available today. According to (Burgoyne, Fujinaga, and Downie, 2015), the tasks involved in MIR can be classified based on the types of input and output data of the system. For this study, we are using audio files as input for our MIR system and are concerned with the sequence labeling output form. In particular, we are interested in beat/bar tracking as this has been identified as a technique to be used to easily populate *Bongo Beats: Tap With Me* with more songs (Potamianos, 2022).

2.1.1 Beat Tracking

Beat Tracking systems were defined in (Goto and Muraoka, 1994) as systems that “recognize temporal positions of quarter notes” in a given audio signal and “organize music into a hierarchical beat structure” (Goto and Muraoka, 1999). In other words, the task of a beat tracking algorithm is to indicate the sequence of beat instants in the signal where a human listener would tap their foot (Davies, Degara, and Plumbley, 2009; Ellis, 2007). The task of beat tracking from audio files is non-trivial, and some difficulties defined by Goto and Muraoka in (Goto and Muraoka, 1994) include:

1. Audio files consist of sounds containing different instruments.

2. Energy peaks in the signal do not necessarily correspond directly to a beat.

3. A beat may not directly correspond to a note onset or a real sound.

4. It is difficult to distinguish a strong beat from a weak beat.

Furthermore, Ellis describes two constraints that must be satisfied when developing an accurate beat tracking system; the beat instants should correspond to the moment in the audio when a beat is indicated (i.e. by a note onset from one of the instruments), and the beats should have regularly spaced locally-constant interval (Ellis, 2007).

Some early applications for beat tracking systems included video editing, audio editing, stage lighting control, and other music interaction systems (Goto and Muraoka, 1994). However, today the need for accurate beat tracking has increased as this is usually part of higher-level MIR tasks such as chord extraction, structural audio segmentation, and music similarity (Davies, Degara, and Plumbley, 2009). Recent years have also seen an increase in the popularity of rhythm games, and beat-tracking systems are also proving to be a useful tool in helping to speed up the development process of these games, or to procedurally generate the game world from a music signal (Meier, Schwär, et al., 2022; Yeh and Jang, 2019; Yeh, H.-H. Li, and Roger, 2014).

In the remainder of this section, different approaches to solving the beat-tracking problem, and state-of-the-art beat-tracking algorithms will be discussed.

Early Attempts

The first attempts at beat tracking on audio files in particular were done in (Goto and Muraoka, 1994). Prior to this, studies were only made on MIDI files. The proposed algorithm used frequency analysis to achieve two objectives, firstly to detect the note onset times and secondly to detect bass drum and snare drum sounds. Multiple agents running different hypotheses are then used to calculate the inter-beat-interval (IBI) and predict the next beat time. The agent with the strongest hypothesis is then chosen and generates the Beat Information (BI). The algorithm assumed a 4/4 time signature and constant tempo. Furthermore, Goto and Muraoka's algorithm relied on the presence of drum sounds, this method was further refined and extended in (Goto and Muraoka, 1999) and (Goto, 2001) to remove this reliance.

Deep Learning Algorithms

Many recent studies and state-of-the-art technologies with respect to beat tracking have seen an increase in reliance on deep learning methods. Böck and Schedl were the first to publish a paper on this where they implemented two algorithms using Bidirectional Long Short-Term Memory (BLSTM) recurrent neural networks to perform beat classification (Böck and Schedl, 2011). The first algorithm assumed a constant tempo, while the second one allowed changes in tempo and time signature. Gkiokas and Katsouros were the first to use Convolutional Neural Networks (CNN) for the task of beat tracking and with their implementation also claim that it will be possible for this algorithm to work online in a real-time fashion (Gkiokas and Katsouros, 2017). The field of deep learning for MIR tasks and beat tracking, in particular, is very rich nowadays and one can find many papers proposing different architectures, methods, and refinements to previously designed networks including (Böck and Davies, 2020; Böck, Davies, and Knees, 2019; Böck, Krebs, and Widmer, 2016; Pinto et al., 2021). (Fuentes et al., 2018) provide a good overview of some of the different approaches available at the time of writing.

Online Systems

A vast amount of research has also been made in developing online beat tracking systems. The challenge with developing online systems is that they need to be causal, meaning that the system needs to infer whether a beat is present using only present and past information, as access to future information is not possible. (Oliveira et al., 2010) were the first to develop an online beat tracking system. This approach made use of Convolutional Recurrent Neural Network (CRNN) structures in their system. Plenty of other examples using different machine learning approaches can also be found in literature, including (Heydari, Cwitkowitz, and Duan, 2021; Heydari and Duan, 2021). One drawback with machine learning approaches for causal systems in particular is that these systems are usually black boxes where their inner workings are hard to understand and cannot be easily fine-tuned to adjust the beat tracker. Thus, (Meier, Krump, and Müller, 2021) state that model-based approaches are more suitable for the task of real-time beat tracking algorithms. Their proposed system is based on an algorithm that exploits Predominant Local Pulse (PLP) information in order to extract the beat. This method was an extension of a previous study on extracting PLP information from music recordings (Grosche and Muller, 2010).

Chosen Algorithm

In order to implement beat tracking for this study, the algorithm to be used should reflect the current state-of-the-art, and provide good results for a variety of music styles. The initial idea for this study was to utilize the PLP method proposed by (Grosche and Muller, 2010), as this was used in an online fashion in the design of a rhythm game in (Meier, Schwär, et al., 2022). The source code for this algorithm however is not yet published, and after communicating with the authors of the paper, it was suggested to seek an alternative algorithm as implementing the PLP algorithm would prove to be out of scope for this study.

The 'BeatNet' algorithm proposed in (Heydari, Cwitkowitz, and Duan, 2021) has many advantages. Apart from the source code being available in an open-source Python package[2], this algorithm can be used in both online and offline ways, and returns beat, downbeat, and meter information. Furthermore, the algorithm does not need to be primed with a time signature and makes use of an information gate that reduces the computation costs of the inference step without compromising accuracy. The 'BeatNet' algorithm follows the machine learning approach to beat tracking which utilizes a neural network with causal convolutional and recurrent layers. During the inference stage, two Monte Carlo particle filters are applied to generate the beat and down-beat activations respectively. Results of this algorithm have shown the generalization ability of this algorithm to work well for multiple different genres and during testing, the accuracy of 'BeatNet' was found to be better than the state-of-the-art[3] in both its online and offline tests. The generalization ability of this algorithm to work over multiple genres and the superior accuracy of 'BeatNet' is what led to the decision to utilize this algorithm in this master thesis. Preliminary experimentation with the algorithm was also conducted where songs of different genres were used and their respective beat time-stamps were generated. These were then loaded in Sonic Visualiser (see subsection 2.1.2) and the beat times lined up well with how one would naturally tap their feet for all the songs tested.

2.1.2 Vamp Plugins and Software

Alongside choosing a beat tracking method to be used when developing an algorithm to quickly add new music and levels to the game, a tool was needed to check the correctness and validity of the results from this algorithm. The

[2] https://github.com/mjhydri/BeatNet
[3] At the time of writing in 2021

Vamp plugins and software[4] have been identified as tools that could be used during the development of the game, due to their availability and direct link to the field of MIR. (Cannam et al., 2019) lists information on available plugins at the time of writing. In particular, Sonic Visualizer[5] will be used after extracting the beats of a song to visualize and check these results, and to help in the process of adjusting the parameters[6] of 'BeatNet'.

In this section, beat tracking algorithms were discussed with the goal of developing an algorithm to quickly add multiple songs into *Bongo Beats: Tap With Me* as this is one of the ways in which the experience of the game will be improved in this study. The next section will cover related literature in the field of games for the visually impaired with the goal of understanding how the feedback given to the players when playing the game can be improved.

2.2 Games for the Visually Impaired

Visual impairment refers to reduced vision or loss of vision which cannot be corrected (Khaliq and Torre, 2019). According to (*Vision Impairment and blindness* 2022), 2.2 billion people suffer from a form of visual impairment, while (Laser Eye Surgery Hub, 2022) further state that it is thought that 237 million people suffer from moderate or severe visual impairment and 30 million people are believed to be blind[7].

Making games accessible for the VI will improve the person's quality of life, as this will allow them to participate in an activity that a sighted person takes for granted, thus reducing their feeling of isolation (or not belonging) and emotional pain (Khaliq and Torre, 2019). In a focus group study with qualified teachers of children with visual impairments (QTVI) and teaching assistants (TAs) of students with visual impairments, (Metatla et al., 2020) found that computer games can provide an outlet for supporting social engagements among children, where one QTVI stated that a lot of children are always talking about the games that they are playing. A TA also mentioned computer games can be a platform for allowing VI children (who find themselves always being monitored by an adult) to engage in "healthy mischievous behavior", stating in particular that "it's something for blind kids to be a little naughty, I caught him playing an audio game while wearing headphones when he was supposed to be using the calculator!"

One major issue when designing accessible solutions for games for VI people is that VI players cannot perceive primary stimuli, thus this has an

[4]https://vamp-plugins.org/
[5]https://www.sonicvisualiser.org/
[6]See section 4 for information on the available parameters
[7]According to data released in 2020

impact on perceiving the feedback necessary to play the game (Yuan, Folmer, and Harris, 2011). A solution to this is to design games from scratch with the intention of using audio as the primary stimulus required to play the game. These are known as audio games. There are plenty of examples of audio games available to play which can be accessed online[8]. Many studies have also incorporated the use of haptic stimuli to convey information to VI players, either through custom hardware or by making use of commercial hardware from virtual reality (VR) systems whose controllers have good haptic/ tactile interfaces. These games cover a broad range of styles. Some examples of games for the VI include:

- *Tim's Journey* (Friberg and Gärdenfors, 2004): A non-linear exploration audio-based game where you move around different soundscapes to unravel a hidden mystery.
- *AudioQuake* (Atkinson et al., 2006): A recreation of the first-person shooter game *Quake* made to be fully playable by blind gamers.
- *Speed Sonic Across the Span* (Oren, 2007): A platformer game that can be played through an audio-only interface.
- *Blind Hero* (Yuan and Folmer, 2008): A music/rhythm game that replaces visual stimuli with haptic stimuli.
- *The Enclosing Dark* (Gluck and Brinkley, 2020): An adventure-maze game using audio as the main source of information
- *Racing in the Dark* (Gluck, Boateng, and Brinkley, 2021): A fast-paced racing game that exploits the technology and availability of VR hardware to provide a non-visual racing experience.

2.2.1 Designing Games for the Visually Impaired

(Friberg and Gärdenfors, 2004) provided an early study into the design of audio games, claiming that these games could be as entertaining and engaging as the mainstream games released at the time of writing. The authors stress the importance of emphasizing the differences between different types of auditory information similar to how one would separate different interface objects in a graphical interface with borders and color coding. They achieved this when developing their own audio games by categorizing the sounds into the following groups:

[8]https://www.audiogames.net/

- *Avatar Sounds*: these refer to sounds produced by the in-game character controlled by the player and include footsteps, shooting, bumping into objects, etc.

- *Object Sounds*: these sounds indicate the presence of an object in the game world. Depending on the strategy taken, these should either be brief recurring sounds, or long continuous sounds.

- *Character Sounds*: sounds generated by the non-player characters in the game.

- *Ornamental Sounds*: These sounds are not necessary for the gameplay, but provide an aesthetic to the game. An example of this would be background music.

- *Instructions*: These are usually speech recordings that provide information to help the players perform tasks and play the game.

(Friberg and Gärdenfors, 2004) also state the importance of communicating to the player whether a sound is generated by the game world, or by their actions.

(Yuan, Folmer, and Harris, 2011) propose the use of *sonification* (manipulating audio to convey information by changing the pitch, amplitude, or tempo) techniques as a strategy to replace visuals with audio. They define the terms:

- *Auditory Icons*: using sound effects to indicate different objects or actions,

- *Ear Cons*: using tones, or a combination of tones, to indicate different objects or actions, and

- *Sonar*: using a sonar-like technique to convey spatial information in the game world.

In their study on audio games, (Garcia and Almeida Neris, 2013) list several design guidelines to consider when designing audio games for VI players. The guidelines include suggestions from setting up an easy installation process to advice on representing objects in the game. The following guidelines have been identified as being relevant to this study as they directly deal with improving the feel, feedback, and understanding of the game:

- Ensuring that all events are emanated by sound, so that the player knows what happened, what will happen, and be given appropriate time to respond to it;

- Providing immediate sound feedback from player action in order to convey to the player how the game has responded to their input;
- Allowing input controls to be customized by the player. This is done to ensure the player is familiar where the buttons are and which buttons to press;
- Avoiding playing many sounds at once, as this can be confusing to the player;
- Describing events with the most accurate sounds where possible, where this is not possible ensuring that the sounds are informative of the event that has occurred;
- When introducing a new sound in the game, present the option to the player to describe the sound (i.e. play the sound and provide an audio description of its meaning). This can be achieved by providing an interface where the user can listen to all the sounds in the game separately;
- Providing either an audio tutorial to first-time players or enough information during gameplay to help them play.

(Khaliq and Torre, 2019) give a good overview of recent approaches and solutions that should be considered when attempting to make games accessible for the VI. They categorize accessibility features into three sections; visual, audio, and tactile, and support their findings with examples from literature and case studies. The visual section focuses on designing visual filters to make games accessible for colorblind players, they stress the importance of customizable font sizes and font types and provide examples of released games where these features are implemented well, and also examples where the font of the user interface is unreadable to VI players. In the audio section, the authors show how audio descriptions can be implemented in games and game menus using a voice-over to verbally convey visual information, and show examples of games where sonification is used to convey spatial information. The tactile section focuses mainly on providing the reader with a summary of how the game *Blind hero* (Yuan and Folmer, 2008) was developed. Furthermore, the authors highlighted the game *1-2 Switch*[9] whose information is prompted through visual methods (the graphics on screen), audio cues, and haptic feedback from the controller, thus allowing players to play the game without needing the graphics displayed on the screen.

[9] https://www.nintendo.com/store/products/1-2-switch-switch/

This research has helped identify areas that can be improved in *Bongo Beats: Tap With Me*. Mainly, the following areas need to be improved to create a better experience for VI children:

- Use more distinct sounds to indicate if a note was hit correctly and if it wasn't hit correctly, ensure that there are distinct sounds to inform the player whether the note was hit too early, or too late (by making use of auditory icons).

- Create a tutorial that will introduce the player to the game, and during this tutorial introduce all the different auditory icon sounds.

The following section will cover research that focus on how feedback can be improved in the game.

2.2.2 Game Feedback

In (Potamianos, 2022), a common theme was the lack of feedback received by the player which affected the overall game experience and the players' performance, with one of the observers in the study stating

"*I feel some participants were eager to get better performance, but they failed because there weren't enough hints on whether the player receiving vibration or the player hitting keyboard being too late. I recall one participant even asked the reason of their poor performance.*"

This also affected the players' motivation to play, specifically as the blind participants could not see the score when playing. One participant even stated "*I don't know the score so no reason for extra motivation.*"

Thus when designing games for the VI, it is very important to pay special attention to developing good feedback to communicate what effect their input has on the game world. When creating games for the VI, immediate audio feedback is of vital importance (Archambault and D. Olivier, 2005). Furthermore, (Chakraborty, Hritz, and Dehlinger, 2014) state that distinct sounds should be used when informing players on a matter that requires precision. The guidelines already mentioned from (Garcia and Almeida Neris, 2013; Yuan, Folmer, and Harris, 2011) provide a good background in designing good audio and haptic feedback in an accessible game for the VI. Mainly, the following guidelines will be implemented in *Bongo Beats: Tap With Me* to improve the feedback of the game:

- Make all menus accessible in the game through speech.

- Replace visuals with audio and haptic responses. This is already implemented in *Bongo Beats: Tap With Me*, but it could be improved for the musician role in the game by switching the input mode to use a controller that provides haptic feedback such as vibrating when hitting the notes correctly.

- Create more distinct audio feedback when informing the player of the results of their actions.

Furthermore, in this study, an audio scoreboard (where a button press on the controller would result in the current score of the player being read out loud similar to the scoreboard implemented in (Gluck, Boateng, and Brinkley, 2021)), and improved sound feedback during gameplay will be added to directly address the comments mentioned in (Potamianos, 2022).

In this section so far, areas that can be improved to improve the game experience for VI players have been found, and strategies on how to improve these have been highlighted. This information will be used to improve the experience of the game by improving the feedback for VI players. However, this thesis is a study on collaborative multiplayer games designed for VI children to play with their sighted peers too. Thus, literature on the collaborative play of VI and sighted children will now be discussed, to understand how collaborative games can be improved to increase the engagement of sighted players.

2.3 Collaborative Play

An important aspect of this study is the design of a game that encourages collaboration between the players. This section covers previous studies on collaborative play and games that focus on inclusivity for VI children. (Trust, n.d.) provide some general guidelines on designing inclusive play opportunities for all children. Their guidelines include:

- Creating a rich mix of play opportunities by combining different skill sets. This can be achieved by combining different skills into different roles.

- Engaging the senses to maximize play opportunities. They recommend designing play spaces that also include things to explore through touch, sound, and vibration.

- Several examples of how to make play spaces more accessible, including;

  - thinking about how children can make the whole journey to the play area and back alone, and

- conveying information using different accessible signage such as braille, and using audio.

In a small-scale article on exploring the socio-emotional needs of VI children through research and practice written by (Roe, 2008), the author discusses the importance of social and emotional growth of children during their development and mentions the role of developing friendships during this time. Issues for VI children throughout this period are then discussed, and the article is concluded by providing guidelines on how one can promote social inclusion, interaction, and development of VI children through activities such as games, group discussions, and storytelling. These activities should ensure that VI children:

- "participate actively as a member of a group;"
- "solve problems in a team effort;"
- "understand others' perspectives;"
- "help others (rather than being helped by others);"
- "gain attention, initiate and maintain conversation;"
- "develop strategies to cope with conflict; and"
- "develop self-esteem and confidence."

Bongo Beats: Tap With Me is thus designed to follow these guidelines. The authors also performed a small-scale study interviewing 5 blind children (aged between 10 and 15 years) focusing on their experiences at school, and their friendships. They found that the interests of VI children aligned with the interests of other children. These included computer games and friendships. The children mentioned some difficulties in social interactions, including a widening culture gap due to their inability to play mainstream games with their peers. In the same study, Roe also highlighted the fact that VI children need opportunities to have unsupervised social interaction with their peers. This could come in the form of collaborative play,

(Verver, Vervloed, and Steenbergen, 2020) studied the effect of introducing an augmented toy to increase collaborative play between VI and sighted children. They found that their designed augmented toy initially caused an increase in solitary play and a decrease in cooperative play for VI children in particular, stating that one of the main causes of this was that the VI children spent more time exploring the sounds coming from the toy, and less time interacting with the sighted child they were paired up with. They however state that the VI participants showed similar social interaction behaviors as the sighted children.

(Metatla et al., 2020) also conducted a study on inclusive play. The authors note that it is important to distribute and share roles between players to encourage empathy, widen the children's perspectives, and promote discussion. They also stated that designing accessible games specifically for VI players could lead to them being another source of isolation for VI children as these games are not played by their sighted peers. This is because these games focus on the needs of a VI audience and are thus simplified to a point where they are not found to be engaging to sighted players.

Through this research on collaborative play, it can be concluded that collaborative games should encourage different forms of communication between the players, and ensure that the VI player will have an important role and feel empowered. The game should be designed in a way that children can play it without the need of adult supervision and in a way that promotes conversation between the players. This should help the reader understand better the design choices made when developing *Bongo Beats: Tap With Me* in (Potamianos, 2022), this will also help make better decisions when implementing changes to the game.

(Metatla et al., 2020) state that the pace of games can be a barrier to making them enjoyable for any audience as children with different skills might find games that are too easy to not be engaging. This is another important aspect to consider when designing a collaborative game between VI children and sighted children. This can be achieved in a rhythm game like *Bongo Beats: Tap With Me* by ensuring the difficulty level is always challenging enough for the players playing the current level. Thus, adjusting the game's difficulty in an online fashion to ensure that the players are always engaged is beneficial to designing inclusive games. This will be discussed in the next section.

