The University of York Department of Computer Science

Submitted in part fulfilment for the degree of MEng.

What’s it like to be a bat gamer?

Huw Talliss

1st May 2018

Supervisor: Dr. Paul Cairns

Number of words = 15, 734, as counted by `wc -w`.
This includes the body of the report, but not the appendices.

Abstract

Audio is widely used in video games as a way to improve them, but it is not known what effect audio has on the experience that games provide in terms of immersion. It is also unclear whether when deprived of a visual interface, players are still able to become immersed. I.e. can audio-only games be immersive?

In this report a study is performed to evaluate what effect audio has on immersion and whether players can become immersed in an audio-only game, through the creation of a bespoke game with three modes: 'audio-only', 'video-only', and 'both'. An analysis of the various types of games for which audio is an important modality is also presented.

The results of th study indicate that players are still able to become immersed in games when deprived of a visual interface, and that audio provides a small increase in immersion over a video-only experience. Particularly in the challenge and emotional involvement immersive factors.

These results correspond with the common assumption that audio improves video game experience and demonstrates that immersive, audio-only games can be indeed created.

Contents

1 Introduction **11**

2 Literature Review **13**
2.1 Audio as an Important Modality 13
2.1.1 Rhythm as an Important Modality 14
2.1.2 Pitch as an Important Modality 16
2.1.3 Spatiality as an Important Modality 17
2.2 Topologies for analysing game audio 21
2.2.1 Auditory Icons and Earcons 21
2.2.2 Categorisation of audio in audio-only games 22
2.2.3 Listening Modes 23
2.2.4 Diegetic Sounds 24
2.2.5 Interactivity 26
2.2.6 IEZA Framework 27
2.3 Immersion in video games 31
2.4 Summary 32

3 Creating the Game **34**
3.1 Requirements 34
3.2 Prototyping 35
3.3 Bounce Game Development 38
3.3.1 Timing & Bytecode 38
3.3.2 Design 41
3.3.3 Tutorial 43
3.3.4 Logging 46

4 The Study **48**
4.1 Hypothesis 48
4.2 Design 48
4.3 Participants 48
4.4 Materials 49
4.5 Procedure 50
4.6 Pilot 50

4.7 Results . . . 51
4.7.1 Total Immersion . . . 52
4.7.2 Challenge . . . 53
4.7.3 Control . . . 53
4.7.4 Cognitive Involvement . . . 54
4.7.5 Emotional Involvement . . . 55
4.7.6 Real World Dissociation . . . 56
4.7.7 Player Performance . . . 57
4.8 Discussion . . . 58
4.9 Evaluation of Bounce Game . . . 60
4.10 Experimental Validity and Limitations . . . 63
4.10.1 Internal Validity . . . 63
4.10.2 External Validity . . . 63
4.10.3 Ecological Validity . . . 63

5 Conclusions 65
5.1 Contributions to Literature . . . 65
5.2 Further Work . . . 65
5.3 Summary . . . 66

6 References 67
6.1 Bibliography . . . 67
6.2 Ludography . . . 70

A Bounce Game 72
A.1 Bytecode . . . 72
A.1.1 Regular Gameplay . . . 72
A.1.2 Tutorial . . . 73
A.2 Tutorial text . . . 73
A.2.1 Audio-Only . . . 73
A.2.2 Video-Only . . . 74
A.2.3 Both . . . 74
A.3 Logging . . . 75

B The Study 77
B.1 Study Documents . . . 78
B.1.1 Audio-Only Information Sheet . . . 78
B.1.2 Video-Only Information Sheet . . . 80
B.1.3 Both Information Sheet . . . 82
B.1.4 Informed Consent Form . . . 84

B.1.5 Experimenter's Notes 85
B.2 Study Questionnaires 86
B.2.1 Demographic Survey 86
B.2.2 IEQ Questionnaire 88

C Statement of Ethics **90**

List of Figures

2.1 In Crypt of the Necrodancer the floor flashes like a dancefloor to show the timing of the beat. 14
2.2 In Mother 3, the number of spinning music notes signifies the length of the current combo. 15
2.3 A diagram classifying games for which rhythm is an important modality in terms of the link between rhythm and expected player input. 17
2.4 For this puzzle in The Witness, the player heard two high pitches, a low pitch and then a middling pitch. 18
2.5 A diagram classifying games for which spatialised audio is an important modality in terms of how audio is spatialised and how these spatialised sounds are played. 20
2.6 A diagram of the IEZA Framework [33] 28
2.7 Figure to show increasing/decreasing optimisation and dynamisation with respect to the IEZA diagram [1]. 30