2.4 Adaptive Difficulty

A big challenge when designing games is maintaining player engagement. The general approach to ensure players are having fun is to provide the right level of challenge (i.e. the difficulty of the game must be in balance with the players' skills) (Lopes and Bidarra, 2011). This can be achieved by adapting the difficulty of the game to the needs of the particular player in real-time during gameplay (also known as dynamic difficulty adjustment (DDA) (Hunicke, 2005)). However, this must be done without disrupting the player experience since the player might feel cheated if they are aware that the game's difficulty is changing based on their skill (Hunicke, 2005).

Various techniques and models exist in literature on how to adapt game difficulty. (Zohaib, 2018) provides a study of the latest techniques proposed at the time of writing (covering the period of 2009-2018). The author categorizes these techniques based on their approach, and a summary can be seen in table 1. Most of these techniques are rather complex and focus on developing systems for games consisting of multiple mechanics.

Approach	**Definition**
Probabilistic methods	A technique that views DDA as a problem of optimization, attempting to maximize player engagement.
Single and Multi-Layered Perceptrons	A neural-network model which maps characteristics of player behavior and emotion, and level design parameters.
Dynamic scripting	An online unsupervised-learning approach aimed at being computationally efficient and utilizes a rule-based approach where every opponent type has a rulebase. These rules are designed manually using domain-specific information.
Hamlet System	Specifically designed for games that contain an inventory management system (such as available ammunition in a shooter game). Adjusts the supply and demand of the inventory resources in game to adjust the difficulty.
Reinforcement Learning	An adaptive game AI aimed at creating varied gameplay experiences for different playing styles using data generated from players already playing the game.
Upper Confidence Bound for Trees and Artificial Neural Networks	Utilizes artificial neural networks from data derived from the upper confidence bound for trees.
Self-organizing System and Artificial Neural Networks	Adjusts the difficulty level by creating a self-organizing system of AI characters.

Table 1: A summary of different approaches for achieving DDA (adapted from Zohaib, 2018)

Studies have also been made on adapting difficulty for serious games. The needs of these types of games require further emphasis on understanding the difficulties and competencies of the particular player, to make smarter decisions on which aspects of the game to adjust (Hendrix et al., 2018). This technique is known as 'Player Modeling' and has been developed as a way to assess the performance of a player.

(Yannakakis et al., 2013) defines player modeling as "the study of computational means for the modeling of the player's cognitive, behavioral, and affective states which are based on data (or theories) derived from the interaction of a human player with a game" and this can be achieved using different approaches. The top-down approach uses emotional models derived from emotion theories in the field of psychology (for example the arousal and valence model (Russell, 1980)) where the emotional state of the player is tracked and a direct link between the emotional state and "fun" is assumed. A different approach to player modeling would be the bottom-up approach which is similar to an unsupervised learning approach in machine learning. Player data is collected and analyzed without assuming anything about the model, player states are then classified from this data, and thus a model is created that relates the player data to one of the defined player states. Hybrid approaches are also possible which would be a mix of both the top-down and bottom-up approaches.

(Hooshyar, Yousefi, and Lim, 2019) provide an overview of the latest innovations at the time of writing in data-driven approaches to player modeling for educational games. The authors found that there were three objectives of player modeling; to understand the behavior of the player, to recognize the goals of the player, and procedurally generate content for the player that will be optimal for the educational game's purpose. They also provided a list of all the machine-learning techniques available in literature that are currently being applied to player modeling. This has been summarized in table 2.

Category	**Number of Papers**	**Algorithms**
Supervised	15	Hidden Markov Models, Markov Logic Networks, Naive Bayes, Dynamic Bayesian network, Deep feed forward neural network, Deep-LSTM network, Support Vector Machine, Decision Tree
Unsupervised	6	K-means, DBSCAN, Genetic Algorithm, Reinforcement Learning

Table 2: A summary of the different techniques being applied in player modeling of educational games.

(Zook and Riedl, 2012) show an example of using data-driven player modeling to dynamically adjust the difficulty of the game by extending two-dimensional matrices representing the player's performance into three-dimensional matrices representing the time of that performance measure using tensor factorization techniques. Using the added temporal data, these tensors are then used to give an insight into how a player will perform given a particular sequence of challenges, and the difficulty of the game is adjusted accordingly. The authors further show that this method achieved state-of-the-art results for adapting difficulty for a skill-based role-playing action game that they developed. (Prendinger, Puntumapon, and Madruga, 2014) further extends the use of player modeling for dynamic difficulty adjustment to the scenario of multiplayer games where multiple players of different skill levels are playing together. The authors solve the problem where adjusting the difficulty of one player might affect the other players in a sub-optimal way by using the approach of distributed constraint optimization.

The models described in literature on implementing player modeling are very specific to the particular game being made. (Hendrix et al., 2018) thus provides a six-step general framework for implementing adaptive difficulty into a game. The authors demonstrated how this can be achieved both for games that have already been designed, and for designing a new game from scratch, thus these steps will be followed when implementing adaptive difficulty for *Bongo Beats: Tap With Me*. The steps are listed below:

1. Identify variables that are good indicators of the player's performance in the game. One should also define how these variables indicate performance (for example, if you are counting the number of notes hit in succession in a rhythm game, then this indicates that the player is performing well).

2. Identify the variables that will adjust the difficulty of the game.

3. Locate these variables in the implementation (if an implementation already exists).

4. If the game features multiple mechanics, identify these and list which variables correspond to each mechanic.

5. Decide how the variables will be used to adjust the difficulty.

6. Decide on the best values to use as the starting variables.

Thus a method has been identified for implementing adaptive difficulty in *Bongo Beats: Tap With Me*. In the next section, literature on difficulty in rhythm games will be discussed to identify which variables should be tweaked when adjusting the difficulty of the game.

2.4.1 Difficulty in Rhythm Games

Not much research has been conducted on developing adaptive difficulty for rhythm games in particular. In a fascinating study, (Gold and A. Olivier, 2010) proposed using machine translation to convert tracks to have easier difficulty levels after creating the hardest difficulty level. This was achieved by using data from the levels in *Guitar Hero I* and *II* to train the sequence of notes as a language, where each difficulty level was treated as the same "text" but in a different language. Thus, the developer will only have to create the hardest difficulty mode of the level, and the model would "translate" that to generate all the other difficulty modes. This technique avoids the need for the designers to create different difficulty levels and preserves the feel of the game.

In their study of procedural content generation for rhythm games (Liang, W. Li, and Ikeda, 2019) defined a list of variables that would affect the difficulty of a level. These are:

1. The complexity of the actions and combinations of actions to perform. This can include how many notes there are available to hit, how many notes can be used in combination that should be hit, and the timing of the notes with respect to the rhythm of the music (for example including notes on the off-beat).

2. The density of actions (i.e. the number of actions per second)

3. The speed of the action marker (i.e. the reaction time of the player)

4. The strictness of judging misses (i.e. how much delay time is allowed)

5. The number of misses permitted

In the case of *Bongo Beats: Tap With Me*, the density of actions, and complexity of action combinations would be the ideal variables to adjust. Altering the speed of the action marker or the strictness of judging misses on the fly without warning the player will lead to confusion and frustration from the players. Thus, through the research presented in this section, a technique for dynamic difficulty adjustment has been found and the parameters to control the levels of difficulty have been identified. In the next section, the research questions will be discussed, and the methodology and experiments will follow in section 4.

3 Research Questions

In order to continue on and improve the research done in (Potamianos, 2022) the main set of research questions aims to re-run the experiments done in (Potamianos, 2022) with VI students studying at Bartiméus School, and provide a deeper understanding of how VI players perform, collaborate with other players, and experience an inclusive collaborative multiplayer game designed with accessibility features for the VI. Mainly:

RQ1.1 How did VI children perform in the inclusive multiplayer rhythm game *Bongo Beats: Tap With Me*?

RQ1.2 How did VI children collaborate in the inclusive multiplayer rhythm game *Bongo Beats: Tap With Me*?

RQ1.3 How did VI children experience the inclusive multiplayer rhythm game *Bongo Beats: Tap With Me*?

The next set of research questions relates directly to the changes implemented in the game. Two different pilot studies were performed to evaluate each change and ensure that they have impacted the gameplay experience in the desired way. For the first research question of the pilot study, the quality of the generated levels should be indistinguishable from that of the handcrafted level. If the generated notes feel more random and do not feel like they make sense in the context of the song then the immersion of the player will be broken. Thus the first research question is:

RQ2: Is there a perceived difference in immersion between students playing the handcrafted levels (from the previous study) and students playing the generated levels (from the algorithm used for this study)?

H0: There was no significant difference in immersion scores between students who played the handcrafted levels and students who played the generated levels.

For the second research question of the pilot study, the experiences of the participants playing the musician role using the new improved feedback were collected through a questionnaire in order to understand how well they understood what was happening in the game and how their actions affected the game world. In particular:

RQ3: How did the players playing the version of *Bongo Beats: Tap With Me* with improved feedback rate their understanding of the level?

The final research question of the pilot study covers the implementation of adaptive difficulty in the game. These research questions are inspired by the experiment setup in (Hendrix et al., 2018), mainly:

RQ4.1: Is there a difference in the performance of the players when playing the adaptive difficulty version of *Bongo Beats: Tap With Me* when compared to playing the original version with 1 difficulty setting?

RQ4.2: Which version of *Bongo Beats: Tap With Me* did players find more fun and engaging to play; the adaptive difficulty version, or the original version?

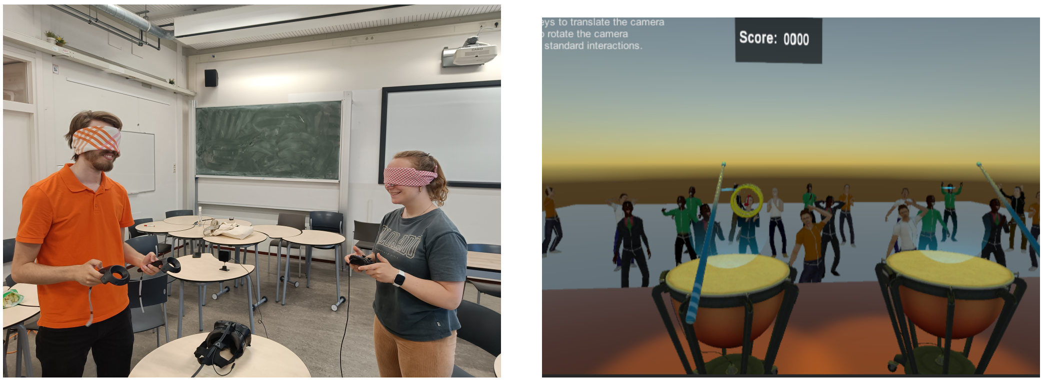
In the next section, the methodology and experiment setup will be discussed in detail, focusing on how the proposed changes were implemented as well as how the experiments were carried out, and how data was collected.

4 Methodology

4.1 The Game

Bongo Beats: Tap with Me is a rhythm game developed with inclusive features for the VI. The game is a collaborative multiplayer game, as two players work together to hit as many notes as possible. One player takes on the role of the conductor who receives feedback from the game informing them which bongo needs to be hit and when. The conductor must then relay this information to the musician, who will hit the bongos at the correct time. Thus, the game is also a serious game, as it has the potential to teach social and rhythmic skills (Potamianos, 2022).

The game is inclusive as it can be played in audio-haptic mode, audio-visual mode, or audio-visual-haptic mode. In the haptic modes, a "stimuli replacement" technique is used to replace visual feedback with haptic feedback. This is inspired by the work done in (Yuan and Folmer, 2008). In order to enable the haptic mode, the player must hold an HTC Vive controller in each hand. The controllers will vibrate before a bongo needs to be hit to convey this information to the player. The sighted players also have the option to wear the HTC Vive headset and stand on stage in front of the bongos to see the notes approaching, doing so will enable the visual mode. Both holding the controllers and wearing the headset will thus enable both visual and haptic modes, this is specifically designed for VI players who are not blind but still need extra haptic feedback for assistance. Figure 1a shows two players playing the game in audio-haptic mode, while figure 1b shows the visual mode of the game.



(a) Two participants playing the game in audio-haptics mode. The player on the left is the conductor, and the player on the right is the musician.

(b) A screenshot taken in-engine showing the visuals of the game, This is what the player would see in VR if they use the visual mode.

Figure 1: Showcasing the different ways to play *Bongo Beats: Tap With Me*

4.2 Developing the Game

Before running the experiments, the new features must first be developed. A month of development time was dedicated to this. The original source code for the game that was developed in (Potamianos, 2022) can be found online[10]. The initial idea was to use this project as a basis to add new features to it, however, it was decided to rewrite the code for the game from scratch for a number of reasons, mainly:

1. The new features added to the game required more processing overhead and the optimization of the code of the original game was rather poor.
2. Certain game events would need to be updated and changed as the game is run (for example when the difficulty of the game changes on the fly when running the game using DDA), and the original code was not designed for this purpose.
3. All the hard-coded events in the original game depended on the game time and not the audio track time, which could cause latency on lower-end computers.

The final code is available online[11]. The remainder of this subsection will cover the design and implementation of each of the new features added to the game.

[10]https://github.com/evangelospot/BongoBeats-Tap-with-me
[11]https://github.com/mfer0010/Bongo-Beats-Tap-With-Me-V2-Public-

4.2.1 Developing an algorithm to generate a level given a particular song

The first step in the process to generate a level given a particular song was to choose the song. All the songs were chosen from royalty-free sources to ensure that the game can be shared easily. After choosing a song, and converting it to *.wav* format, the beat instances for that song were generated. This was done using a method based on the findings in section 2.1.1. In particular, the *BeatNet* algorithm was used to generate the beat times for each song and saved in a *.csv* file. The quality of the output was then checked using Sonic Visualizer to ensure that all timestamps generated are on the beat and in time with the song. If needed, the hyperparameters of the beat tracking algorithm were altered to get the best result for that particular song. The following hyperparameters of the *BeatNet* algorithm could be adjusted:

1. *Model*: Since no new model was trained, BeatNet comes with three different pre-trained models that can be utilized. This was the main hyperparameter that was changed for each song.

2. *Mode*: This parameter selects whether the algorithm should work in an online fashion (i.e. only using information from the current time stamp and previous time stamps) or offline fashion (i.e. the whole audio file will be considered when extracting the beats). For the songs used in this study, the results were always more accurate using the *offline* mode.

3. *Inference Model*: This parameter selects which model will be used to determine the beat activations; the DBN model (a pre-trained model used by the authors of (Heydari, Cwitkowitz, and Duan, 2021)) which was found to be more accurate yet more time-consuming, or the Particle Filtering Model (uniquely developed in (Heydari, Cwitkowitz, and Duan, 2021)) which sacrifices some accuracy for faster processing. Note that the DBN model cannot be used in online processing.

Figure 2 shows this whole process using a flowchart. This process was repeated for all 14 songs that were added to the game. The files containing the beat instances of each song were then used together with the audio file of the song as inputs to the algorithm to generate the notes for the game.

Figure 3 details the algorithm for generating the notes for a level, given a song, its associated beat instances, and by defining values to the following parameters:

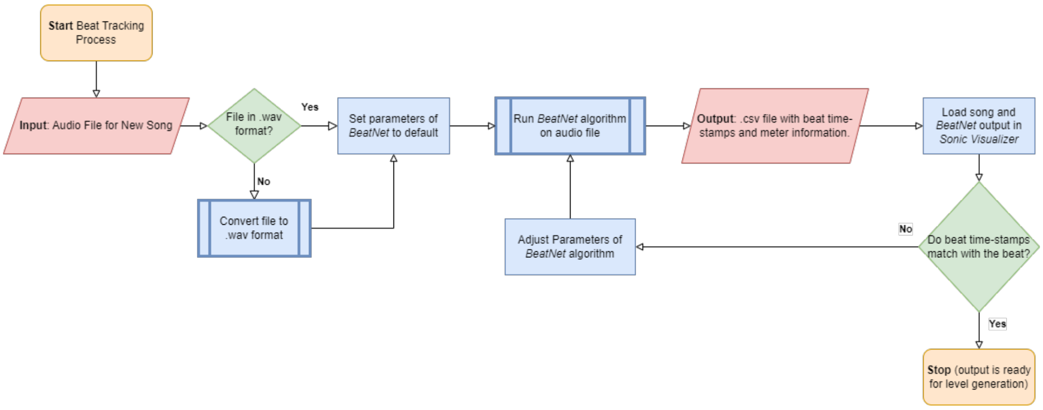


Figure 2: A flowchart showing the process taken to generate the beat instances for a given song audio file.

- *songStartBuffer*: This parameter is used to control when the earliest possible note can be generated, all beat instances before *songStartBuffer* will be ignored. This was set to 5 seconds for the final version of the game.

- *minTimeBetweenNotes*: When considering a beat, if the last note generated is closer than *minTimeBetweenNotes*, then the beat is ignored. This is done to ensure that the generated notes are not too close together, thus making the game unplayable or too difficult. This is one of the variables identified as a candidate to vary depending on the difficulty level (see subsection 4.2.3)

- *RMSThreshold*: An algorithm to calculate the Root Mean Square (RMS) Energy of the audio file was written. This is calculated in order to ensure that no note is generated when the song is quiet (usually towards the end of the song, or in pauses in the song). Thus the *RMSThreshold* parameter is used to ensure that at the time of generating a note, the RMS Energy of the song is greater than this threshold. This value varied depending on the song, but a generally good starting point was 0.05.

- *ProbNoteGenerated*: For each beat that was identified as a possible candidate to generate a note, *ProbNoteGenerated* is used to determine if a note should be generated or not. The probability value varies as the song progresses, where the probability is lower in the first 30 seconds of the song and keeps increasing to be at its highest after 90 seconds. This was done to ease the player into the level by ensuring

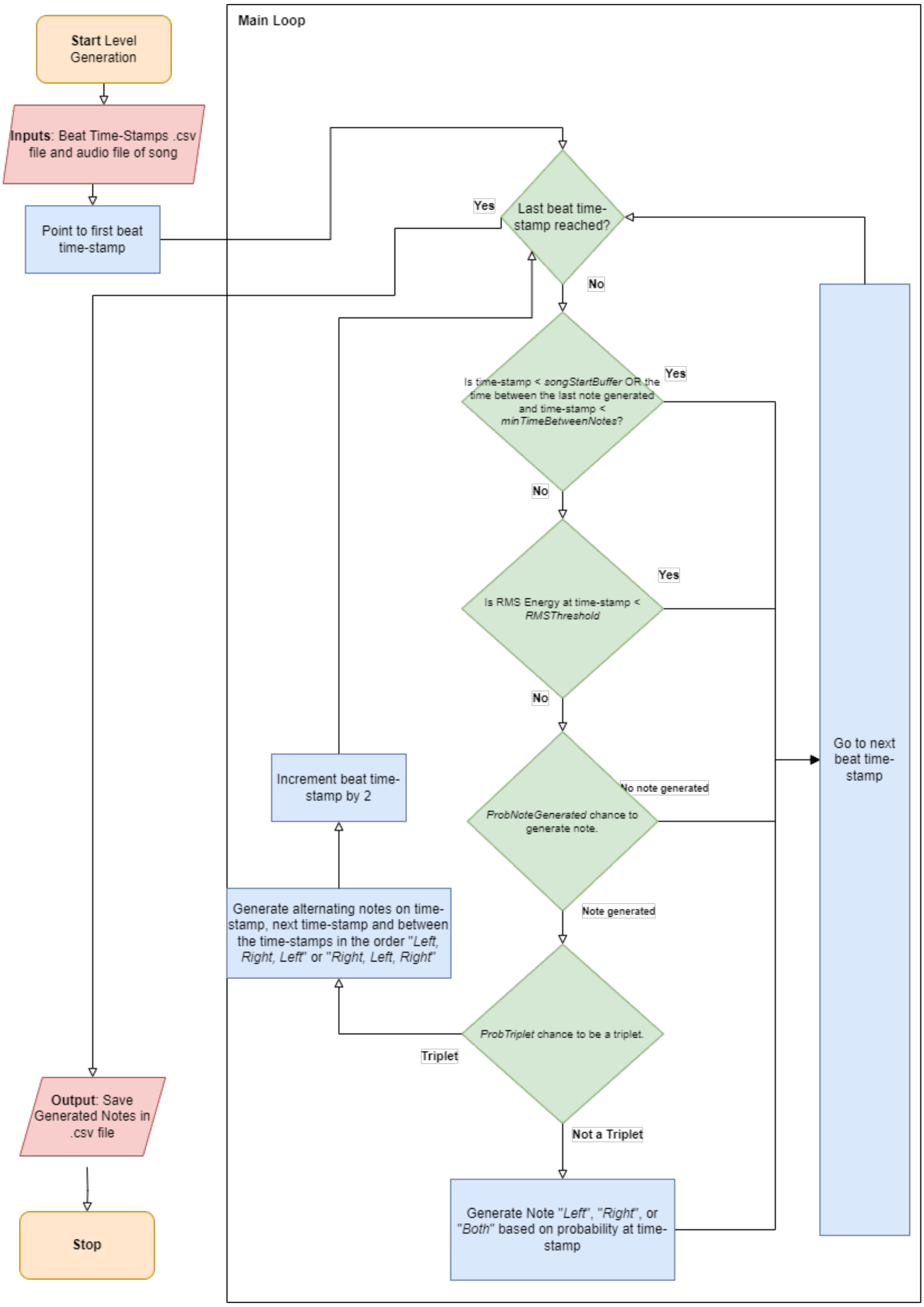


Figure 3: A flowchart showing the process taken to generate the notes of the level given a song audio file and its associated beat instances *.csv* file.

that there are fewer notes at the start of the song, and more notes as the song progresses. Furthermore, these probability values also change depending on the difficulty level, and more information on the exact values used can be found in subsection 4.2.3.

- *ProbDoubleNote*: If a note is generated, the note can be either a "left bongo" note, a "right bongo" note, or "both bongos". *ProbDoubleNote* decides the probability that a "both bongos" note is generated, thus the probability of a "left bongo" or "right bongo" note is $(1 - \text{ProbDoubleNote})/2$. This was implemented as "both bongos" was identified as being more challenging than playing a single bongo. The exact value of this parameter also varies depending on the difficulty level, more information can be found in subsection 4.2.3.

- *ProbTriplet*: If a note is generated, there is also a slight chance that it is a triplet note. A triplet note means that the player must play the bongos in the order "Left, Right, Left" or "Right, Left, Right" where the middle note lies between two beats (i.e. on the offbeat). This is implemented to add more challenge in the higher difficulty levels and its exact value can be found in subsection 4.2.3.

The algorithm works by looping over all the beat instances, and when finding a beat instance that is a suitable candidate for a note to be generated, the probability parameters are used to decide if a note is generated, and what type of note to is generated. This resulted in generated notes that followed a similar pattern to the handcrafted level developed by (Potamianos, 2022).

4.2.2 Improving the quality of feedback given to the player during gameplay

The design and implementation of the improved feedback features were based on all the findings discussed in section 2.2. Mainly, the guidelines proposed in (Chakraborty, Hritz, and Dehlinger, 2014; Garcia and Almeida Neris, 2013; Yuan, Folmer, and Harris, 2011) and advice given in (Potamianos, 2022) were followed to make an informed decision on how to improve the quality of feedback for the game *Bongo Beats: Tap With Me*. In particular, the following features were implemented:

1. All the sound effects used in the previous version of the game were changed to use better-quality sounds. These sounds were taken from a royalty-free website[12], or a sound pack[13].

[12] freesound.org

[13] purchased from gamedevmarket.net

2. Modulation was also added to all the feedback sounds that are played, where the volume and the pitch of the sound are randomly altered by a small amount to increase immersion and dynamism of the sounds. This makes the sounds sound like they're coming from more natural sources, and thus feel less repetitive.

3. The visual mode was not working as intended in (Potamianos, 2022), this was thus implemented to work properly to allow users to use this mode.

4. Visual feedback was also added to the game which correlates with all the audio feedback given in order to be consistent in audio-visual mode. This included adding a stage and a crowd to the game and updating the notes to indicate if they have been hit or missed.

5. In the previous version of the game developed by (Potamianos, 2022), the only feedback sounds were a drum sound to inform the player that they are hitting the bongos, and a very similar drum sound informing the player that the note was hit correctly. As suggested in (Friberg and Gärdenfors, 2004), the feedback sound informing the player that a note was hit correctly was *separated* from the game world sound by changing it to the sound of a 'ride cymbal' as this is a more distinct sound. Furthermore, feedback sounds were also added to indicate the following situations:

   - when a note is missed completely,
   - when the players hit a note too early,
   - when the players hit a note too late,
   - a background 'crowd' noise was added to the level to increase immersion,
   - a 'crowd cheering' sound is played when the level is complete.

6. Spatial audio was also added to the game, where the feedback sounds related to the left drum are panned to the left, and the feedback sounds related to the right drum are panned to the right.

7. An audio mixer was added to the game which dynamically ensures that the feedback sounds are always heard and separated from the other sounds in the game (such as the song and the crowd) by subtly lowering the volume of these sounds when a feedback sound is played. This was achieved using the 'Duck Volume' effect in Unity's audio mixer.

8. Controller support was added to the musician role as suggested in (Garcia and Almeida Neris, 2013), as when using a controller, it is easier for a VI player to be aware of what buttons they are pressing. Haptic feedback was also added to the controller, which would vibrate when a note is hit correctly.

9. A menu was created which allows the musician to select the song they want to play by moving the left joystick on the controller up or down and listening to a snippet of the song, before selecting one they enjoy.

10. An audio scoreboard was also implemented, where at any point in the game, the musician can press any of the face buttons on the controller and the score would be read out (in Dutch) to the players.

11. All interface objects were made accessible through speech. In particular, at the end of each level, the end-of-level screen is read out to the players, congratulating them on completing the level and informing them of their final score, how many notes they missed, and how many notes they hit.

12. A tutorial scene was created where all the sounds of the game are introduced to the player. Unfortunately, due to time constraints, it was not possible to implement this feature as intended, however before each experiment, the sounds were still introduced to the participants in order to replicate this.

4.2.3 Implementing realtime DDA

The findings in section 2.4 were used as a basis for implementing adaptive difficulty for *Bongo Beats: Tap With Me.* Particularly, the six-step framework provided in (Hendrix et al., 2018) was followed and the variables chosen to be adapted to affect the difficulty were inspired by those mentioned in (Liang, W. Li, and Ikeda, 2019).

For the algorithm to work in real-time while the game is being played, a window of notes is taken where the performance of the players is assessed and the decision on how to adapt the difficulty is taken at the end of each window. That is, if during the window the players are performing well, then the difficulty level will increase, if the players are performing badly then the difficulty level will decrease, and if the players' performance is average then the difficulty will remain the same. After testing the game with different window lengths, it was decided to use a window size of 8 notes as this would give plenty of opportunities when playing the song during the experiments to assess the difficulty and adjust it accordingly. Any value lower than this would not give a good overview of the performance of the players, and any value higher than this would not give enough opportunities in the limited time available for each experiment to adjust the difficulty.

The details of the six-step plan as defined in (Hendrix et al., 2018) that was used in the development of *Bongo Beats: Tap With Me* are as follows:

1. The first step is to identify the variables that will be a good indicator of the performance of the players. The mistakes ratio calculated for the current window was used and this variable is defined in equation 1. This will give a mistakes ratio $r \in [0, 1]$, where a value of 0 implies a perfect performance (i.e. all notes were hit), and a value of 1 implies the worst possible performance (i.e. none of the notes were hit).

$$\text{Mistakes Ratio (Window) } r_{\text{w}} = \frac{\text{\# notes missed}}{\text{window size}} \tag{1}$$

2. The second step was to identify what would influence the difficulty of the game. According to (Liang, W. Li, and Ikeda, 2019), the density of actions, and the complexity of actions should be adjusted to alter the difficulty in rhythm games.

3. Following this, the variables that adjust the difficulty were located in the code. In order to adjust the density of the actions, the variables *minTimeBetweenNotes*, and *ProbNoteGenerated* (as already discussed in section 4.2.1) should be altered. In order to adjust the complexity of actions, the variables *ProbDoubleNote* and *ProbTriplet* (as already discussed in section 4.2.1) should be adjusted accordingly.

4. Step four is only relevant if the game features multiple game mechanics, and thus was skipped for *Bongo Beats: Tap With Me*.

5. The next step is to define how the performance variables will be used to calculate the difficulty. In the case of *Bongo Beats: Tap With Me*, after some balancing and an initial pilot test, it was decided to use the values defined in equation 2 to determine the difficulty level for the next window where r_w is the mistakes ratio as defined in equation 1.

$$r_w \begin{cases} < 0.25 \implies \text{Raise Difficulty by 1 Level} \\ > 0.55 \implies \text{Lower Difficulty by 1 Level} \\ \in [0.25, 0.55] \implies \text{Keep the Same Difficulty Level} \end{cases} \tag{2}$$

   The game features 5 total difficulty levels, these 5 versions of the game are all generated when the levels are generated (during the process outlined in section 4.2.1), they are loaded into the scene at runtime, and a pointer is kept using the song time to dynamically swap between the different difficulty levels at the end of each window. The values used for each variable for each difficulty level can be seen in table 3.

6. The final step is to decide on the starting value when loading the game. It was decided to start the game on the middle difficulty (i.e. level 3), with the idea that the game would automatically adapt to the correct level for the players after a few windows.

Difficulty Level	**1**	**2**	**3**	**4**	**5**
minTimeBetweenNotes	2s	1.5s	1.25s	1s	0.8s
ProbNoteGenerated <30s	0.05	0.05	0.05	0.1	0.15
ProbNoteGenerated 00:30-01:00	0.08	0.08	0.1	0.15	0.2
ProbNoteGenerated 01:00-01:30	0.1	0.125	0.15	0.2	0.3
ProbNoteGenerated >01:30	0.15	0.18	0.2	0.3	0.35
ProbDoubleNote	0	0.05	0.15	0.33	0.33
ProbTriplet	0	0	0	0.01	0.05

Table 3: A table showing how the different level generation variables change depending on the difficulty level. Note that the *ProbNoteGenerated* variable has 4 different values depending on the song time.

Following the implementation of the new features in the game, pilot experiments were conducted to ensure that the new features were implemented correctly and have a positive impact on the game experience. The final experiments were then conducted at Bartiméus school with VI children. The remainder of this section details the design and processes taken for each experiment.

4.3 Experiments

4.3.1 Pilot Studies

Initially, pilot studies were conducted where students were gathered and paired into dyads. The participants were split into 2 groups, where each group was part of a different experiment. Group 1 participated in an experiment aimed at answering RQ2, and Group 2 participated in an experiment aimed at answering RQ4.1 and RQ4.2. RQ3 was answered through a survey that was handed out to all the participants who played the musician role in the game.

To answer RQ2, Group 1 was once again split randomly into a control group and an experiment group. Each dyad played the game twice, where the participants swapped their roles in the game on their second playthrough (i.e. the player that played the coordinator role in the first game played the musician role in the second game and vice-versa). The control group played the handcrafted level that has been implemented in (Potamianos, 2022) (however with the improved feedback features developed in this study), and the experiment group played the level with the same song as the control group, however, the notes had been generated using the algorithm as explained in section 4.2.1. The participants were not informed whether they are part of the control or experiment group. The game experience questionnaire (GEQ) was handed out to the participants immediately after playing the game.

The game experience questionnaire (GEQ) (IJsselsteijn, De Kort, and Poels, 2013) is commonly used in literature to evaluate player experience (Johnson, Gardner, and Perry, 2018). It has been used to assess serious games (De Lima, Lima Salgado, and Freire, 2015), evaluate sound effects and music in games (Nacke, Grimshaw, and Lindley, 2010) and evaluate the degree of challenge in games (Nacke, Stellmach, and Lindley, 2011) amongst other studies. For the experiments of Group 1, the Core Module and the Post-Game Module of the GEQ were handed out to the participants. The core part of the GEQ scores the game experience in 7 categories, where $3 - 6$ questions are presented for each category and the average value of all questions is considered during the analysis. The categories are:

1. Competence (5 questions),
2. Sensory and Imaginative Immersion (6 questions),
3. Flow (5 questions),
4. Tension/ Annoyance (3 questions),
5. Challenge (5 questions),
6. Negative Affect (4 questions),
7. Positive Affect (5 questions).

The post-game module consists of 4 categories, where each category consists of $2-6$ questions. The categories are:

1. Positive Experience (6 questions),
2. Negative Experience (6 questions),
3. Tiredness (2 questions),
4. Returning to Reality (3 questions).

The experiment setup for Group 2 (to answer RQ4.1 and RQ4.2) was inspired directly from (Hendrix et al., 2018) where the authors tested the implementation of adaptive difficulty for two games they developed. In particular, participants from Group 2 were also split into two groups where the first group played a fixed-difficulty version of the game, then proceed to play the version of the game with the implemented adaptive difficulty, and the second group started with the adapted difficulty version and play the fixed-difficulty version after. This was done with the aim to reduce learning effects. The players this time kept the same roles in the game when playing the different versions, and were also not aware of the differences in the versions of the game they played. Note that the same song was used for all experiments of this group in order to reduce the effect of any confounding variables. Their performance was tracked using equation 3, where a lower value of r_s indicates a better performance than a higher value.

$$\text{Mistake Ratio (Entire Song) } r_s = \frac{\text{\# notes missed}}{\text{Total \# notes}} \tag{3}$$

This was used to answer RQ4.1. A questionnaire was handed out to the participants after they have played both versions of the game in order to answer RQ4.2. The questionnaire[14] has questions about the background of the player and questions related to their enjoyment of each version of the game.

For all the participants who participated in the pilot studies, if they had the opportunity to play the game in the musician role, then they were also handed a separate questionnaire[15] asking them about the feedback they received when playing the game. This was used to answer RQ3.

[14]Please refer to the appendix for a copy of this questionnaire.
[15]This can also be found in the appendix.

4.3.2 Final Experiments

Following the pilot studies, experiments were carried out with VI students from Bartiméus school in order to answer RQ1.1, RQ1.2, and RQ1.3. A consent form along with an information sheet[16] was sent to the parents of the participating students prior to the experiments.

The experiment setup was similar to the final experiment done in (Potamianos, 2022). Participants were split into dyads and played the final version of the game. Each participant played the game twice, one time as the musician, and one time as the conductor. The musician chose the song. For the conductor role, the participants were also given the option to play in haptic mode, visual mode, or haptic and visual mode.

Immediately after playing the game, the participants answered a questionnaire about their experience with the help of an adult. The questionnaire contained the In-Game Module of the GEQ, as well as the Social Presence Module. It was decided to use the shorter In-Game Module rather than the full version as we wanted the questionnaire to be as short as possible in order to put less stress on the participants and retain their focus. This module was used to answer RQ1.3. The Social Presence Module investigates the behavioural and psychological involvement of each player in relation to the other player and was used to answer RQ1.2. The participants also had the opportunity to mention any other comments or feedback about their experience. As is defined in (IJsselsteijn, De Kort, and Poels, 2013), the In-Game module of the GEQ scores the game experience in the same 7 categories as the core module, however only 2 questions are asked for each category. The Social Presence Module scores the social aspect of the game experience into 3 categories:

1. Psychological Involvement – Empathy (6 questions)
2. Psychological Involvement – Negative Feelings (5 questions)
3. Behavioural Involvement (6 questions)

4.3.3 Expert Interviews

Expert interviews were also conducted with a VI international student studying at Utrecht University, a member of the Dutch Visio organisation[17], and 3 PhD Students working in the field of assistive technologies for the VI. For these sessions, the participants played the game as many times as they

[16] Refer to the Appendix
[17] www.visio.org/en-gb/home

wished, ensuring that they played both the musician and the coordinator roles and this was followed by a semi-structured interview where the researcher asked questions about their experience of the game, what they liked, and what they feel should be improved.

In the next section, the results from the pilot and final experiments will be presented, as well as the results from the expert interviews.

5 Results

5.1 Pilot Studies

In total, 25 participants were recruited to participate in the pilot studies using opportunity and snowball sampling. All of the participants were students studying at Utrecht University or Hogeschool Utrecht. These participants were split into two groups as explained in section 4.3.1, where 14 participants were in Group 1 (to answer RQ2), and 11 participants were in Group 2 (to answer RQ4.1 and RQ4.2). A total of 15 participants experienced the game as a musician, and thus they also filled in the extra questionnaire aimed at answering RQ3.

5.1.1 Research Question 2

The 14 participants that were placed in Group 1 participated in the pilot study to answer RQ2 (*Is there a perceived difference in immersion between students playing the handcrafted levels (from the previous study) and students playing the generated levels (from the algorithm used for this study)?*). These participants were split into the control and experiment groups. In total, 6 participants were in the control group, and 8 participants were in the experiment group. As mentioned in section 4.3.1, the core and post-game modules from the GEQ were handed out to the participants after they played the game for 2 rounds. The gathered data was then analysed, and the results are discussed in this section.

The scores for each category of the core and post-game modules of the GEQ were calculated as explained in (IJsselsteijn, De Kort, and Poels, 2013). Table 4 shows the descriptive statistics for each category in the core module of the GEQ and these are visualised in the box plot shown in figure 4. Table 5 shows the descriptive statistics for each category in the post-game module of the GEQ and these are visualised in the box plot shown in figure 5. Despite the small sample size, this data shows that the results for each category of the GEQ do not vary significantly. In order to test this claim, the Mann-Whitney

	Group	min	max	range	median	mean	var	std. dev
Competence	*Ctrl.*	2.2	3	0.8	2.7	2.6	0.112	0.335
	Exp.	0.2	3.6	3.4	2.5	2.3	1.577	1.256
Sensory and Imaginative Immersion	*Ctrl.*	1	3.167	2.167	2.167	2.083	0.575	0.758
	Exp.	1.667	3.333	1.667	2.917	2.708	0.482	0.694
Flow	*Ctrl.*	1.2	3.8	2.6	2.7	2.567	0.727	0.852
	Exp.	1.2	3.2	2	2.5	2.45	0.626	0.791
Tension/ Annoyance	*Ctrl.*	0	1	1	0.333	0.444	0.207	0.455
	Exp.	0	2	2	0.5	0.625	0.490	0.700
Challenge	*Ctrl.*	0.2	2.4	2.2	1.8	1.6	0.544	0.738
	Exp.	1.6	3	1.4	2	2.15	0.167	0.411
Negative Affect	*Ctrl.*	0.25	1	0.75	0.5	0.583	0.117	0.342
	Exp.	0	2	2	0.625	0.844	0.499	0.706
Positive Affect	*Ctrl.*	3.25	4.75	1.5	4.25	4.083	0.492	0.701
	Exp.	3	4.5	1.5	3.875	3.844	0.249	0.499

Table 4: Descriptive statistics for the core categories of the GEQ from the pilot study to answer RQ2 (rounded to 3 decimal places)

	Group	min	max	range	median	mean	var	std. dev
Positive Experience	*Ctrl.*	1.667	2.667	1.5	2.083	2.056	0.285	0.534
	Exp.	0.833	3	2.167	2.25	1.958	0.856	0.925
Negative Experience	*Ctrl.*	0	0.833	0.833	0.083	0.194	0.195	0.323
	Exp.	0	1	1	0.25	0.563	0.603	0.776
Tiredness	*Ctrl.*	0	1.5	1.5	0	0.333	0.367	0.606
	Exp.	0	2	2	0.25	0.563	0.603	0.776
Returning to Reality	*Ctrl.*	0	1.667	1.667	0.5	0.556	0.385	0.621
	Exp.	0	1.333	1333	0.333	0.542	0.379	0.616

Table 5: Descriptive statistics for the post-game categories of the GEQ from the pilot study to answer RQ2 (rounded to 3 decimal places)

U test (Mann and Whitney, 1947) was used for pairwise comparisons between the control and experiment groups for each category. It was decided to use this non-parametric test due to the very small sample size. This test assumes that all samples from both groups are independent of each other and that the data can be ordered (i.e. a score of 3.2 $>$ a score of 2.9). For each pairwise comparison done, the hypotheses are as follows:

H_0: The distributions of both populations are identical.

H_1: The distributions are not identical.

The results of the Mann-Whitney U test for all the categories of the GEQ can be seen in table 6. We do not have enough evidence to reject the null hypothesis that the distributions are identical (with confidence of $p = 0.05$) in all the cases. Thus, from the evidence presented, the distributions of each control and experiment group for every category are similar.