3.1 Screenshots of the three prototypes 36
3.2 An example line of the bytecode csv file. 38
3.3 Initial command execution regularity 39
3.4 Final command execution regularity 39
3.5 Rhythms for the three Balls 40
3.6 Screenshots of the three ball types to show the colour differences. 41
3.7 High-level game state diagram 42
3.8 Civilisation IV uses a great deal of text split into many text boxes in its tutorials due its many complex mechanics. . . 44
3.9 One of the few pieces of text in Hyper Light Drifter, teaching the player they can shoot switches through the use of a gameplay situation. 44
3.10 The mode select screen. Which allowed the tutorial to be disabled and the game mode to be selected. 45
3.11 Screenshot of the game's tutorial. Tutorial text was displayed in the textbox. 47

4.1 Boxplot of Total Immersion scores across the three experimental conditions (audio-only, video-only, and both). . . . 51
4.2 Boxplot of Challenge scores across the three experimental conditions. 52
4.3 Boxplot of Control scores across the three experimental conditions. 53
4.4 Boxplot of Cognitive Involvement scores across the three experimental conditions. 54
4.5 Boxplot of Emotional Involvement scores across the three experimental conditions. 55
4.6 Box plot of Real World Dissociation scores across the three experimental conditions. 56
4.7 Box plot of Player Performance across the three experimental conditions. 57

List of Tables

2.1 Breakdown of diegetic sounds in games [29]. 25

4.1 Descriptive statistics of IEQ scores across the experimental conditions. 51
4.2 Descriptive statistics of Challenge scores across the experimental conditions. 52
4.3 Descriptive statistics of Control scores across the experimental conditions. 53
4.4 Descriptive statistics of Cognitive Involvement scores across the experimental conditions. 54
4.5 Descriptive statistics of Emotional Involvement scores across the experimental conditions. 55
4.6 Descriptive statistics of Real World Dissociation scores across the experimental conditions. 56
4.7 Descriptive statistics of Player Performance across the experimental conditions. 57

A.1 List of bytecode commands for regular gameplay 72
A.2 List of bytecode commands for the tutorial 73
A.3 Example log file including the three practice sections in the tutorial followed by the main gameplay. 76

1 Introduction

Audio has been used in video games for a long time, becoming an extremely important part of modern games and game design, yet research into audio in games is still quite new. There is a general assumption that the use of well-designed audio in games improves them and makes them more immersive, yet there is a lack of research and studies to investigate this assumption that audio has an effect on immersion. One of the few examples of work done in this area is by Huiberts [1], but while this does address the effect of audio on immersion (largely analysed in terms of the IEZA framework [2]), findings were limited to a user survey [3] and no experimental study was performed. There are various questions related to the use of audio in games that are yet to be fully answered, such as 'can audio actually increase immersion in games?', 'how can audio be used to improve game experience?', and 'how can we best create fun and immersive games in which audio is a key aspect of gameplay?'.

Additional questions stem from the fact that the large majority of video games extensively use visual components, which means that those with visual impairments cannot enjoy these games fully. In order to make accessible games for those with visual impairments it is important to first know whether it is possible to make an effective and enjoyable audio-only game. Answers to the questions 'What is the experience of those playing an audio-only game?' and 'can players become immersed in audio-only games?' could be very beneficial for those people that cannot currently fully enjoy many video games due to visual impairments. However little formal research has been done into this area of study.

In order to attempt to perform a study to investigate some of these questions, we require a game in which we have tight control over the audio and video outputs. This game must also work equally well audio-only and video-only. We must therefore create a bespoke game to fulfil these requirements which we can then use to perform a study looking into the immersive experience of playing the game in these different modalities.

As a result of this study we find that audio does indeed have an effect on immersion, causing a slight increase in immersion over a video-only experience. Additionally, we find that it is indeed possible for players to become immersed in a game when deprived of a visual interface.

In addition to this study, an analysis of the types of games in which audio is an important modality is performed, in particular how audio is used in these games. Three principle areas of audio that are used in such games are found: rhythm, pitch and spatiality. Additionally, two categorisation frameworks and diagrams that can be used in analysing such games are defined.

In terms of structure for this report, the following chapter contains a review of literature that supports and relates to the study of audio and immersion in games, as well as various topologies for analysing game audio. Chapters 3 and 4 detail the creation of a game and its subsequent use in a study into the effect of audio on immersion. The final chapter suggests a body of future work to be done and concludes the report.

2 Literature Review

The field of research into the effect of audio in video games is relatively young, and very few studies have been performed to examine this area. A large amount of prior work is concentrated on devising topologies and frameworks for analysing and categorising game audio, rather than on what effect audio has in games. While formal research is somewhat lacking, the large amount and wide variety of different video games in existence means that looking at how they differ in their use of audio can give good insight into the effects audio can have.