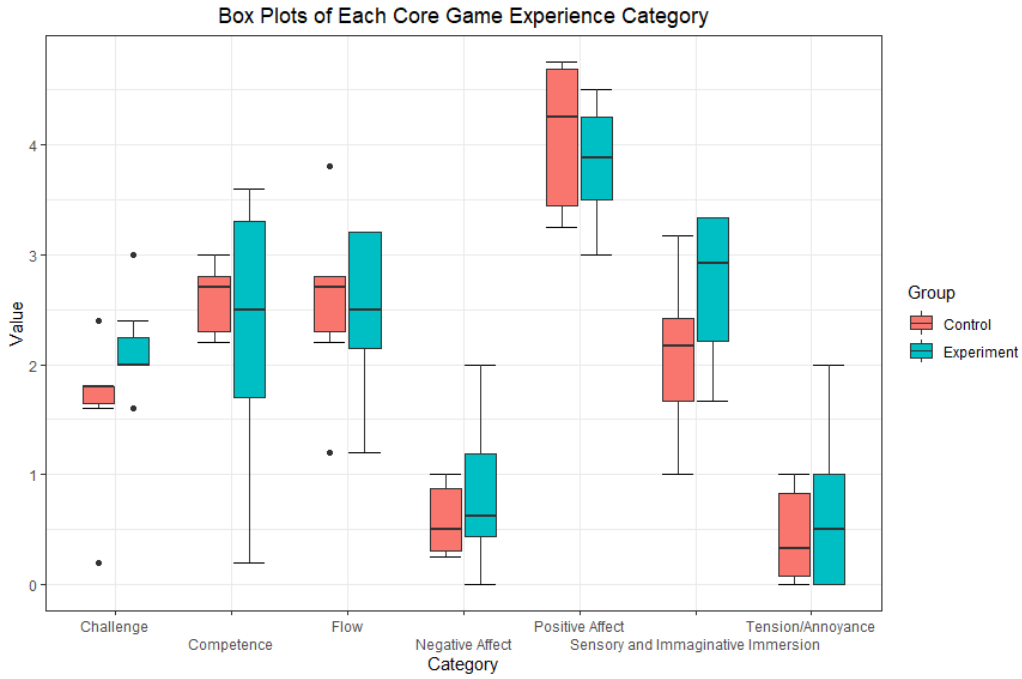


Figure 4: Box Plots showcasing the difference in experience between the control and experiment groups for the experience categories in the core module of the GEQ.

The hypothesis for this pilot experiment is that there would be no significant difference between the immersion scores of players who played the handcrafted level when compared to players who played the generated levels. This is shown particularly in the results for the 'Sensory and Imaginative Immersion' and the 'Flow' categories. Thus, we have evidence to suggest that the players do not feel a large difference when playing the handcrafted or generated levels.

5.1.2 Research Question 3

The 15 participants that played the game in the musician role also answered a questionnaire asking them about their understanding of the game and their opinions of the feedback given to them by the game while they play. This was done in order to answer RQ3 (*How did the players playing the version of Bongo Beats: Tap With Me with improved feedback rate their understanding of the level?*). The gathered data is presented in figure 6.

The results in figure 6 show a positive trend for the new feedback features implemented. Mainly, it was clear to the participants what effect their

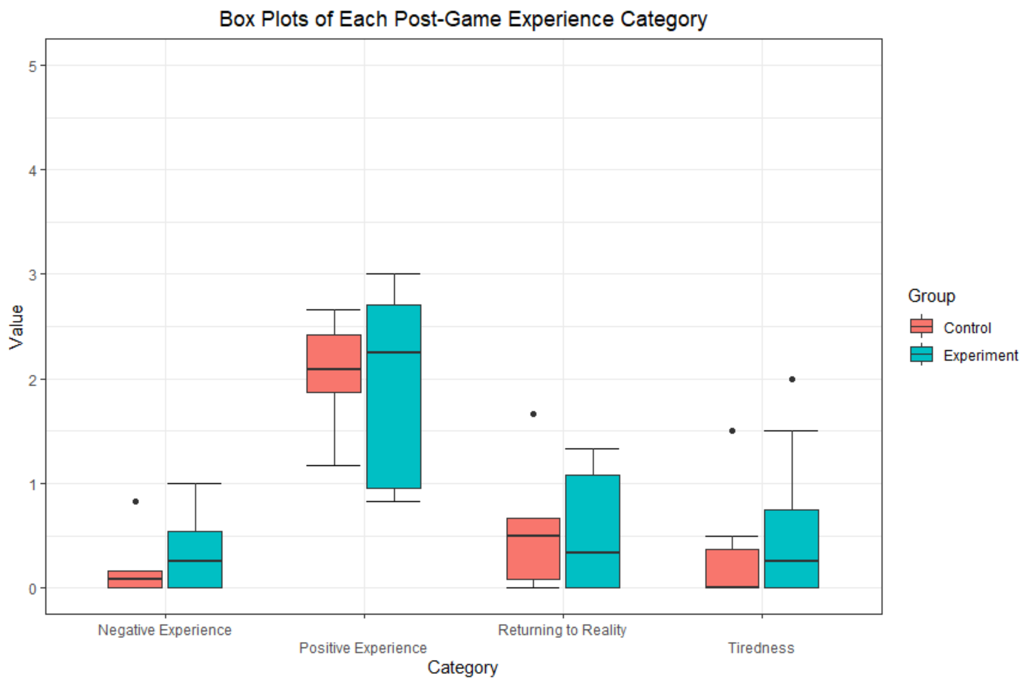


Figure 5: Box Plots showcasing the difference in experience between the control and experiment groups for the experience categories in the post-game module of the GEQ.

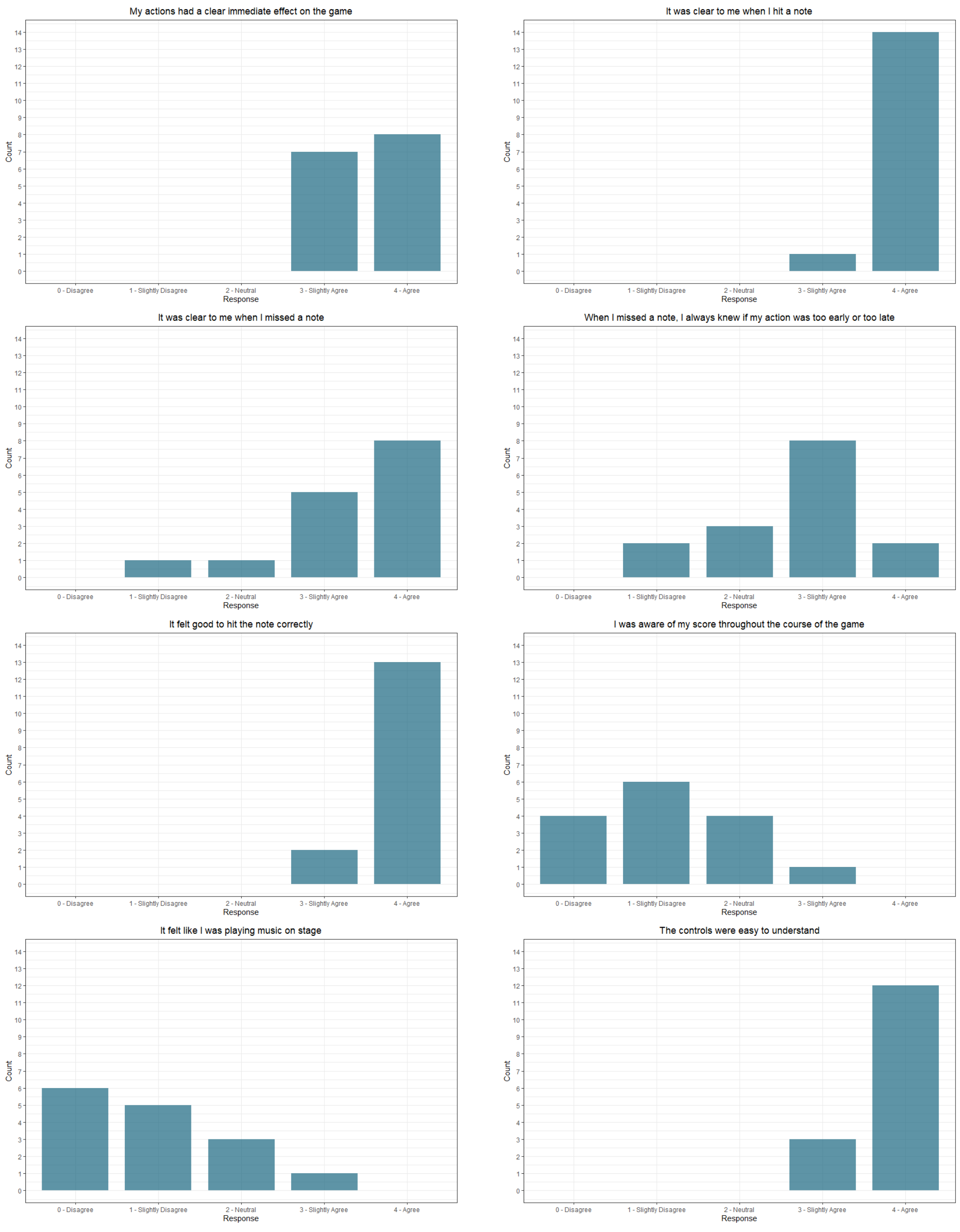


Figure 6: Bar-plots showing the results of the survey handed out to all participants who played in the musician role

	Median (control)	Median (experiment)	Mann-Whitney U	p-value
Competence	2.7	2.5	25.5	0.897
Sensory and Imaginative Immersion	2.167	2.917	115	0.119
Flow	2.7	2.5	25	0.9481
Tension/ Annoyance	0.333	0.5	22	0.841
Challenge	1.8	2	10	0.076
Negative Affect	0.5	0.625	20	0.645
Positive Affect	4.25	3.875	30	0.475
Positive Experience	2.083	2.25	23	0.949
Negative Experience	0.083	0.25	18.5	0.500
Tiredness	0	0.25	19.5	0.565
Returning to Reality	0.5	0.333	25	0.946

Table 6: The results of the Mann-Whitney U test for all categories of the GEQ, showcasing that we do not have enough evidence to reject the null hypothesis ($p < 0.05$, two-tailed)

actions had on the game, the new controls were easy to understand, and the users reported that it was satisfying to hit a note correctly. This shows clear improvements to the previous version of the game implemented in (Potamianos, 2022), where in that study, users reported that "it was unclear if the performer was early or late", "participants, in general, had a very low correctness rate on hitting the notes", the controls were "difficult at first", and "several participants had difficulty with finding the left/right arrow key on the keyboard".

The questions "I was aware of my score throughout the course of the game", and "It felt like I was playing music on stage" were the lowest-scoring questions. Regarding the users being aware of the score, none of the participants made use of the audio-scoreboard feature that had been implemented, however, some users still managed to be aware of their performance as they were aware of how many notes they hit and missed while playing. Regarding the participants not feeling like they were playing music on stage, all the experiments utilized laptop speakers while playing the game, this could be improved by using a better-quality sound system.

Although there are no direct results to compare to, the generally positive trends of the results presented in figure 6 show that there has been an improvement in the feedback given to the players when compared to the previous version of the game. Thus, all the implemented feedback features were used in the final experiments.

5.1.3 Research Questions 4.1 and 4.2

The 11 participants that were placed in Group 2 participated in the pilot study to answer RQ4.1(*Is there a difference in the performance of the players when playing the adaptive difficulty version of Bongo Beats: Tap With Me when compared to playing the original version with* 1 *difficulty setting?*) and RQ4.2 (*Which version of Bongo Beats: Tap With Me did players find more fun and engaging to play; the adaptive difficulty version, or the original version?*). Of the 11 participants, 6 were female and 5 were male. The majority of participants (6 participants) identified as "non-gamers", 2 participants identified as "Casual Gamers", 2 participants as "Game Developers", and 1 as a "Games Researcher". This information is presented in figure 7.

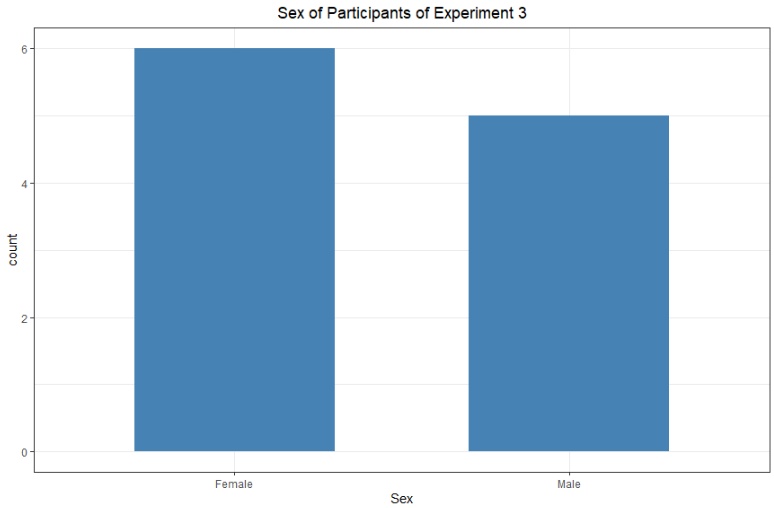
These 11 participants were further split into 2 groups as explained in section 4.3.1 where the first group (Group 2.1) played the game first in fixed-difficulty mode and then proceeded to play the game in adaptive-difficulty mode, and the second group (Group 2.2) played the game first in adaptive-difficulty mode and then in fixed-difficulty mode. 5 participants were in Group 2.1, and 6 participants were in group 2.2. A questionnaire was then handed out to all the participants. This data was then analysed, and the results are discussed in this section.

Research Question 4.1

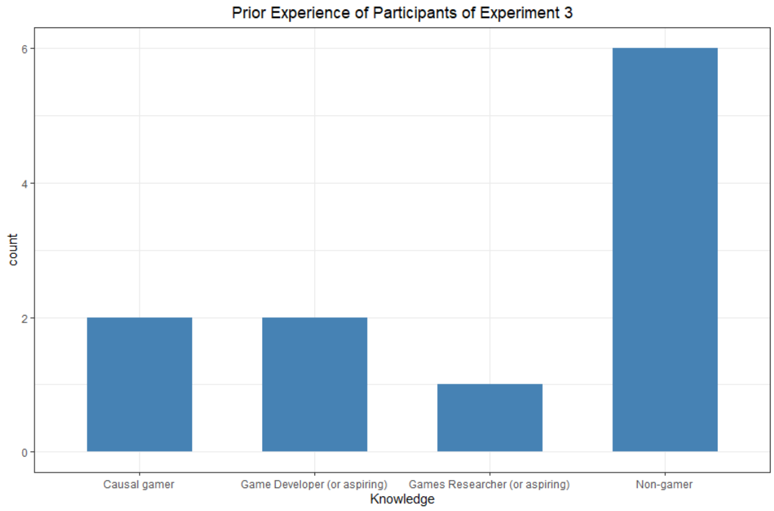
In order to answer RQ4.1, data for each play session was gathered including the game mode (fixed-difficulty or DDA), the number of notes hit, and the number of notes missed. These were then aggregated into the mistakes ratio variable r_s as defined in equation 3.

Table 7 shows the resulting descriptive statistics of the mistakes ratio (r_s) variable for both the fixed-difficulty and DDA game modes. These results can be seen summarised in figure 8. From these figures, it seems that there is no significant difference in performance between the two difficulty modes. In order to test this claim, the Mann-Whitney U test was once again used for pairwise comparison between the two difficulty modes. The median values for r_s were 0.392 for the fixed difficulty mode and 0.386 for DDA mode. The results of the test show that we do not have enough evidence to reject the null hypothesis that the distributions are identical (*Mann-Whitney* $U = 22.5$, $p = 0.8477 > 0.05$).

Thus, we cannot conclude that there is a significant difference in the performance of the players when playing the adaptive difficulty version of



(a) The sex of the participants who participated in the pilot study to answer RQ4.1 and RQ4.2



(b) The prior experience of the participants who participated in the pilot study to answer RQ4.1 and RQ4.2

Figure 7: Background information of the participants who participated in the pilot study to answer RQ4.1 and RQ4.2

	Group	**min**	**max**	**range**	**median**	**mean**	**var**	**std. dev**
Mistakes Ratio	*Fixed Difficulty*	0.314	0.902	0.588	0.392	0.465	0.042	0.204
	DDA	0.311	0.710	0.399	0.386	0.433	0.019	0.139

Table 7: Descriptive statistics for the mistakes ratio (r_s) variable (rounded to 3 decimal places)

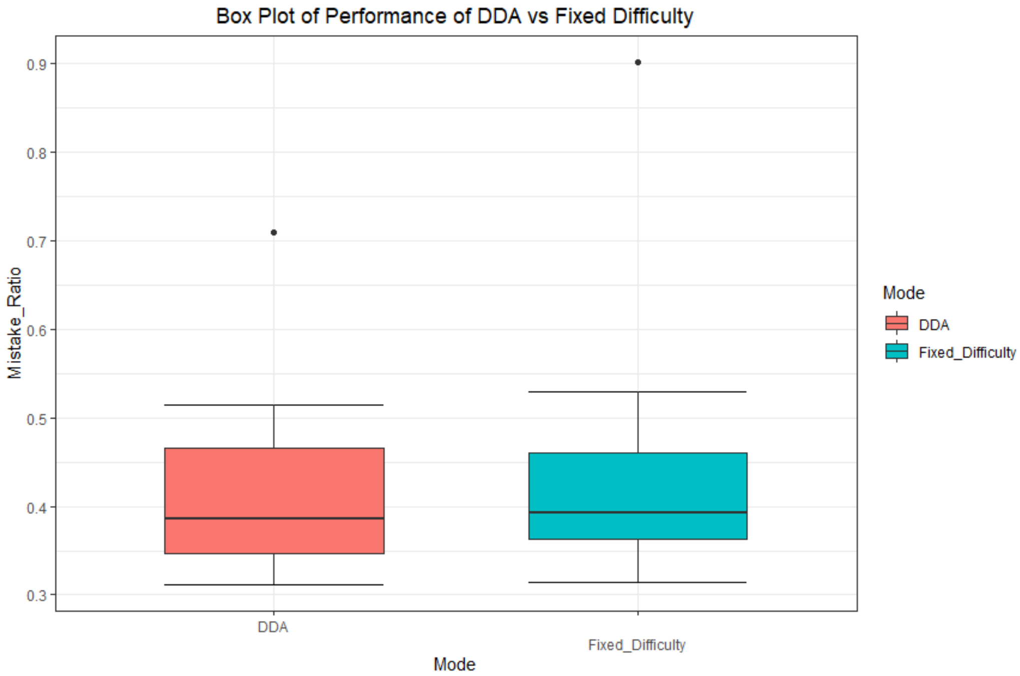


Figure 8: Box Plots showcasing the difference in performance between the fixed-difficulty and DDA game modes.

Bongo Beats: Tap With Me when compared to playing the original version with the fixed difficulty setting. An interesting question that emerged during the analysis of this data was: "*To what degree do the learning effects affect the participants' performance when playing Bongo Beats: Tap With Me*".

Figure 9 shows the performance of all teams over both game rounds. It is interesting to note that 3 teams performed significantly worse in their second round (teams 9, 21, and 23), and all these teams started by playing the DDA version first and then performed worse in the fixed-difficulty version of the game. There were 2 teams that performed better in the second round, the team that performed significantly better (team 5) also started with the DDA version and improved in the fixed-difficulty version, and the team that performed only slightly better in the second round (team 10) started with the fixed-difficulty version and then improved in the DDA version. The remaining teams had a very similar performance in the two rounds, and both started with the fixed-difficulty version before playing the DDA version. Unfortunately, no data was collected on how the DDA adjusted the game's difficulty for each session, so no conclusions can be made.

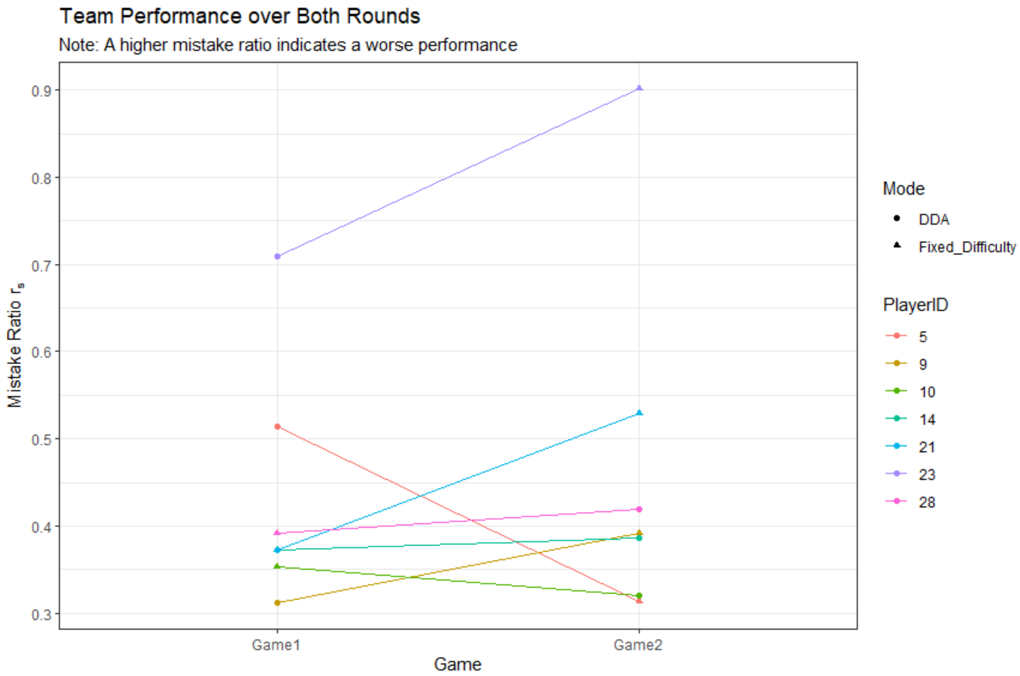


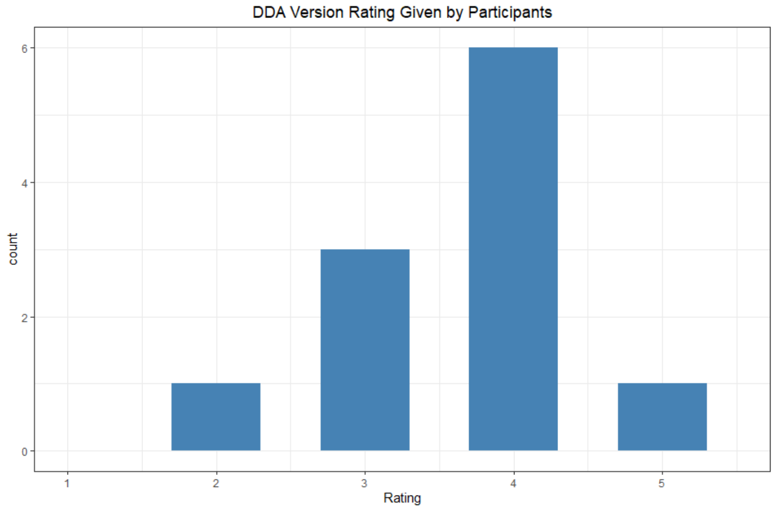
Figure 9: The team performance over both game rounds.

Research Question 4.2

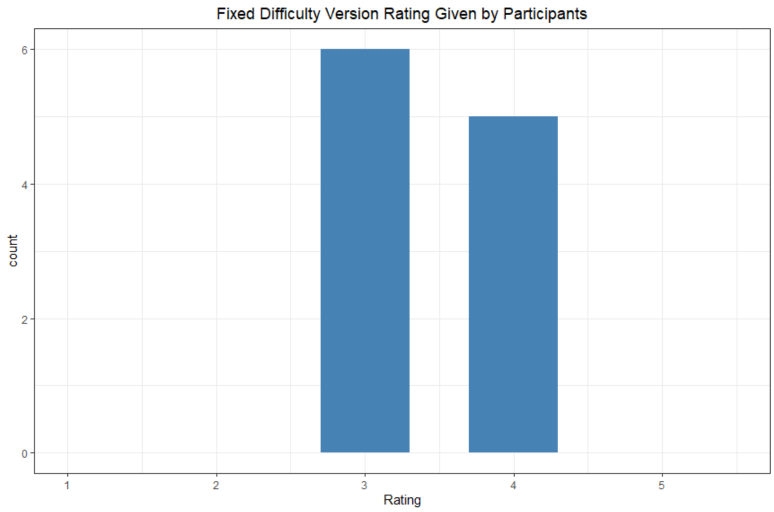
In order to understand which version the participants preferred, data was taken from the survey that was handed out to the participants after playing the game. Participants were asked to rate each version of the game and they were not aware of the differences between each version. Figure 10 shows the ratings for both the DDA and fixed-difficulty versions of the game. Furthermore, participants were also asked directly which version of the game they preferred. The results of this question can be seen in figure 11.

From the results presented in figures 10 and 11, it can be seen that the DDA version of the game was generally the preferred version. After running the Mann-Whitney U test to compare the ratings of each version, where the median rating of the DDA version is 4 and the median rating of the fixed-difficulty version is 3, the results of the test show once again that we do not have enough evidence to reject the null hypothesis that the distributions are identical (*Mann-Whitney* $U = 71$, $p = 0.4651 > 0.05$). Thus, we can make no direct conclusions on which version of the game was preferred by the participants.

The participants were also given an opportunity to provide further feedback on the game. This feedback was then categorised into *positive* or *neg-*



(a) The ratings of the DDA version of the game given by the participants (on a 5 point Likert scale)



(b) The ratings of the fixed-difficulty version of the game given by the participants (on a 5 point Likert scale)

Figure 10: The ratings given by the participants on each version of the game. Note that the participants were not aware of the differences between the versions.

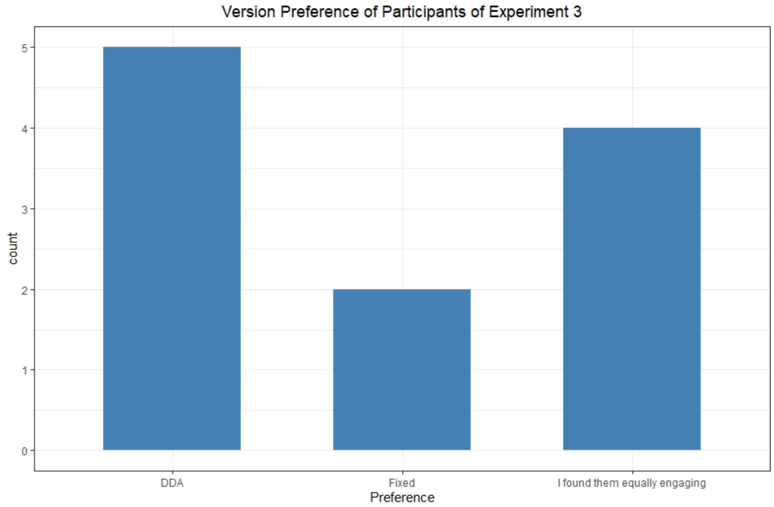


Figure 11: The results of asking the participants directly which version of the game they preferred between the DDA and fixed-difficulty versions of the game. Similarly, they were not aware of the differences between the versions.

ative sentiment and is presented in table 8. The DDA version of the game received more positive remarks (5 positive remarks) than the fixed-difficulty version (2 positive remarks) and also received fewer negative remarks (2 negative remarks) than the fixed-difficulty version (3 negative remarks).

It is interesting to note that two of the negative remarks from the fixed-difficulty version of the game state that "it got a little boring by the end" and "after some time (*it was*) quite slow as there are not as many hits in (*the fixed version*)", and the other remark noted that "in the middle (*I*) got a little frustrated that we missed a lot". These comments justify the implementation of DDA as for the first two remarks the game felt too easy for the participants and thus they felt bored, while for the other remark, the game was too difficult in the middle which led to their frustration. This is further justified by positive comments made for the DDA version that "some parts felt a lot faster which maybe made it more fun", "I liked this because it seemed to have more notes to hit so you had to be very concentrated", and "more fast and engaging ... therefore maybe more fun".

Some negative comments for the DDA version are related to the pacing of the game, for example "it felt like there was less variation" and "the rhythm was slower than the second version". For these cases, the DDA must have made the game too easy for them. This shows that there is still room to improve the difficulty adjustment parameters as it didn't work well for all users. This was also repeated in a general comment where one of

Fixed Difficulty	*Positive*	Loved it; I like the variance in the speed with which the notes followed each other, as well as the variance in note types (left, right, both);
	Negative	It got a little boring by the end; I also really liked this one, but in the middle got frustrated that we missed a lot. But I also like that cause it's part of the game experience.; Good to get to know the game, but after some time quite slow as there are not as many hits as in the second version;
DDA	*Positive*	It was easier because I had more practice but some parts felt a lot faster which maybe made it more fun; I liked this because it seemed to have more notes to hit so you had to be very concentrated; Loved it; I really liked it; Really exciting! More fast and engaging, more points possible and therefore maybe more fun;
	Negative	It felt like there was less variation (a lot of left in a sequence), which made the game a bit boring.; the rhythm was slower than the second version;
General	*Positive*	Very cool game; The song put me in a great mood and made me want to dance a little while I was playing; I like the set up as well. It's good to remind the musician to close their eyes I think. For me it helped with the game experience; Really interesting game;
	Negative	I found it hard to figure out whether I was giving instructions too fast or too slow; I would enjoy a bit more difficulty, by increasing the tempo for instance;

Table 8: Feedback given by the participants in Group 2, categorised by the version of the game they are referring to and whether the feedback had *positive* or *negative* sentiment.

the participants mentioned that they “would enjoy a bit more difficulty, by increasing the tempo for instance”.

Thus, from the results presented in figure 11 and the feedback given by the participants of the experiment, there is still a good indication that the DDA version of the game was preferred over the fixed-difficulty version. It was therefore decided to use the DDA version in the final experiments.

5.2 Final Study

The final study was conducted at Bartiméus school for the VI. A total of 11 VI children participated in the experiment, where 5 of the participants were in the desired target age group (12-16 years old), and 6 participants were younger than the target age group (7-10 years old). The children participated in the experiment in dyads where every dyad consisted of children with mixed visual abilities (blind or different types of visual impairments), and each participant played the game one time as the musician and one time as the coordinator. The musician was asked to choose the song using the controller. After each participant played the game twice, they were taken to a separate room where an adult volunteer helped them fill in the questionnaire which contained questions from the In-Game and Social Presence Modules of the GEQ.

5.2.1 Research Question 1.1

Data about each play session was collected in order to answer RQ1.1 (*How did VI children perform in the inclusive multiplayer rhythm game Bongo Beats: Tap With Me?*). The data collected included the song chosen by the participants, the number of notes hit, and the number of notes missed. This data was aggregated into the mistakes ratio variable r_s as already defined in equation 3. Data on each play session for all the dyads can be seen summarised in figure 12. It can be seen that the performance for the groups of VI children is worse than the performance in the pilot studies, this is expected as the participants of the final study are younger than the participants of the pilot studies. Nothing can be said about the performance between the two rounds for each group as the participants changed roles between the rounds so they were learning a new role each time they played.

Since two different distinct age groups participated in the experiments, it is also interesting to see how the performance differed between the age groups. Table 9 shows the resulting descriptive statistics of the mistakes ratio (r_s) variable for both the “Young” group (aged between 7-10 years old), and the “Teen” group (aged between 12-16 years old). These results can also be seen

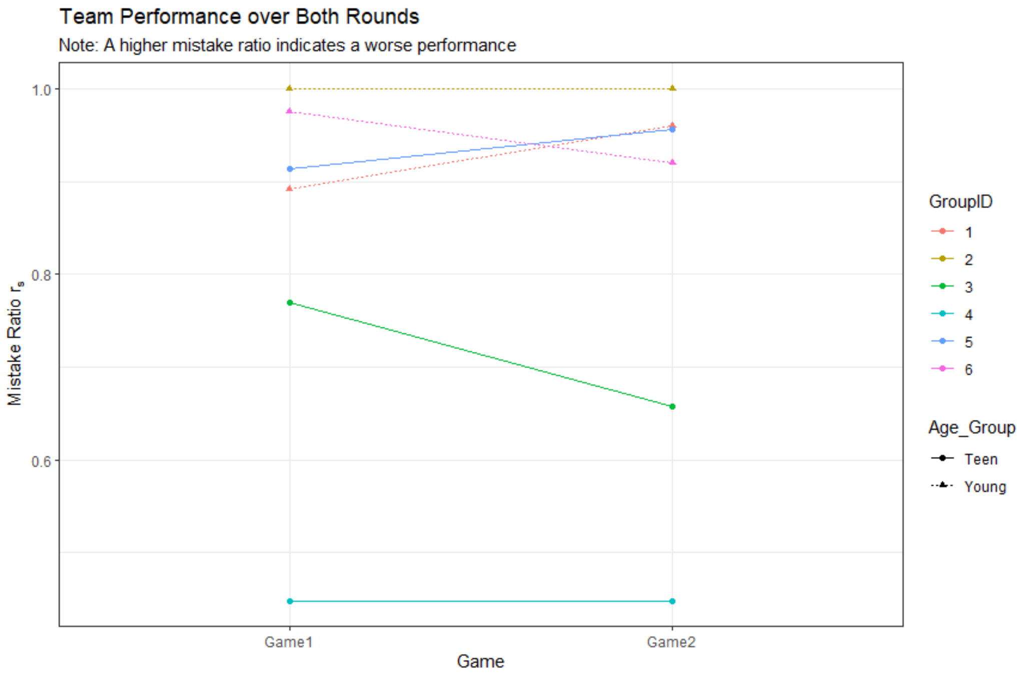


Figure 12: The team performances over both game rounds during the final experiment.

summarised in figure 13. From this information, it can be seen that there is a clear difference in performance between the two groups. In order to test this claim, the Mann-Whitney U test was once again used for pairwise comparison between the two age groups. The median values for r_s were 0.769 for the "Teen" group and 0.968 for the "Young" group. The results of the test show that we have enough evidence to reject the null hypothesis that the distributions are identical (*Mann-whitney* $U = 3$, $p = 0.03534 < 0.05$). Thus it can be concluded that the performance between the age groups differs significantly, where the "Teen" age group performs significantly better than the "Young" age group. This justifies the claim that the game is designed for VI children who are in their teenage years.

	Age Group	**min**	**max**	**range**	**median**	**mean**	**var**	**std. dev**
Mistakes Ratio	*Teen*	0.448	0.957	0.508	0.769	0.749	0.042	0.206
	Young	0.892	1	0.108	0.968	0.958	0.002	0.044

Table 9: Descriptive statistics for the mistakes ratio (r_s) variable of the two age groups of the final study (rounded to 3 decimal places)

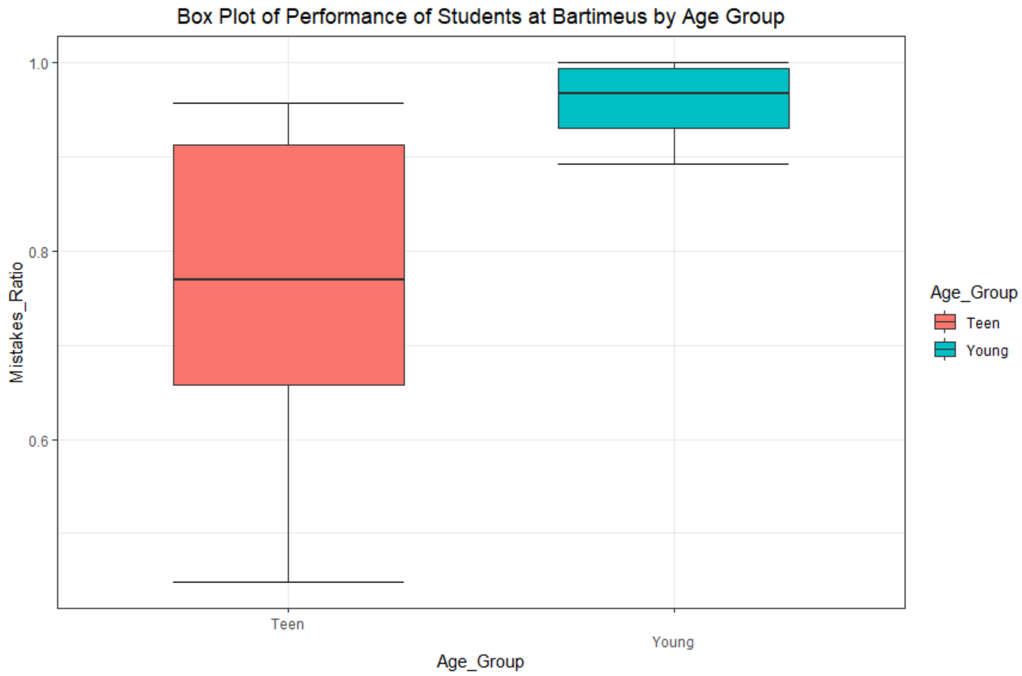


Figure 13: Box Plots showcasing the difference in performance between the "Young" and "Teen" age groups.

5.2.2 Research Question 1.2

In order to answer RQ1.2 (*How did VI children collaborate in the inclusive multiplayer rhythm game Bongo Beats: Tap With Me?*), data was analysed from the social presence module of the GEQ that was handed out to the children of Bartiméus school after they played the game. The scores of the social presence module of the GEQ were calculated as explained in (IJsselsteijn, De Kort, and Poels, 2013). Table 10 shows the descriptive statistics for each category in the social presence module of the GEQ and these are visualised in the box plot shown in figure 14.

	Age Group	**min**	**max**	**range**	**median**	**mean**	**var**	**std. dev**
Psychological Involvement – Empathy	*Teen*	1.5	3.667	2.167	3.167	3	0.792	0.890
	Young	0.667	3.833	3.167	3.333	2.889	1.441	1.200
Psychological Involvement – Negative Feelings	*Teen*	0.6	1.2	0.6	1.2	0.96	0.108	0.329
	Young	0	2.2	2.2	1.4	1.267	0.539	0.734
Behavioural Involvement	*Teen*	1.667	3.667	2	3.167	2.767	0.772	0.879
	Young	0.667	3.833	3.167	3.167	2.667	1.9	1.378

Table 10: Descriptive statistics for the score variables of the social presence module of the GEQ (rounded to 3 decimal places)

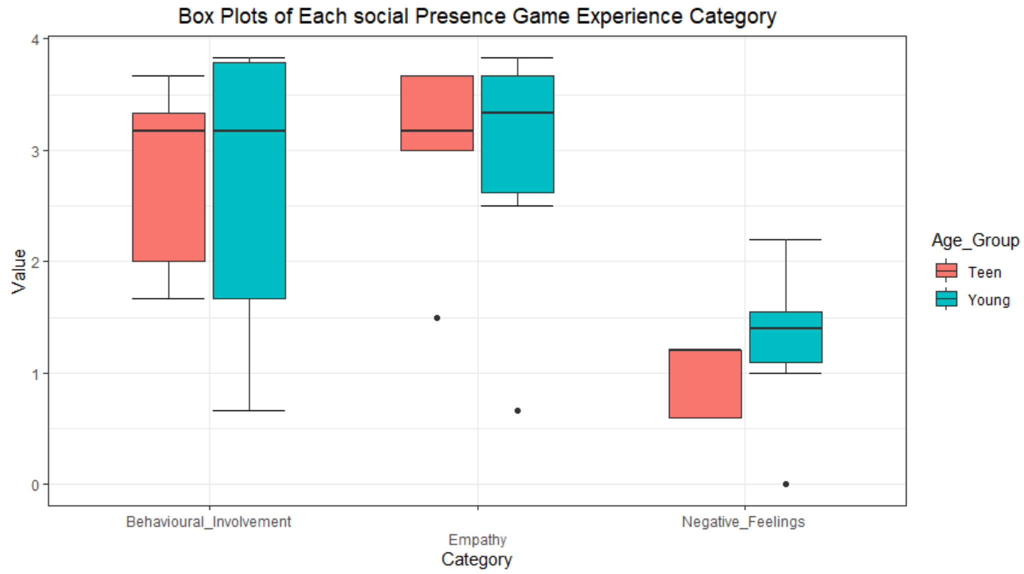


Figure 14: Box Plots showcasing the difference in social presence between the teen and young age groups for the categories in the social presence module of the GEQ.