2.1 Audio as an Important Modality

Most games use audio in some way or another, often to enhance the experience and improve immersion [4, p. 257] (playing a game with the sound on and then off provides two very different game experiences). The quality of a game's audio can improve the player's perception of other parts of the game, such as the graphics [5, p. 351]. However there are some games where audio is not used simply as a kind of decoration (as a way to enhance the core experience), but as an integral part of the game's design and the experience it affords.

Audio in these games is described as being an important modality, a key channel of communication through which the player interacts with the game. There are many genres of games that use audio as an important modality and they often use audio in many different ways. Games such as bit Generations: Soundvoyager [6], and Papa Sangre [7] are designed specifically to be playable audio-only, so audio is the most important aspect of the game. These games that have audio as their forefront gameplay mechanic are quite uncommon, much more common is for audio to be an important, but supplementary channel of communication.

Competitive team games such as Dota 2 [8] and Overwatch [9] are examples of this latter kind of game, where audio is not the primary modality but still extremely important, especially at high-level play. The audio in these games is not just decorative but a very important channel through which information about the state of the game is continually communicated to players.

These games for which audio is an important modality can be categorised in three main ways: Rhythm, Pitch, and Spatiality, by looking at which aspects of the audio are most important when they are communicating to the player as well as what aspect of audio their gameplay is based around.

2.1.1 Rhythm as an Important Modality

Rhythm-based games are often what we first think of when considering games where audio is an important modality. Games such as the Guitar Hero [10] series in which the core gameplay involves playing music, where the player is required to enter input with specific timings, corresponding with the rhythms of the song being played. Rhythm in these games is used to communicate to the player about the input that is required. This input can be split primarily into the categories of input type and input timing. For example, in Crypt of the Necrodancer [11] for player input to be accepted, it must be entered in time with the beat of the music. The player is free to choose what actions they want to perform (their input type) but their input timing is specified by the game's audio, the beat of the music.

Figure 2.1: In Crypt of the Necrodancer the floor flashes like a dance-floor to show the timing of the beat.

Another example is Mother 3, a role-playing game with turn-based combat. Every battle has music playing in the background which is different based on the type of enemies the player is battling against. When attacking, the player can can tap the attack button in time with the music to chain their attack into a combo in order to deal additional damage to the enemy. Initially the rhythms of the background music are simple and relatively slow 4/4 rhythms. But as the game progresses, songs with more complex rhythms and time signatures are added which often include additional beats designed to throw off the player. Later songs use 15/8 and 29/16 time signatures [12] at high tempos, increasing the difficulty of the later battles in the game. Similarly to Guitar Hero, rhythm in Mother 3 is used to specify the input timing but not the input type (It is always the attack button that needs to be pressed for the rhythm-based combo).

Figure 2.2: In Mother 3, the number of spinning music notes signifies the length of the current combo.

The sub-games in Rhythm Heaven: Fever [13] have a strong link between not only rhythm and required input timing, but also between rhythm and input type. Players are taught to respond to specific rhythmic cues with responses of specific input types and timings; different rhythmic cues signal different input timings and types. The various input types in Rhythm Heaven: Fever are: pressing the 'A' button, pressing the 'B' button or pressing the 'A' and 'B' buttons together at the same time (sometimes buttons must be held down for a short time as well). An example of rhythms specifying input types and timings in this game is the Samurai Slice sub-game where the rhythmic cue of a demon spawning indicates to the player that they must press the 'A' button (the input type) on the beat after the cue finishes playing (the input timing). Another rhythmic cue occurs when multiple demons spawn and indicates that the player must press and hold both the 'A' and 'B' buttons at the same time on the beat after the cue finishes, releasing the buttons after one beat.

There is a strong relationship between the rhythmic cues and player input in the Rhythm Heaven games, rhythm *specifies* player input. Whenever an input is required of the player, one of the rhythmic cues will always play first. In contrast to this, in Guitar Hero [10] the relationship between rhythm and player input is much weaker. Required input is much more tied to the visual position of the on-screen note markers than the music. Required input timings can also change based on difficulty settings rather than exclusively due to changes in the music. For this case, rhythm only *implies* the player input, it doesn't specify it.

A diagram can be constructed in order to classify games based on the link they have between rhythm and expected player input (figure 2.3).