Behavioural involvement refers to how well the players felt that they were collaborating together. The high value of this result shows that in general, the players felt that they were collaborating well while playing the game. The high variance for the "Young" age group further shows that the game is designed for an older target group as not every participant felt like they collaborated well with their partner while playing the game. A high empathy score shows that the players empathised with each other, felt connected to the other player and enjoyed playing together. Negative feelings refer to how the mood of the players affects each other, and if they had any negative feelings towards each other. A low value is a positive result for this study as it shows that in general, the players worked together to help each other without malicious intent. This value is higher for the "Young" age group and further shows that *Bongo Beats: Tap With Me* should be played by an older audience. These results show that collaborative multiplayer games can help VI children socialise and form better relationships.

The Mann-Witney U test was performed to test if any values significantly differ between age groups. The results can be seen in table 11 and show that we do not have enough evidence to reject the null hypothesis that the distributions are identical between the age groups. Thus, although the values differ between the age groups, we cannot conclude that either age group collaborates better or worse than the other age group when playing *Bongo Beats: Tap With Me*.

	Median (Teen)	**Median (Young)**	**Mann-Whitney U**	**p-value**
Psychological Involvement -Empathy	3.167	3.333	14.5	1
Psychological Involvement - Negative Feelings	1.2	1.4	8	0.229
Behavioural Involvement	3.167	3.167	13.5	0.855

Table 11: The results of the Mann-Whitney U test for the social presence categories of the GEQ for the final experiments, showcasing that we do not have enough evidence to reject the null hypothesis ($p < 0.05$, two-tailed)

5.2.3 Research Question 1.3

In order to answer RQ1.3 (*How did VI children experience the inclusive multiplayer rhythm game Bongo Beats: Tap With Me?*), data was analysed from the in-game module of the GEQ and from the feedback given by the VI children who participated in the experiments. The scores of the in-game module were calculated as explained in (IJsselsteijn, De Kort, and Poels, 2013). The categories for the in-game module are the same as those of the core module used for the pilot studies, however, the questionnaire consists of only two questions per category and was designed for assessing the game experience at multiple intervals during the game session. The questionnaire was only handed out at the end of the play session in order to minimise the stress put on the VI children participating in the experiments by minimising the number of questions asked to them. Table 12 shows the descriptive statistics for each category in the in-game module of the GEQ and these are visualised in the box plot shown in figure 15.

From these results, it can be seen that the "Negative Affect" category scores very low, and the "Positive Affect" category scores very high. This is a very good indication that the VI children enjoyed the game and had a very positive experience overall. Regarding the "Challenge" and "Flow" categories, it can be seen that the responses were mixed and varied depending on the person playing. The median value for "Challenge" was 2 for both age groups, which sits perfectly in the middle of the scale, showing that the game experience for the VI children, in general, was not too easy or too difficult. It is also interesting to note that the "Sensory and Imaginative Immersion" scores were high, and the "Tension" scores were low for all the participants.

	Age Group	min	max	range	median	mean	var	std. dev
Competence	*Teen*	0	3	3	3	2.3	1.7	1.304
	Young	3.5	4	0.5	3.75	3.75	0.075	0274
Sensory and Imaginative Immersion	*Teen*	2	3	1	2.5	2.6	0.175	0.418
	Young	2.5	4	1.5	3	3.083	0.442	0.665
Flow	*Teen*	0	3.5	3.5	3	2	2.625	1.620
	Young	0	4	4	2.75	2.25	2.675	1.636
Tension/ Annoyance	*Teen*	0	3	3	0.5	0.9	1.55	1.245
	Young	0	0.5	0.5	0	0.083	0.042	0.204
Challenge	*Teen*	0.5	3	2.5	2	1.9	0.925	0.962
	Young	0	3.5	3.5	2	1.583	1.842	1.357
Negative Affect	*Teen*	0	0	0	0	0	0	0
	Young	0	0.5	0.5	0	0.167	0.067	0258
Positive Affect	*Teen*	2	4	2	4	3.4	0.8	0.894
	Young	1	4	3	4	3.333	1.467	1.211

Table 12: Descriptive statistics for the categories of the in-game module of the GEQ (rounded to 3 decimal places)

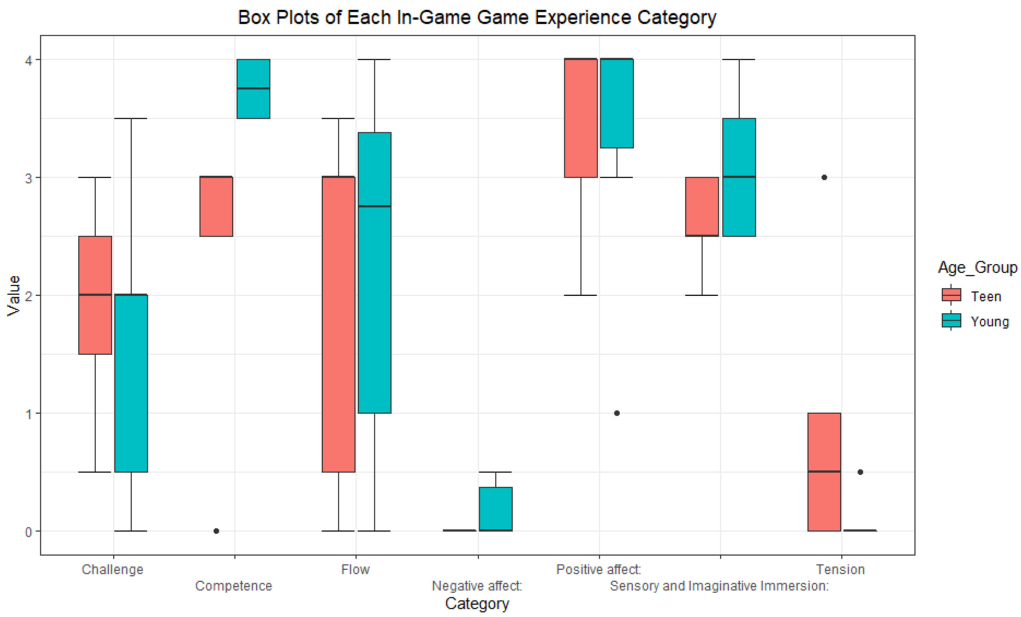


Figure 15: Box Plots showcasing the difference in game experience between the teen and young age groups for the categories in the in-game module of the GEQ.

Although the “Young” age group performed worse than the “Teen” age group (refer to the results in section 5.2.1) when playing the game, it is interesting to note that they still felt that they performed well. In fact, the participants in the “Young” age group rated themselves higher in competence than those in the “Teen” age group.

The Mann-Whitney U test was once again performed to test if any of the values differ significantly between age groups and the results can be seen in table 13. There is enough evidence to reject the null hypothesis that the distributions are identical between age groups for the “Competence” category. Thus it can be concluded that the “Young” age group felt more competent when playing the game than the “Teen” age group when playing *Bongo Beats: Tap With Me*. Otherwise, the game was experienced similarly for both age groups.

	median (Teen)	**median (Young)**	**Mann-Whitney U**	**p-value**
Competence	3	3.75	0	0.006
Sensory and Imaginative Immersion	2.5	3	9	0.290
Flow	3	2.75	13.5	0.853
Tension/ Annoyance	0.5	0	225	0.138
Challenge	2	2	17.5	0.708
Negative Affect	0	0	10	0.221
Positive Affect	4	4	14.5	1

Table 13: The results of the Mann-Whitney U test for the in-game categories of the GEQ for the final experiments, showcasing that we have enough evidence to reject the null hypothesis for the “Competence” category ($p < 0.05$, two-tailed)

The VI children who participated in the experiment were also given the opportunity to provide further comments or feedback about the game. Their responses were split into 3 categories; positive, negative, and feedback. These can be seen in table 14 (note that they were translated to English from Dutch). Although not every participant gave feedback, it can be noted that there were a total of 5 positive comments all expressing their enjoyment of playing the game. The 2 negative comments refer to the lack of connection the players felt with the music, this was found to be a recurring theme and is discussed further in section 6. There were also 5 suggestions on how the game can be further improved. One VI participant played the game in visual-haptic mode, thus they recommended improvements to the visuals of the game, stating that “the game should be adjusted in speed and colour” by using "more contrast colours". Other comments include adjusting the game parameters such as increasing "the time between the vibration and the click", and implementing a "slower version of the game" to practise. One participant suggested increasing the variety of drum sounds, and one participant also mentioned implementing a single-player version of the game.

	Comments
Positive	The game was fun to play. It was good to do; I found it fun to play; I really enjoyed doing it; It was just a lot of fun; I would like to do it more often;
Negative	The game doesn't really have anything to do with the music. I felt I could also play the game without the music; I miss the connection with the music.
Improvements	Some things in the game should be adjusted in speed and colour. It is helpful if more contrast colours are shown. The visually impaired can see that better; It would be nice if there were more drum sounds in the game; The time between the vibration and the click should maybe be a bit longer; It would help if you could play it in a slower version first. Then you can practise with the beat. After that, the tempo could go up; I would like it if you can also play the game alone. For example, feeling vibrations in the Xbox controller;

Table 14: Feedback given by the VI children who participated in the final experiment. Note that this was translated to English from Dutch.

5.3 Expert Interviews

Expert interviews were conducted with a VI student studying at Utrecht University, 2 PhD students who are working in the field of developing assistive technologies for the VI, their supervisor, and our contact person from the Dutch Visio organisation. The VI student played the game for 40 minutes, playing various different songs, and both the coordinator and musician roles before participating in a semi-structured interview.[18]The other participants all joined in one session together where they each played the game once in each role before participating in an informal discussion.

Following the expert interviews, the recordings were transcribed and the NVivo[19] software was used to conduct a thematic analysis. The transcribed interviews were coded inductively, where the interviews were analysed and codes were created from this. The resulting codebook can be seen in table 15.

Code	**Description**	**Nr. of Participants**	**References**
Collaboration	Points about collaborating with the other play	5	12
Connection with the Music	Did the song help with the gameplay? Did it feel like a rhythm game?	4	6
Controllers	Comments on the game controllers.	1	3
Difficulty	How did the participants experience the difficulty of the game and what they find to be difficult/ easy.	3	6
Feedback	The thoughts on the quality of the feedback given to the players when playing the game	3	8
Gameplay	General comments on the gameplay	1	5
Games for the VI	Any comments about making games for the VI, or how a VI person experiences this game.	2	6
Song	Comments on the song choice and music in the game.	2	9

Table 15: The resulting codebook from the thematic analysis done on the expert interviews

Three themes then emerged from the thematic analysis of the expert interviews, these are Collaborative Game, Game for the VI, and Rhythm Game. Table 16 showcases how each code fits into each theme. These themes will be discussed in the remainder of this section.

5.3.1 Collaborative Game

The first theme that emerged from the thematic analysis was 'Collaborative Game'. All comments related to collaborative games or playing the game with others make up this theme. A word cloud of the most frequent words can be seen in figure 16.

[18]The interview outline can be found in the appendix.

[19]https://lumivero.com/products/nvivo/

Theme	Codes	Nr. of Participants	References
Collaborative Game	Collaboration Gameplay	5	17
Game for the VI	Controllers Feedback Games for the VI	4	17
Rhythm Game	Connection with the Music Difficulty Song	4	21

Table 16: The themes that emerged from the thematic analysis of the expert interviews

The attitude of the VI student was very positive towards the collaborative aspect of the game, and about playing the game with people of mixed abilities together. Some comments mentioned by the VI student include "*It is really a teamwork", "as a kid, I can bother people saying "Play with me, play with me this game." I think it is, you know it is not all repetitive, you know you can do (different things)? Yeah, probably I would want to play more.*", "*If you have a friend, maybe you can try to have some hints*", "*I personally like it (the collaboration) because it is all about, you know, what we are doing in university as well, what we are doing in life as well. We have to depend on people sometimes and mostly otherwise it doesn't make sense to be perfect on your own, you know, this type of thing. And I think it is a good practice for children. Especially they are really egocentric at that moment, probably you know. And yeah, I love it and I think it makes it more fun. I don't know. I love it.*", and "*Maybe it'll be (improve your relationship with the other player) because you can have some memories, fun memories about the game*".

The PhD students and Dutch Visio expert had a more mixed response, where half of them felt that it was collaborative while the other half didn't feel like they were collaborating with the other player. Some comments regarding the game being collaborative include "*it was like I calibrated myself to ... when you said things and my response from hearing it to time was calibrated as well.*", "*I think a little bit (collaborative) I got used to the cadence yeah*", "*if we're like, really going together, what was what I felt as collaborative is that you were tapping. And I was a little bit dancing, right? So we established this joint space.*", and "*So when ... 'PhD Student' said, like "nice", and that was like really very much feeling like we're doing it together. And we started to give feedback to each other somehow. That really felt like a collaboration. And so when you were doing at some moment I tried to also tell you "nice".*". The comments where the PhD student and Dutch Visio expert mentioned that they did not feel like they were collaborating were "*I don't think we bothered creating the joint space*", " *we didn't use this joint space to react*",

Figure 16: A Word Cloud highlighting the most frequent words mentioned in the 'Collaborative Game' theme.

and "*I didn't feel like it was collaborative.*".

5.3.2 Game for the VI

Another theme from the thematic analysis was 'Game for the VI'. Every comment regarding the features implemented to make the game accessible for the VI and the improved feedback features make up this theme. A word cloud of the most frequent words can be seen in figure 17.

Regarding making games for the VI, comments were mentioned by the VI student and the Dutch Visio expert. All the mentioned comments were of a positive sentiment and it was especially mentioned that they liked the idea of making games that people of mixed abilities can play together. The VI student mentioned "*I like the idea that anyone can play even without vision and with vision*". When speaking about games that are already available on the market, the VI student mentioned that "*if a kid, a blind kid wants to play with a sighted friend, after some time, I think the sighted one is getting bored first.*" and " *my focus is that you know not creating something different, but adapting (games) that are already existing to everyone*". The Dutch Visio expert mentioned "*I think that the blind can play it as well as the sighted, that's also very nice because mostly they have to play something of a lower level*".

Regarding the improved feedback features that have been implemented

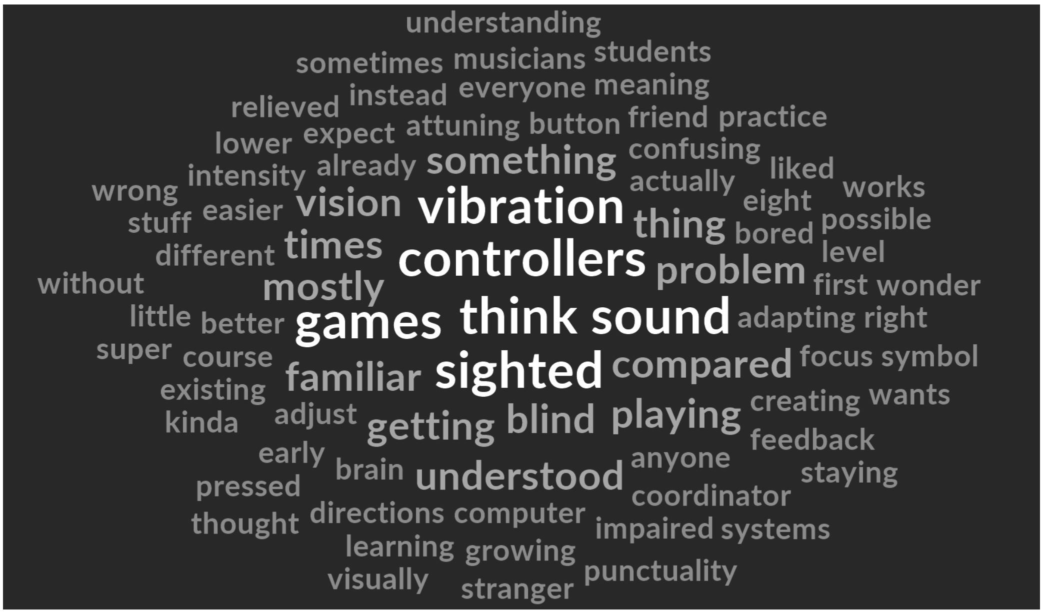


Figure 17: A Word Cloud highlighting the most frequent words mentioned in the 'Game for the VI' theme.

in the game and the use of controllers, there were comments with positive sentiments made by one PhD student and the VI student, as well as suggestions on how to further improve these features made by the PhD students. The positive comments made by the VI student were "*I didn't expect to do it with, you know, controllers ... that's why once I see the controllers I understood what we are doing and got relieved a bit.*", "*I find controllers as a kid easier compared to computer directions*", and "*I understood when I pressed the wrong button*". One PhD student mentioned that "*I liked the vibration feedback when getting it right*". The suggestions made by the PhD students were to change the sound of the cymbal that plays when you hit the note correctly as they "*didn't like the cymbal because it makes the song sound even stranger to me*", and "*is it possible to adjust (increase) the intensity of vibration?*" when referring to the vibration that is felt by the coordinator.

5.3.3 Rhythm Game

The final theme from the thematic analysis was 'Rhythm Game'. This theme is made up of all the comments related to the game feeling like a rhythm game, where the difficulty comes from in this game, and the song choices of the game. A word cloud of the most frequent words can be seen in figure 18.

Regarding the choice of songs and variety of music in the game, the

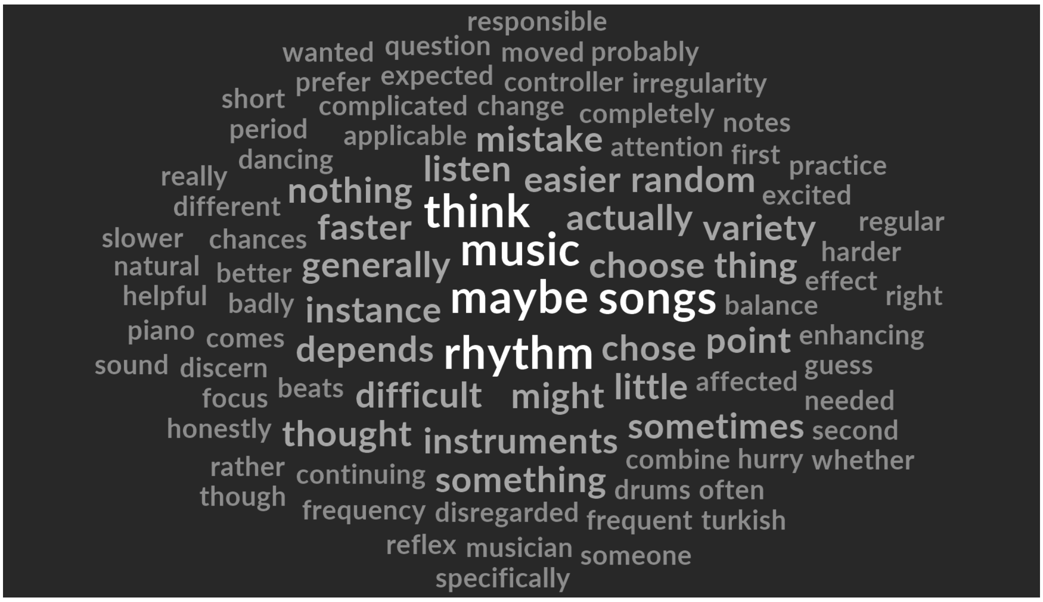


Figure 18: A Word Cloud highlighting the most frequent words mentioned in the 'Rhythm Game' theme.

VI student gave positive comments as well as some suggestions on different music styles that can be added to the game, and one PhD student also made a neutral comment on the variety of songs available. Regarding the music variety, the VI student mentioned that they wanted to try different songs and "*I like it because I choose the song myself*", "*Is it the question of whether I like the songs or not? It depends. I like the ones I chose actually.*", "*and sometimes I like the song you chose. ... I think it is good to have a variety of songs here.*". With regards to different songs that they would like to see in the game, the VI student said that they would like to see "*much more variety of music, for instance, faster music*", "*maybe kids songs*", and "*maybe I would prefer for instance a rock song in it*". One PhD student also made the comment that "*for me, in the first song, music was helpful because (it had a lot of drums) that it comes on the rhythm. So it was kind of enhancing. ... But in the second one, the rhythm was difficult to discern, so my attention was moved. So it really depends*".

The rhythm game aspect was a big point of discussion between the PhD students and Dutch Visio expert, they made mostly negative comments on this aspect, and also gave suggestions, stating that "*it felt like a game of reflex rather than a game of rhythm*", "*I just completely disregarded the song*", "*but we both didn't listen to the music very much, I think that will change when you do it more often that you listen more to the music*". Their suggestion was to increase the frequency of the notes to make regular intervals on the beat so that it feels like you are playing the drums to the beat of the song mentioning "*what if the beat would be generally slower, or maybe the beat is more regular and we would have a bit more time*", and "*the beats didn't feel as on the rhythm is I think I would have expected it to be and that might be the frequency*". The VI student also made similar comments on the frequency of the notes, stating "*I felt this is random notes because in any case, the song was continuing. Maybe it might have some effect on the song itself, you know. Maybe if you made a mistake, the song should be affected badly.*", and "*maybe you can add some more (instruments) for instance, a drum and a piano*".

There was also a related discussion on defining the difficulty of the game. One PhD student mentioned that the game felt harder when there were fewer notes in the level as "*the irregularity of it was sometimes a little bit harder to combine with the sound*". The VI student also commented on the difficulty of the game and stated that "*even with a short period of time, you can be better with practice*", "*(The musician role) was easier, and even though you make mistakes, you just thought that this is not only about you, but the coordinator is specifically responsible*", and suggested increasing the difficulty by making"*it more complicated, you know, add more instruments*".

6 Discussion

In this section, the results will be discussed and contextualised with respect to the literature for the implemented improvements and experiments conducted. The talking points of collaboration versus rhythm and the target age group will then be discussed. Following this, the limitations of the study will be mentioned, as well as ideas for future research.

6.1 Results - Implemented Improvements

The results from the pilot study for all the implemented changes to the game all show a positive trend towards improving the overall experience of the game. The note-generation algorithm was successful in generating levels based on new songs that are similar to the level that was implemented in the previous study. This aligns well with the research presented in section 2.1.1, and showcases the accuracy of the state-of-the-art beat tracking algorithms available today as beat instances were successfully generated for various songs of different styles.

The implemented improved feedback features were also very well received.

The majority of the participants understood well what was happening in the game when playing in audio-haptics mode, and they claimed that the gameplay was clear and easy to understand. Furthermore, although not formally participating in any experiment, an adult blind gamer working at Bartiméus School who played the old version of *Bongo Beats: Tap With Me* said that he was pleasantly surprised after playing the new version of the game because all of his suggestions to improve the game were implemented. Another feature that was very well received was enabling the musician role to be played using a controller as many participants stated that it also felt good to hit the note correctly. These results further show that the guidelines specified by (Garcia and Almeida Neris, 2013) and (Yuan, Folmer, and Harris, 2011) provide a good basis to consider when implementing accessibility features for the VI. Although it was explained to all the participants, it was interesting to note that none of them used the audio-scoreboard feature. This is probably because there is no downtime while playing the level where the participants have time to check their score without having to look out for the next note. The quality of the recording for the audio-scoreboard and interface items could also be improved as this was done using a laptop microphone and some words are not very clear.

With regards to the implementation of DDA in the game, there are signs to indicate that this feature has resulted in an overall better experience. Particularly from the qualitative feedback presented in section 5.1.2, it has been shown to be more engaging to the players. This aligns well with the research presented in section 2.4 regarding implementing adaptive difficulty in games. However, finding the correct parameters for adjusting the difficulty levels is tricky as the values of the parameters are unique for every game and require a lot of user testing to implement correctly, and thus there is still room for improvement in the current implementation. In particular, DDA was implemented with the idea that players would only play 1 or 2 levels in the game, thus a very small window size was chosen to ensure that they would feel the effect of DDA. For a full release of the game, it would be better to track the players' performance over multiple levels to get a better understanding of the difficulty level that suits them.

The response to the difficulty parameters that were adjusted also received a mixed response. While some participants claimed that they found an increase in notes to be more challenging, some participants claimed that having fewer notes in the level made the game harder as they were not as regular. This result contradicts the guidelines mentioned in (Liang, W. Li, and Ikeda, 2019), however, the authors in that study focused on single-player rhythm games with visuals, and *Bongo Beats: Tap With Me* is a collaborative multiplayer game with accessibility features for the VI, thus it can be argued that different elements define the difficulty in this case. There were also suggestions to increase the variety of note types or instruments that could be added to the game to make it more challenging. This could be achieved by including different drums in the game (for example adding also a cymbal, a snare drum, and a bass drum to the game) to increase the number of note types, however, this would require a redesign for the coordinator role in haptics-mode. Thus, it would be also interesting to also explore how to make a collaborative rhythm game where players feel connected to the music, but are also connected to each other.

6.2 Results - Final Experiment

The main aim of this study was to provide a deeper understanding of developing collaborative multiplayer games with accessibility features for the VI. Through the results of the final experiments, it has been shown that the game *Bongo Beats: Tap With Me* was successful in creating a space where VI children can collaborate well together while still having fun. Furthermore, the level of challenge of the game was perceived to be at the right level by the players.

This is further backed by observations made during the experiments, where it was clear (especially in the older age group) that the children thoroughly enjoyed the game and played well together. The excitement in some of the participants was also very noticeable during the experiments. These results support the findings in literature about developing collaborative games presented in section 2.3. In particular, by developing a game where the guidelines for creating an inclusive space presented by (Trust, n.d.) and encouraging collaboration amongst VI children presented by (Roe, 2008) are used as the pillars of the game's design, it has been shown that the game had a very positive response from the VI children who participated in the experiments.

The experiment conducted in this study which was similar to the experiment done in (Potamianos, 2022) further backs the claims made by the author about the reception of the game by the VI, and the importance of research in this field. Playing video games and sharing these experiences with friends is something that children take for granted, and is something that VI children are not capable of experiencing. This highlights the need for developers to create accessibility features in their games to enable VI players to play them, as (Metatla et al., 2020) showed that this will help reduce the culture gap between VI and sighted children.

6.3 Results - Expert Interviews

Although the game was received very positively by the participants in the pilot studies and those in the final experiments, the results from the expert interviews were more mixed.

The VI university student was generally very positive and excited about the game and the study. When speaking about past experiences playing games, the student stated that "*if a kid, a blind kid wants to play with a sighted friend, after some time, I think the sighted one is getting bored first*". This further support the argument made by (Metatla et al., 2020) that designing games specifically for VI players would lead to them being another source of isolation for them. This also highlights the importance of creating games that could be enjoyed by VI and sighted people equally and showcases the value of continuing research in this field. The student also made comments highlighting the importance of encouraging collaboration between VI and sighted children as they "*have to depend on people sometimes*".

It is also encouraging to note that many of the implemented feedback features to make the game accessible for the VI were not mentioned during the expert interviews. These comments tend to come up when something doesn't feel right or is unclear, as was the case in (Potamianos, 2022). Thus, this is a good indication that the gameplay is clear and understandable.

The majority of the comments with a more neutral or negative sentiment were made by the PhD students. These comments focused more on the rhythm aspect of the game. These comments will be discussed in further detail in the next section.

6.4 Collaboration vs. Rhythm

The expectation from the PhD students and the Dutch Visio expert who participated in the expert interviews was that the game was designed to teach rhythm skills to VI children. This is understandable as the game is primarily a rhythm game. However, the focus of this study was on creating a collaborative multiplayer game that can be enjoyed by VI and sighted children equally.

Many suggestions were made on how to improve the rhythm aspect of the game. Primarily, when playing the game, the players do not feel like they are contributing to the music, as the bongo notes seem random, and in many cases, the bongos do not sound like they are part of the song being played. In fact, a music teacher who had some time to play the game commented on this aspect. He had a different expectation for the game, yet an interesting one that could be looked into in future studies, to use the game during music lessons. During the expert interviews, it was suggested to increase the frequency of the notes and ensure that they follow a repeating pattern. This would thus add a continuous bongo track to the song and reduce the feeling of randomness. This could however be tricky to implement as it may impact the collaboration aspect of the game since the conductor role would be more predictable and thus less engaging. In fact, during the same interview, it was suggested to create a single-player mode as this would be better to teach rhythm skills. This further highlights that the expectation of this participant was not aligned with the goal of this study.

Another suggestion to improve the rhythm aspect of the game was to inform the coordinator exactly two beats before the note should be hit so that the musician will have to listen to the beat before hitting the note, rather than hit it straight away. This would thus require the players to have a better connection with the music. This feature however could be difficult to implement as people sense the haptic feedback at different rates and the communication time between players can also vary. This could also be a feature that would better suit a single-player game.

Creating a traditional rhythm game to teach rhythm skills to the player will have different design principles and goals to those of a collaborative game and it has been shown that the goals sometimes do not align. Both are valid ideas that can have a positive impact on improving the quality of life of VI people and future studies should therefore consider which principle they would like to focus on.

6.5 Target Age Group

Another point of discussion that emerged from analyzing the data was that of the target age group. The results shown in section 5.2.1 show that the performance of the children in the younger age group (aged between $7 - 10$ years old) was poor. Some of the children were confusing “left hand” and “right hand” which led to some frustration when playing the game. This however did not reflect in the results shown in section 5.2.3 where the younger age group felt like they performed well in the game. Although the game is designed for a slightly older audience, it has thus been shown to still be enjoyable for younger children. However, their regular poor performance might be demotivating and lead to frustration. An interesting solution to this would be to create a difficulty setting for younger children where they do not need to hit a particular “left” or “right” bongo, but rather can hit any bongo. Until such a solution is developed, it is still recommended that the game is played by older children (aged 12 and over).

6.6 Limitations

Following the discussion of the results of the study, some limitations of this thesis will be presented in this section. In particular, there were four main limitations to this study;

1. The sample size and age of participants for the final experiments,
2. The amount of time each participant had with the game,
3. A lack of focus on accessibility features in the sighted mode, and
4. The accessibility of the required hardware.

As is often the case with studies related to children with disabilities, gathering enough participants for the experiments was challenging. There were only 11 participants available for the final study, and the majority of them (6) were younger than the target age group. Thus, the majority of the results of this study were not statistically significant. However, the information presented in this thesis can still provide a better understanding of how VI children experience and collaborate when playing collaborative multiplayer games and could serve as a good guideline for developing games for the VI.

Furthermore, due to the limited time availability of the participants, each participant could only play the game 2 times (one time in each role). The only participant who played the game for a longer period (approximately 40 minutes) was the VI university student who participated in the expert interview. It was noted that the skill of the player improved over time, and they had the opportunity to try out different songs in the game. Thus, it would have been very insightful to have more participants playing the game for a longer period.

Furthermore, due to the scope of the study and time constraints, the majority of research and development time focused on improving the audio and haptic modes of the game. Although some time was spent on adding basic visuals to the game, no focus was given to ensuring that these visuals are fun and accessible. One VI student who participated in the final experiments had chosen to play the game with the visuals enabled, and they said that it was difficult to understand the visuals given their visual impairment. The student suggested adding more contrast between the colours and adding the option to increase the size of all of the visual elements. It would also be interesting to implement a colourblind mode using filters as suggested in (Khaliq and Torre, 2019) to ensure that any player with any form of colourblindness could also play the game.

Finally, the cost of hardware is a limitation to the overall goal of developing a game that is accessible to both VI and sighted people. At the time of writing, a VR headset compatible with *Bongo Beats: Tap With Me* costs a minimum of $500[20], and a high-end computer or laptop is also required to run the game. This is a very high barrier to entry for an individual to invest in to play such a game and makes the game not accessible to anyone who does not afford it. Future work should focus on lower-cost alternatives to the hardware required to play this game.

6.7 Future Research

The results in this thesis show that *Bongo Beats: Tap With Me* had a positive effect on the participants, and it has been shown that the children who played the game collaborated well, enjoyed the game and were generally excited to play. Furthermore, the importance of developing games with accessibility features for the VI has also been highlighted.

With relation to improving the game further, as mentioned in section 6.4, future researchers can further explore the theme of collaboration, or work to improve the rhythm game elements. With regards to improving the game to teach the user rhythmic skills, the level generation algorithm can be further improved by utilizing drum or music generation algorithms (such as (Nuanáin, Boyer, Jordà Puig, et al., 2015), (Dostál, 2005) and (Ji, Luo, and Yang, 2020)) rather than beat tracking to generate better-sounding drum tracks over the songs that will enable the users to feel like they are contributing to the song. The ideas mentioned in section 6.4 with regards to improving the rhythm aspect of the game also have the potential to be explored further.

Future research can also explore the collaborative aspect of the game further. In particular, it would be interesting to explore how VI children and sighted children would collaborate and play *Bongo Beats: Tap With Me* together. As was the case in (Potamianos, 2022), it was difficult to find mainstream schools with VI students who were willing to participate in the experiment and thus it was not possible to test this aspect in this thesis. However, the original aim of the game was to create a game where VI and sighted children can play together. This has several advantages including educating sighted children on what it is like to have a visual impairment and reducing the feeling of isolation for VI children by giving them a platform to express themselves with their sighted peers. Furthermore, it would be interesting to perform an experiment similar to the one done in (Verver,

[20]according to https://www.vive.com/us/product/

Vervloed, and Steenbergen, 2020) where participants would be given plenty of time to play the game and their interactions would be recorded and coded to better understand how *Bongo Beats: Tap With Me* can facilitate interaction between VI and sighted children.

Finally, researchers could focus on developing and creating a version of *Bongo Beats: Tap With Me* that would be available to access and play by anyone. This would include improving the visuals as mentioned in section 6.6, and creating a version of the game that does not rely on expensive hardware. The final goal of this research would be to share the game with schools and individuals so that it can be experienced by anyone.

7 Conclusion

The lack of opportunities and experiences that teach and encourage collaboration and cooperation between VI and sighted children is evident. This skill will play an important role in the lives of the VI, as was stated by a VI university student who participated in this study "*We have to depend on other people sometimes*". The main goal of this thesis was to provide insights into creating games with accessibility features for the VI that encourage collaborative behaviour. The presented results show that such a game supports collaborative behaviour amongst the players, while overall retaining a fun and exciting experience.

To continue the work done in the previous study by (Potamianos, 2022), several improvements were made to the game *Bongo Beats: Tap With Me* including;

1. Adding 14 new songs to the game by using an algorithm based on beat tracking, to create a more varied library of songs and ensure that every player will find a song that they enjoy.

2. Creating several improvements to the feedback given to the players while playing the game, which made the game mechanics easy to understand.

3. Implementing adaptive difficulty in the game, which helped to increase player engagement.

The results of the pilot studies show that all these improvements have positively affected the gameplay experience of *Bongo Beats: Tap With Me*, while still retaining the design and goals of the original implementation. In particular, the added songs were successful in maintaining the feel of the original handcrafted level, the feedback improvements made the game easier to understand, and the adaptive difficulty algorithm contributed to ensuring the game is exciting for players of various skill levels and capabilities.

To answer the main research questions in understanding how VI children perform, collaborate, and experience the improved version of *Bongo Beats: Tap With Me*, experiments were conducted at Bartiméus school with VI students aged between 8 and 16 years old. With regards to performance, the results of the experiments show that the children in the target age group (12 and older) performed fairly well, however, the scores of the children in the younger age group were very low. The collaboration scores were evaluated using the Social Presence Module of the GEQ, and these showed that the participants experienced positive behaviour towards each other, and showed a high degree of empathy towards each other while exhibiting a low score for negative feelings towards each other. The In-Game Module of the GEQ was used to understand how the VI children experienced the game. The results from this questionnaire showed that the challenge of the game was perceived to be at the right level, and the participants scored themselves highly in competence, positive affect, and immersion scores. The negative effect and tension categories scored very low.

The insights provided in this thesis show that by incorporating guidelines presented in the literature on encouraging collaborative play between children and creating accessible experiences for the VI into the pillars of a game's design, then it is possible to create collaborative games with accessibility features for the VI that are also enjoyed by a sighted audience. If these principles are followed by studios working on mainstream games, then it is possible to create high quality games that can be played by both VI and sighted children, and thus bridge the quality gap between games specifically designed for the VI and mainstream games. Furthermore, although some work is still required to release the game *Bongo Beats: Tap With Me*, the source code for the current version of the game has been made available online with the hopes that this study can be continued and that the game can be released and made available for VI and sighted children.

Glossary

Term	**Definition**
AI	Artificial Intelligence
BI	Beat Information
BLSTM	Bidirectional Long Short-Term Memory
CNN	Convolutional Neural Network
CRNN	Convolutional Recurrent Neural Network
DDA	Dynamic Difficulty Adjustment
GEQ	Game Experience Questionnaire
IBI	Inter-Beat-Interval
MIDI	Musical Instrument Digital Interface
MIR	Music Information Retrieval
PLP	Predominant Local Pulse
QTVI	Qualified Teachers for the Visually Impaired
RMS	Root Mean Square
RQ	Research Question
TA	Teaching Assistant
VI	Visually Impaired
VR	Virtual Reality

Table 17: List of terms and abbreviations used.

References

Archambault, Dominique and Damien Olivier (2005). “How to make games for visually impaired children”. In: *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pp. 450–453.

Atkinson, Matthew T, Sabahattin Gucukoglu, Colin HC Machin, and Adrian E Lawrence (2006). “Making the mainstream accessible: redefining the game”. In: *Proceedings of the 2006 ACM SIGGRAPH Symposium on Videogames*, pp. 21–28.

Böck, Sebastian and Matthew EP Davies (2020). “Deconstruct, Analyse, Reconstruct: How to improve Tempo, Beat, and Downbeat Estimation.” In: *ISMIR*, pp. 574–582.

Böck, Sebastian, Matthew EP Davies, and Peter Knees (2019). “Multi-Task Learning of Tempo and Beat: Learning One to Improve the Other.” In: *ISMIR*, pp. 486–493.

Böck, Sebastian, Florian Krebs, and Gerhard Widmer (2016). “Joint Beat and Downbeat Tracking with Recurrent Neural Networks.” In: *ISMIR*. New York City, pp. 255–261.

Böck, Sebastian and Markus Schedl (2011). “Enhanced beat tracking with context-aware neural networks”. In: *Proc. Int. Conf. Digital Audio Effects*, pp. 135–139.

Burgoyne, John Ashley, Ichiro Fujinaga, and J Stephen Downie (2015). “Music information retrieval”. In: *A new companion to digital humanities*, pp. 213–228.

Cannam, Chris, Emmanouil Benetos, Matthew E. P. Davies, Simon Dixon, Christian Landone, Mark Levy, Matthias Mauch, Katy Noland, and Dan Stowell (2019). “Mirex 2019: VAMP Plugins from the centre for digital music”. In: *Queen Mary, University of London*.

Chakraborty, J, J Hritz, and J Dehlinger (2014). “Preliminary results in the understanding of accessibility challenges in computer gaming for the visually impaired”. In: *Inclusive Designing: Joining Usability, Accessibility, and Inclusion*. Springer, pp. 83–92.

Davies, Matthew EP, Norberto Degara, and Mark D Plumbley (2009). “Evaluation methods for musical audio beat tracking algorithms”. In: *Queen Mary University of London, Centre for Digital Music, Tech. Rep. C4DM-TR-09-06*.

De Lima, Luis Gustavo Rotoly, André de Lima Salgado, and André Pimenta Freire (2015). “Evaluation of the user experience and intrinsic motivation with educational and mainstream digital games". In: *Proceedings of the Latin American Conference on Human Computer Interaction*, pp. 1–7.

De Schipper, Tessa, Lauren J Lieberman, and Brigitte Moody (2017). ""Kids like me, we go lightly on the head": Experiences of children with a visual impairment on the physical self-concept". In: *British Journal of Visual Impairment* 35.1, pp. 55–68.

Dostál, Martin (2005). "Genetic Algorithms As a Model of Musical Creativity–on Generating of a Human-Like Rhythmic Accompaniment". In: *Computing and Informatics* 24.3, pp. 321–340.

Ellis, Daniel PW (2007). "Beat tracking by dynamic programming". In: *Journal of New Music Research* 36.1, pp. 51–60.