2.1.2 Pitch as an Important Modality

The group of games which use pitch as an important modality is much smaller than those which use rhythm. (It is possible that the reduction in the number of pitch-based games compared to rhythm games may be to do with a larger variance in the ability of people with regards to pitch than with rhythm). The Witness [14], a puzzle game with a variety of puzzle types has a series of puzzles based on pitch (figure 2.4), where the player must discern a simple bird call made up from two or three different pitches, from other distracting sounds and then enter the call that they heard onto an in-game panel. Compared to the other puzzles present in The Witness which are much more visual-based, the perceived difficulty of this type of puzzle seems to have a much larger variance between different players judging by online discussion [15], [16]. Some players found it very easy to discern between the relative pitches and solve the puzzle, but others found it hard, verging on impossible. There is clearly some variation in the pitch-ability of players here, but it is not

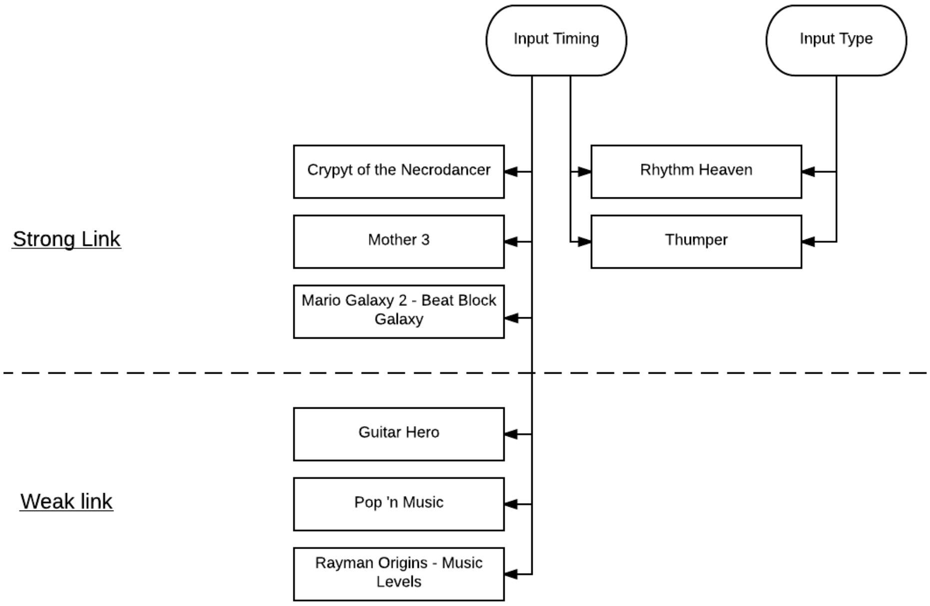


Figure 2.3: A diagram classifying games for which rhythm is an important modality in terms of the link between rhythm and expected player input.

obvious how representative this is of the wider population.

Another Puzzle-Style game String Theory [17], presents multiple guitar strings which each play a note that makes up a chord that matches with the backing music, however one of the notes is dissonant. The player must correctly identify and select the dissonant string to remove using exclusively audio information. This game heavily relies on the musical and pitch-recognition ability of the player and acts more as a training game for these skills rather than a more mainstream game.

2.1.3 Spatiality as an Important Modality

Spatiality is an aspect of audio that is used a great deal in games. Audio in nearly all major new games is spatialised, usually using stereo audio for direction, and volume for proximity. Sounds are usually attached to an in-game source, (an animal or an object for example) and will be spatialised based on the position of the sound source relative to the player [4, p. 148] (UI sounds and background music are usually not spatialised). Spatiality is most regularly used to improve the game by

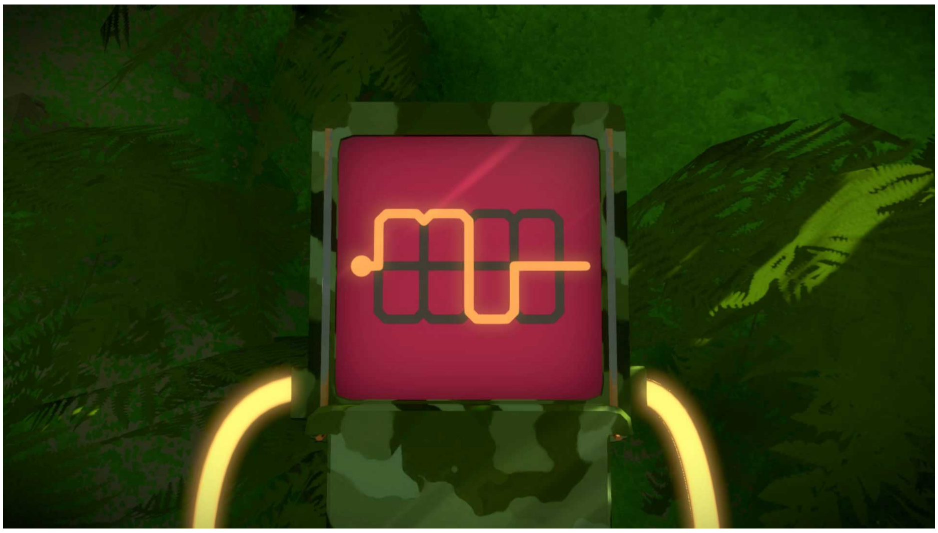


Figure 2.4: For this puzzle in The Witness, the player heard two high pitches, a low pitch and then a middling pitch.