Friberg, Johnny and Dan Gärdenfors (2004). "Audio games: new perspectives on game audio". In: *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pp. 148–154.

Fuentes, Magdalena, Brian McFee, Hélene C Crayencour, Slim Essid, and Juan P Bello (2018). "Analysis of common design choices in deep learning systems for downbeat tracking". In: *The 19th International Society for Music Information Retrieval Conference.*

Garcia, Franco Eusébio and Vânia Paula de Almeida Neris (2013). "Design guidelines for audio games". In: *Human-Computer Interaction. Applications and Services: 15th International Conference, HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part II 15.* Springer, pp. 229–238.

Gkiokas, Aggelos and Vassilis Katsouros (2017). "Convolutional Neural Networks for Real-Time Beat Tracking: A Dancing Robot Application." In: *ISMIR*, pp. 286–293.

Gluck, Aaron, Kwajo Boateng, and Julian Brinkley (2021). "Racing in the Dark: Exploring Accessible Virtual Reality by Developing a Racing Game for People who are Blind". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.* Vol. 65. 1. SAGE Publications Sage CA: Los Angeles, CA, pp. 1114–1118.

Gluck, Aaron and Julian Brinkley (2020). "Implementing 'The Enclosing Dark': A VR Auditory Adventure". In: *Journal on Technology and Persons with Disabilities* 8.2020, pp. 149–159.

Gold, Kevin and Alex Olivier (2010). "Using machine translation to convert between difficulties in rhythm games". In: *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment.* Vol. 6. 1, pp. 27–32.

Goto, Masataka (2001). "An audio-based real-time beat tracking system for music with or without drum-sounds". In: *Journal of New Music Research* 30.2, pp. 159–171.

Goto, Masataka and Yoichi Muraoka (1994). "A beat tracking system for acoustic signals of music". In: *Proceedings of the second ACM international conference on Multimedia*, pp. 365–372.

— (1999). "Real-time beat tracking for drumless audio signals: Chord change detection for musical decisions". In: *Speech Communication* 27.3-4, pp. 311–335.

Grosche, Peter and Meinard Muller (2010). "Extracting predominant local pulse information from music recordings". In: *IEEE Transactions on Audio, Speech, and Language Processing* 19.6, pp. 1688–1701.

Hendrix, Maurice, Tyrone Bellamy-Wood, Sam McKay, Victoria Bloom, and Ian Dunwell (2018). "Implementing adaptive game difficulty balancing in serious games". In: *IEEE Transactions on Games* 11.4, pp. 320–327.

Hestenes, Linda L and Deborah E Carroll (2000). "The play interactions of young children with and without disabilities: Individual and environmental influences". In: *Early childhood research quarterly* 15.2, pp. 229–246.

Heydari, Mojtaba, Frank Cwitkowitz, and Zhiyao Duan (2021). "BeatNet: CRNN and particle filtering for online joint beat downbeat and meter tracking". In: *arXiv preprint arXiv:2108.03576*.

Heydari, Mojtaba and Zhiyao Duan (2021). "Don't look back: An online beat tracking method using RNN and enhanced particle filtering". In: *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, pp. 236–240.

Hooshyar, Danial, Moslem Yousefi, and Heuiseok Lim (2019). "A systematic review of data-driven approaches in player modeling of educational games". In: *Artificial Intelligence Review* 52, pp. 1997–2017.

Hunicke, Robin (2005). "The case for dynamic difficulty adjustment in games". In: *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pp. 429–433.

IJsselsteijn, Wijnand A, Yvonne AW De Kort, and Karolien Poels (2013). "The game experience questionnaire". In.

Ji, Shulei, Jing Luo, and Xinyu Yang (2020). "A comprehensive survey on deep music generation: Multi-level representations, algorithms, evaluations, and future directions". In: *arXiv preprint arXiv:2011.06801*.

Johnson, Daniel, M John Gardner, and Ryan Perry (2018). "Validation of two game experience scales: the player experience of need satisfaction (PENS) and game experience questionnaire (GEQ)". In: *International Journal of Human-Computer Studies* 118, pp. 38–46.

Kassler, Michael (1966). “Toward musical information retrieval”. In: *Perspectives of New Music*, pp. 59–67.

Khaliq, Imran and Isabelle Dela Torre (2019). “A study on accessibility in games for the visually impaired”. In: *Proceedings of the 5th EAI international conference on smart objects and technologies for social good*, pp. 142–148.

Laser Eye Surgery Hub, author (Mar. 2022). *Visual impairment blindness global data statistics.* URL: https://www.lasereyesurgeryhub.co.uk/data/visual-impairment-blindness-data-statistics/.

Liang, Yubin, Wanxiang Li, and Kokolo Ikeda (2019). “Procedural content generation of rhythm games using deep learning methods”. In: *Entertainment Computing and Serious Games: First IFIP TC 14 Joint International Conference, ICEC-JCSG 2019, Arequipa, Peru, November 11–15, 2019, Proceedings 1.* Springer, pp. 134–145.

Lopes, Ricardo and Rafael Bidarra (2011). “Adaptivity challenges in games and simulations: a survey”. In: *IEEE Transactions on Computational Intelligence and AI in Games* 3.2, pp. 85–99.

Mann, Henry B and Donald R Whitney (1947). “On a test of whether one of two random variables is stochastically larger than the other”. In: *The annals of mathematical statistics*, pp. 50–60.

Meier, Peter, Gerhard Krump, and Meinard Müller (2021). “A Real-Time Beat Tracking System Based on Predominant Local Pulse Information”. In: *Demos and Late Breaking News of the International Society for Music Information Retrieval Conference (ISMIR). Online. https://archives. ismir. net/ismir2021/latebreaking/000021. pdf.*

Meier, Peter, Simon Schwär, Sebastian Rosenzweig, and Meinard Müller (2022). “Real-Time MIR Algorithms for Music-Reactive Game World Generation”. In: *Mensch und Computer 2022-Workshopband.*

Metatla, Oussama, Sandra Bardot, Clare Cullen, Marcos Serrano, and Christophe Jouffrais (2020). “Robots for inclusive play: Co-designing an educational game with visually impaired and sighted children”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–13.

Nacke, Lennart E, Mark N Grimshaw, and Craig A Lindley (2010). “More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game”. In: *Interacting with computers* 22.5, pp. 336–343.

Nacke, Lennart E, Sophie Stellmach, and Craig A Lindley (2011). “Electroencephalographic assessment of player experience: A pilot study in affective ludology”. In: *Simulation & Gaming* 42.5, pp. 632–655.

Nuanáin, Cárthach Ó, Herrera Boyer, Sergi Jordà Puig, et al. (2015). "Target-based rhythmic pattern generation and variation with genetic algorithms". In: *Proc. of the 12th Int. Conference on Sound and Music Computing (SMC-15); 2015 Jul 30-Aug 1; Maynooth, Ireland. Maynooth; Department of Computer Science, Maynooth University; 2015.* Maynooth University.

Oliveira, Joao Lobato, Fabien Gouyon, Luis Gustavo Martins, and Luis Paulo Reis (2010). "IBT: A Real-time Tempo and Beat Tracking System." In: *ISMIR*, pp. 291–296.

Oren, Michael A (2007). "Speed sonic across the span: building a platform audio game". In: *CHI'07 Extended Abstracts on Human Factors in Computing Systems*, pp. 2231–2236.

Pinto, António S, Sebastian Böck, Jaime S Cardoso, and Matthew EP Davies (2021). "User-Driven Fine-Tuning for Beat Tracking". In: *Electronics* 10.13, p. 1518.

Potamianos, Evangelos (2022). "Developing an inclusive multiplayer rhythm game for visually impaired children". MA thesis.

Prendinger, Helmut, Kamthorn Puntumapon, and Marconi Madruga (2014). "Extending real-time challenge balancing to multiplayer games: A study on eco-driving". In: *IEEE Transactions on Computational Intelligence and AI in Games* 8.1, pp. 27–32.

Roe, Joao (2008). "Social inclusion: meeting the socio-emotional needs of children with vision needs". In: *British Journal of Visual Impairment* 26.2, pp. 147–158.

Russell, James A (1980). "A circumplex model of affect." In: *Journal of personality and social psychology* 39.6, p. 1161.

Smeets, Ed and Anke de Boer (Sept. 2019). *Leerlingen met een visuele of auditieve beperking en/of taalontwikkelingsstoornis.* URL: `https://evaluatiepassendonderwijs.nl/onderzoek/onderzoeksplannen/thematische-casestudie-leerlingen-visuele-auditieve-beperking-enof-taalontwikkelingsstoornis/`.

Trust, Sensory (n.d.). *Guide to inclusive play.* URL: `https://www.sensorytrust.org.uk/resources/guidance/inclusive-play`.

Verver, Suzanne H, Mathijs PJ Vervloed, and Bert Steenbergen (2020). "Facilitating play and social interaction between children with visual impairments and sighted peers by means of augmented toys". In: *Journal of Developmental and Physical Disabilities* 32.1, pp. 93–111.

Vision Impairment and blindness (Oct. 2022). URL: `https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment`.

Yannakakis, Georgios N, Pieter Spronck, Daniele Loiacono, and Elisabeth André (2013). "Player modeling". In: *S. M. Lucas, M. Mateas, M. Preuss,*

P. Spronck, J. Togelius (Eds.), Artificial and computational intelligence in games, DFU 6, pp. 45–59.
Yeh, Tzu-Chun and Jyh-Shing Roger Jang (2019). "AutoRhythm: a music game with automatic hit-timing generation and percussion identification". In: *IEEE Transactions on Games* 12.3, pp. 291–301.
Yeh, Tzu-Chun, Hao-Hsun Li, and JS Roger (2014). "Automatic hit time generation for music rhythm games". In: *15th International Society for Music Information Retrieval Conference, ISMIR2014*.
Yuan, Bei and Eelke Folmer (2008). "Blind hero: enabling guitar hero for the visually impaired". In: *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*, pp. 169–176.
Yuan, Bei, Eelke Folmer, and Frederick C Harris (2011). "Game accessibility: a survey". In: *Universal Access in the information Society* 10, pp. 81–100.
Zohaib, Mohammad (2018). "Dynamic Difficulty Adjustment (DDA) in computer games: A review". In: *Advances in Human-Computer Interaction* 2018.
Zook, Alexander E and Mark O Riedl (2012). "A temporal data-driven player model for dynamic difficulty adjustment". In: *Eighth Artificial Intelligence and Interactive Digital Entertainment Conference*.

Appendix

Questionnaire Used to Answer RQ4.2

Thank you for taking part in my Master's Thesis study! This study is on developing collaborative multiplayer games for the visually impaired. Please take your time to answer each question truthfully, the survey is only 3 pages long. If you have any questions, please do not hesitate to ask!

1. Please enter your participant ID.
2. Please specify your sex.
   (a) Male
   (b) Female
   (c) Prefer Not To Say
3. Pick the option that best describes you.
   (a) Non-gamer
   (b) Casual gamer
   (c) Hardcore gamer
   (d) Game Developer (or aspiring)
   (e) Games Researcher (or aspiring)
4. What role did you play as?
   (a) Musician (i.e. playing the notes)
   (b) Conductor (i.e. guiding the musician on what notes to play)
5. Which version of the game did you find the most engaging?
   (a) Version 1
   (b) version 2
   (c) I found them equally engaging
6. On a scale of 1 (boring) to 5 (exciting):
   - Please rate your experience with the first version of the game.
   - Please rate your experience with the second version of the game.
7. Do you have any further feedback on the first version of the game?

8. Do you have any further feedback on the second version of the game?
9. Any other general feedback?

Questionnaire Used to Answer RQ3

- Please give your opinions on the following statements when playing as the **musician** (on a scale of 0 (disagree) to 4 (fully agree)):
    1. My actions had a clear immediate effect on the game.
    2. It was clear to me when I hit a note.
    3. It was clear to me when I missed a note.
    4. When I missed a note, I always knew if my action was too early or too late.
    5. It felt good to hit the note correctly.
    6. I was aware of my score throughout the course of the game.
    7. It felt like I was playing music on stage.
    8. The controls were easy to understand.

Consent Form for Children Participating in Final Experiment

In the following pages, you will find the consent form and information sheet that was sent to the families of the children attending Bartiméus school who participated in the final experiment. Note that these were translated into Dutch before sending them out.

Universiteit Utrecht

Consent form for participation in the research project

Improving the Player Experience of Collaborative Multiplayer Games for Visually Impaired Children

Please complete the form below by ticking the relevant boxes and signing on the line below. A copy of the completed form will be given to you for your own record.

- I confirm that I am 18 years of age or over.

- I confirm that the research project ***"Improving the Player Experience of Collaborative Multiplayer Games for Visually Impaired Children"*** has been explained to me. I have had the opportunity to ask questions about the project and have had these answered satisfactorily. I had enough time to consider whether I consent my child to participate.

- I consent to the material I contribute being used to generate insights for the research project "**Improving the Player Experience of Collaborative Multiplayer Games for Visually Impaired Children***".*

- I understand that my child's participation in this research is voluntary and that they may withdraw from the study at any time without providing a reason, and that if they withdraw any data already collected from me will be erased.

- I consent to allow the fully anonymized data to be used in future publications and other scholarly means of disseminating the findings from the research project.

- I understand that the data acquired will be securely stored by researchers, but that appropriately anonymized data may in future be made available to others for research purposes. I understand that the University may publish appropriately anonymized data in appropriate data repositories for verification purposes and to make it accessible to researchers and other research users.

- I agree to take part in the above research project on ***"Improving the Player Experience of Collaborative Multiplayer Games for Visually Impaired Children"***.

Name of parent/ guardian	Date	Signature

Name of researcher	Date	Signature

Universiteit Utrecht

Research Participant Information Sheet

Improving the Player Experience of Collaborative Multiplayer Games for Visually Impaired Children

June 2023

1. Introduction

This study is a continuation of a study that was conducted last year, where a game has been designed with accessibility features for Visually Impaired (VI) children. The game is a collaborative multiplayer game where 2 players will work together to play the bongos. One player is the conductor, who will inform the other player (the musician) on what note to play and when. Together they will play along to a song of their choosing.

Last year students from Bartiméus participated in a similar study and the overall feedback was that they really enjoyed the game and wanted to keep playing, furthermore, they mentioned that they would like.to see the game developed further. For this year's study, we have improved the game based on the feedback given to us by the participants from last year's study. We are asking for consent for the participant to take part in a scientific study to gain insights into the implemented changes. The study will take place at Bartiméus.

2. What is the background and purpose of this study?

Through the study will we gain further insights and a deeper understanding of the best practices for implementing accessibility features in games for VI players. The larger goal and overall purpose is to improve the quality of life of VI children by understanding how we can adapt technology, apps, and games for VI people. Furthermore, we aim to create a game that can be enjoyed by sighted and VI children equally, regardless of their abilities.

3. Who will carry out the study?

This study is carried out by Marc Ferriggi (m.ferriggi@students.uu.nl) as part of my master's thesis under the supervision of Prof. Anja Volk (A.Volk@uu.nl).

4. How will the study be carried out?

In this study, students will be paired in groups of 2. They will first be given a short tutorial on how to play the game. Each student will then play the game twice, once as a musician and once as a conductor. Following this, they will be assisted to answer a short questionnaire to gather their opinions on the game and learn about their experience. The entire experiment is expected to last around 20 minutes.

5. What will we do with your data?

No personal data will be collected. We will store the feedback anonymously.

6. What are your rights?

Participation is voluntary. We are only allowed to collect the data for our study if you consent to this. If you do not give consent for your child to participate, you do not have to take any further action. You do not need to sign anything. Nor are you required to explain why you do not want your child to participate. If you decide to give consent for your child to participate, the child can always change their mind and stop participating at any time during the study.

7. Approval of this study

This study has been allowed to proceed by the Research Institute of Information and Computing Sciences on the basis of an Ethics and Privacy Quick Scan. If you have a complaint about the way this study is carried out, please send an email to: ics-ethics@uu.nl. If you have any complaints or questions about the processing of personal data, please send an email to the Faculty of Sciences Privacy Officer: privacy-beta@uu.nl. The Privacy Officer will also be able to assist you in exercising the rights you have under the GDPR. For details of our legal basis for using personal data and the rights you have over your data please see the University's privacy information at www.uu.nl/en/organisation/privacy.

8. More information about this study?

If you have any questions or concerns about this research please contact Marc Ferriggi at m.ferriggi@students.uu.nl or my supervisor Anja Volk at A.Volk@uu.nl.

9. Appendices:

Along with this information sheet, you will find an informed consent form.

Expert Interview Outline

Protocol

Materials

This document, an audio recording device, and a notepad to take notes.

Semi-Structured Interview

Opening the interview

Build rapport with the participant so that they feel comfortable being interviewed:

- Hi, what is your name? And how are you?
- Are you ready to start?

Introducing the Game

Thank you for joining my study, I appreciate it a lot. For my master's thesis, I designed a collaborative multiplayer game with accessibility features for the Visually Impaired. The game you will play is a prototype of a rhythm game where two players must work together to play the bongos on stage in front of a live audience. One player is the coordinator who will guide the other player on what notes to play, and one player is the musician who must play these notes at the correct time. You will now play a minimum of 2 rounds, one time as a coordinator, and one time as a musician. The game has 15 songs, so please also take some time to choose any song you wish. [*If it's a one-on-one expert interview we should let them play for as many rounds as they want*]. The game features both audio and visuals, however, it is designed in a way that it is also possible to play it without the visual features and still have the same experience, [*feel free to try playing the game in different ways (if they are not visually impaired)*]. The experiment will take about 30 minutes [*explain the steps*].

Here is the information sheet and the informed consent form. Take your time to read it carefully. After that, you can decide if you want to sign it or not.

[*if the participant does not sign*]
Thank you anyways for your time. Here are the sheets with my contact information in case you change your mind, it is no problem at all to do the experiment at a later date. Have a nice day.

Before playing the game *simulate* the tutorial by explaining the different feedback sounds and playing them.

After Playing the Game

An important aspect of making games accessible means that they should be able to be played and enjoyed by everyone. The goal of the experiment is to study the effect of 3 new features added to the game, those being an algorithm to automatically generate a level given a song, improved feedback, and accessibility features, and the implementation of an algorithm to dynamically adjust the difficulty of the game as the player is playing to keep it engaging. The target group for the game is children and students. Because you are [*student/ work closely with VI children/ an expert in the field/ visually impaired*], I think your answers are very valuable for the research.

Beginning the Interview

Now that you have played the game, let us begin the interview. The interview takes about $10 - 15$ minutes. Then the recording starts now [*start recording*].

The Interview

I will start with a question about your prior experience playing games.

- Do you play computer/video games?
  - If not, have you ever played games before?
  - If yes, have you ever played any rhythm games?

Now to move on to questions about the game.

- What was your first impression of the game?

Thank you for your answer. My next question is:

- On a scale from $1 - 10$, how much did you feel like you understood what was happening in the game?
- Can you elaborate on your answer?

Thanks!

- What did you think of the difficulty of the game?

That's clear. Let's dive deeper into the different roles.

- Which role was your favourite?
  - Why?
- Do you have any ideas on how we can improve the [*least favourite*] role?

Thank you, this is really helpful!

- About the musician role,
  - Did you feel like you were adding to the song, or did the notes feel random?
  - Were you aware when you missed a note, played it a bit too early/ too late, and played it on time?
- About the coordinator role,
  - Was it clear what was happening in the game?
  - Did you feel like you were guiding the musician well?

To move on to another aspect:

- What did you think of the music in the game?
  - Did you find any songs you liked?
  - Did you feel like you also wanted to try different songs?
  - Is there a style of music you wish you saw in the game?

Alright. I'd like to zoom out, so we're talking about the game in general.

- What did you think about the collaboration part of the game?
  - Do you feel like playing this game with a friend will improve your relationship with them?
  - What are your thoughts on playing it with a complete stranger?

Good. Then my final question is:

- Do you have any other feedback?

Closing the Interview

Wrap up the interview:

- The interview is coming to an end. So we will start to wrap it up. Is there anything else you would like to add?

- Thank you again for participating in the experiment. I appreciate it very much. Your data will be anonymized and used as input for the research of *Collaborative Multiplayer Games for the Visually Impaired.* Feel free to contact me anytime if you have any questions, remarks, or anything else. Have a nice day [*end recording*].