adding depth and helping to "sell" the game world by making the audio behave in a realistic way, thus improving player experience. The reasons to use spatiality are clear: if the player saw a cannon being fired in the distance to their right yet heard the sound of the cannon firing at very high volume centred on their character that would be very jarring and confusing for the player, going against their held-expectations and pulling them out of the experience. For 3D game worlds it is a de-facto standard to spatialise audio, mimicking the real world to allow the player to use their pre-existing knowledge of sound to better understand the virtual game world. Indeed, a great deal of work is put into modern games to make the audio behave as realistically as possible through techniques such as calculating occlusion, sound propagation, reverberation, and filtering [4, p. 89].

Spatialising audio also allows players to be given more information about the state of the game world other than just the section of the world they can currently see. This is particularly important for competitive multiplayer games where having as much information as possible about the current game-state (such as enemy player positioning) is extremely advantageous.

In Overwatch [9] (a team-based first-person shooter), a complex system is used to help convey information to the player using spatialised audio.

Each player-character makes noise when they perform actions such as walking or firing, and the loudness of these sounds is adjusted based on their importance to the player. For example, the footsteps of an enemy that can do a high amount of damage that is directly behind you will be mixed much louder than those of a friendly character that is in front of you. Similarly, player-characters have special abilities which, when used cause a very loud voice sting to play, that can be heard from anywhere in the game to ensure that all players know that this ability has been used. Loudness of a sound is not only based on its proximity to the player in Overwatch, but also on its importance to the player [18].

Find The Invisible Cow [19] spatialises its audio differently to most games, exclusively using volume to indicate proximity. In the game, you move the cursor around the screen while attempting to find the location of an invisible cow. You are guided by a regular audio cue (the spoken word "cow") that increases in volume as the cursor is moved closer to the location of the invisible cow. Not only does the volume of the sound increase, the inflexion of the word also changes; transitioning from regular speech to shouting as the cursor approaches the cow's location. This game gives a good example of what can be achieved through very simple spatialisation of audio. On this 2D plane, volume is the only variable needed in order to be able to navigate the game world towards the goal.

Extending this concept of a game entirely built around spatialised audio into three dimensions, Papa Sangre [7] is a horror-themed game that describes 3D environments for the player to navigate purely using spatialised audio. The player can explore these environments by rotating their avatar in-place and by walking forwards, and must collect musical notes and proceed to the exit whilst avoiding enemies. To account for this additional dimension (compared to Find The Invisible Cow), in addition to the use of volume, audio panning is also used in spatialising the audio. Papa Sangre is a great example of how much information audio -especially spatialised audio- can give the player, allowing them to move around a relatively complex 3D environment with sound alone.

With the relative ubiquity of spatialised audio in games, it becomes hard to define a categorisation system for games in which spatialised audio is an important modality. Figure 2.5 shows a diagram that attempts to classify these games in terms of how the audio is spatialised, along with the way these spatialised sounds are played in the game, either played constantly, or in response to in-game events/actions.

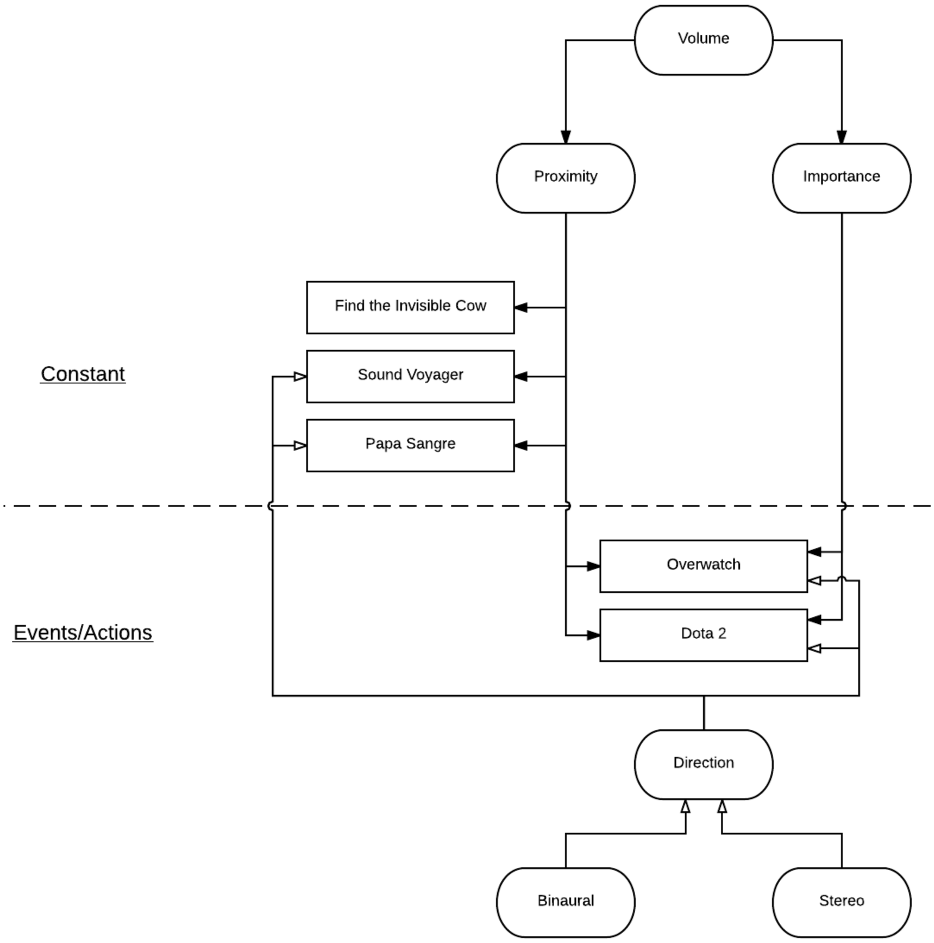


Figure 2.5: A diagram classifying games for which spatialised audio is an important modality in terms of how audio is spatialised and how these spatialised sounds are played.

Constant audio and Event/Action-based audio A standard practice in games is to tie sound to events/actions taking place. For example, in a first person shooter game such as Overwatch [9] or Call of Duty [20], when a gun is fired, a gunshot sound will be played. This is equivalent to how sounds are caused in the real world and helps to give game worlds more realism. If sounds behave as players expect them to based on prior knowledge of the real world, it can help to make the game world feel more real. Any deviations from player expectations for audio would certainly pull the player out of the experience, this is a similar concept to the idea in films that if the audience does not notice the audio in a film, then the sound department have done a good job. It is common that audio exists to complement the rest of the experience, which is often done by staying out of the way and behaving believably.

Some games will utilise constant audio with their spatialised audio (e.g. a whine of a computer fan). This is often used in soundscapes [21], building up a rich ambient texture of constant noises to improve world-depth, but can also be used to guide the player. In audio-only games like Papa Sangre [7], sound is constantly emanating from the location the player needs to move towards. This gives a constant reference point to the player as they explore the space they are in.

2.2 Topologies for analysing game audio

A sizeable amount of work has been performed into creating classification systems for describing how audio is used in games and the type of audio that is used, with many different approaches taken. This section details some of the prior work performed and a number of these approaches.

2.2.1 Auditory Icons and Earcons

Games use a mix of auditory icons and earcons for their sound effects. Auditory icons are defined as "caricatures of naturally occurring sounds" [22], they take advantage of the user's prior knowledge and natural auditory associations between sounds and their results in order to inform the user. The term auditory icon was first proposed by Gaver [22] who defined multiple categories of auditory icon:

Nomic auditory icons (dubbed causal auditory icons by Grimshaw and Schott [23] due to "nomic" having another meaning in games) are strongly tied to the event/action they represent, a direct representation. In games where objects/actions exist in the real-world, real recordings will be used for their auditory icons [23] (e.g. racing games using recordings of the engines of the real cars). Where icons are needed for fictional objects/events, they are usually constructed from combining multiple real recordings to create a sound that would seem a plausible noise for the fictional object/action to produce (e.g. a fireball spell sound may be built up of a mix of real recordings of fire, impacts sounds and whooshing/wind sounds). This kind of sound design is commonplace in games as well as films.

Symbolic auditory icons have no innate connection to the event/action that they are associated with, relying on social conventions for their meanings. Examples given by Gaver [22] are of "telephone bells and sirens.

Earcons are structured sounds that differ from auditory icons in that they are "generally synthesised tones or sound patterns and have no direct relationship to the event they are used with" [24]. As such they use learnt associations, reinforced through their repetition. An example earcon is the coin noise from the original Super Mario Bros. [25] which is a simple synthesised sound, constructed from two tones which has no direct association with real coins. Earcons were used extensively in early video games largely due to the limited audio capabilities of the hardware at the time. But as the audio capabilities of game hardware have increased, auditory icons, have become more popular than earcons in most cases, though earcons are still often used (particularly for UI audio, and in more abstract games).

2.2.2 Categorisation of audio in audio-only games

One framework for classifying audio is to categorise by how the sound assets themselves are used in games.

Whilst creating an audio-only game in which the player explores a virtual 3D environment, Friberg and Gärdenfors [26] found it useful to categorise the different types of sounds in their games "to emphasise the differences between the various auditory messages". They defined five categories:

- Avatar sounds - The effect of avatar activity (walking, attacking etc.)
  - Positioned at the centre of the stereo space and directly tied to player input
- Object sounds - Indicate the presence of objects (can be brief or continuous) dependent on the object.
- Character Sounds - Generated by non-player characters

- Ornamental Sounds - Convey non-gameplay information but enrich the game (ambient music etc.)
- Instructions - Usually speech for providing the player with information

There are a few issues with this framework, firstly that it doesn't consider sound sources such as animals (it is unclear as to whether they would fit under Character or Object sounds), and some of the distinctions between categories seem rather arbitrary. The framework also does not take into account the different functions that sounds can have in a game (though admittedly this was not the aim of the classification system). Whilst these categories may have been useful for the authors own purposes and can give a good baseline framework, it is far from a generalisable to all games.

2.2.3 Listening Modes

The field of film research has provided some very useful terms and concepts about audio that are quite also commonly used to analyse audio in the context of games. Chion [27] outlined three different ways of listening, three modes based on how and why we listen in the context of film sound:

Causal Listening is the most common mode and is the type of listening we use in order to gather information about the cause or source of a sound. We can use causal listening to identify a source we cannot see or to gain additional information about a source that we *can* see. For example, tapping a bottle can provide information about how full it is. Chion notes that causal listening is the most easily influenced and imprecise mode, and also stated that there are multiple levels of specificity that causal listening can give us as to the cause of a sound. It can be very precise in some cases (we can recognise a person's voice or a sound made by a particular object), but this isn't usually the case. More often, the information we gain from causal listening is general. One level of specificity is that we can recognise the cause of a sound without knowing any information about the cause, such as a radio presenter's voice. We recognise their voice and vocal qualities but not their face. Even more generally, we can often simply recognise the general category of source such as "mechanical" or "natural" etc. Most generally, we can often infer the "causal history" of a sound without identifying the source, by understanding the distance, acceleration, weight etc. that the sound implies.

Semantic Listening is any listening where the aim is to understand an auditory code or interpret a message. This could be spoken language or another type of code such as Morse code. It is noted that compared to causal listening, the exact acoustic properties of the sound are less important as long as the sound can be interpreted into one of the constructs of the code (such as a phoneme in natural languages).

Reduced listening was defined by Schaeffer [28] as listening that focuses on the the qualities of the sound itself rather than its cause or meaning. When we use reduced listening, we only concentrate on the sound, being unconcerned with its source, the sound *is* the reason to listen. In games this usually applies to the background music.

It is to be noted that we often listen using multiple modes at once. We might listen to what a person said but also how they said it, which can give additional meaning.

2.2.4 Diegetic Sounds

A particularly important concept commonly used in classifying game audio that was taken from film research is that of diegetic sound. The term "diegetic" is used in films to refer to elements from the fictitious world of the story. Diegetic sound is therefore sound that is part of the fictitious world, part of the diegesis. This same term is very similar when applied to games, a diegetic sound in games is a sound that is "real in the game world" [29] such as the sound of the player character firing a gun. Non-diegetic sounds therefore are sounds that are not real/part of the game world. This usually includes background music and UI sounds. Ekman [29] presents a framework for understanding game sounds using a semiotic viewpoint on sound where a "signal is the sound itself, as heard by the player" and a "referent is the thing that is being told by the sound". Ekman extends these terms to include diegesis.

A diegetic signal is a sound that is real and exists within the game world, whereas non-diegetic signals are sounds that have no in-game source or aren't real in the game world.

A diegetic referent is usually an event that occurs in the game but can also be something ongoing such as the presence of a person/object. The important thing is that the event or information is real in the diegesis. Non-diegetic referents, again are events/information not real in the game world such as game engine information.

Ekman constructs a table (2.1) using these terms as the distinctions. The use of this table also defines four other terms:

	Diegetic signal	**Non-diegetic signal**
Diegetic referent	Diegetic sounds	Symbolic sounds
Non-diegetic referent	Masking sounds	Non-diegetic sounds

Table 2.1: Breakdown of diegetic sounds in games [29].

Diegetic sounds have both a diegetic signal and a diegetic referent. The sounds are real in the game world and they signify information that is also real in the game world. These sounds have real effects in the game world and are not just decorations, e.g. footstep sounds should be audible to in-game characters as well as the player for them to be diegetic sounds.

Symbolic sounds relate to events/information that are real in the game-world but the sounds themselves are non-diegetic signals. An example of this is the narrator from Bastion [30] who narrates events and information that's real in the game world as they happen, but the actual sound signals of the narrator's voice are played non-diegetically. They are not real in the game-world as characters in the game world cannot hear the narrator (it appears).

Masking sounds are where the signal is diegetic but the referent is not. An example of this in games might be of a monster emitting a growl when it spawns into the game. This event is not (usually) part of the game world (unless it is the game's diegesis that monsters spontaneously pop into existence, growling as they do so- a worrying diegesis indeed) but is part of the game engine and game design that exists to let the player know that the monster has spawned in a place they can't see. Ekman [29] argues that all diegetic sounds in games can technically be viewed as masking sounds as they are simply constructs covering up the functionality of the game engine, they aren't part of "real" events (though generally we can ignore this distinction, else the entire concept of the diegesis in games is useless). Different events in games are more diegetic than others however, and a line can be drawn separating diegetic and non-diegetic events.

In non-diegetic sounds, neither signal nor referent are diegetic. Common examples are of user interfaces and backing music. In UIs, both the cause of sounds and the sounds themselves are usually non-diegetic, they are separate from the game-world and just decorations for the method of interaction for some of the games' systems. Backing music, like in most films is usually non-diegetic, we don't expect to find an orchestra just around the corner from the player, playing the music we hear, it is accepted as being outside the game world.

This framework presented by Ekman analyses game audio in terms of what the meaning of sounds are to the player and is a useful approach that mixes together the concept of diegesis with into a framework that works well in describing the types of audio in games.

Grimshaw and Schott [23] further refined the concept of diegetic sounds through the lens of the first-person shooter genre of games. They broke down diegetic sounds into two subcategories: ideodiegetic, and telediegetic.

Ideodiegetic sounds are sounds that the player hears. This category is further broken down into: kinediegetic sounds (sounds that are triggered by the actions of the player), and exodiegetic sounds (all other ideodiegetic sounds).

Telediegetic sounds are defined as sounds that are produced by other players that have significance to the player.

This framework is a good refinement of Ekman's [29], adding provisions for analysing games which have multiple players (something that is increasingly the case in today's games). It does not however, offer any further refinement to non-diegetic sounds, which are ignored, as part of the game's interface.

2.2.5 Interactivity

The frameworks presented so far have a weakness in that they do not particularly take into account interactivity - the key element that distinguishes video games from other media. These prior frameworks largely treat the player as just another thing that exists in the game world.

The framework presented by Collins et al. [31] focuses on participation and non-linearity, audio is not treated as static content but reactionary, and the player is in general largely in control of the game's audio through their actions and decisions. Collins et al. define three different types of audio:

- Adaptive audio - sounds that react to game events, occur due to something within the game (e.g. In Super Mario Kart [32] game where the music speeds up for the final lap)

- Interactive audio - sounds that respond to the player directly, in reaction to gameplay (e.g. the player presses the jump button, causing a jump sound to play).
- Linear audio - sounds that are not adaptive or interactive (e.g. sounds that play during pre-set cutscenes/scripted events)

This framework meshes well with prior concepts of diegetic/non-diegetic audio, allowing for combinations of the two sets of terms such as interactive non-diegetic audio (non-diegetic audio which the player controls directly), an example of which is a player walking into a new area of the world, causing the background music to change. The player has control of this non-diegetic audio and can repeatedly move between these two areas, causing the music to keep swapping.

2.2.6 IEZA Framework

The IEZA framework [2] is perhaps the most well rounded of the frameworks developed to date, consolidating a large body of the earlier frameworks into one.

Huiberts [1] specifies four different types of game audio occurrences:

- Audio during interactive gameplay
- Audio during other interactive moments (e.g. pause menus - there is interaction but not gameplay)
- Audio in non-interactive moments when the game is active (cutscenes etc.)
- Audio outside of the context of the game (e.g. adverts, soundtrack releases etc.)

Huiberts [1] comments that "for immersion, the actual interactive gameplay is most important, as the player is able to fully concentrate on the experience by participating. In the non-interactive parts of the game, the player is mostly a spectator and is not actively participating with the game itself".

The IEZA framework was designed as a topology to describe the first occurrence of audio that Huiberts defined - audio during interactive gameplay.

The framework is 2-dimensional with the dimensions being diegesis and interdependence. The poles of the diegetic dimension are diegetic and non-diegetic as defined in previous frameworks. The interdependence dimension also has two poles, defined thusly:

- Activity - communicates events occurring in the game environment
  - Usually directly reacts to player actions
  - Used to communicate what is happening and what can be interacted with
- Setting - provides a background/context for the activity
  - Doesn't generally respond directly to player actions
  - Used to communicate information about the surroundings (feel, mood, atmosphere etc.)

As can be seen in figure 2.6, this 2D view results in four intersecting areas, dubbed domains.

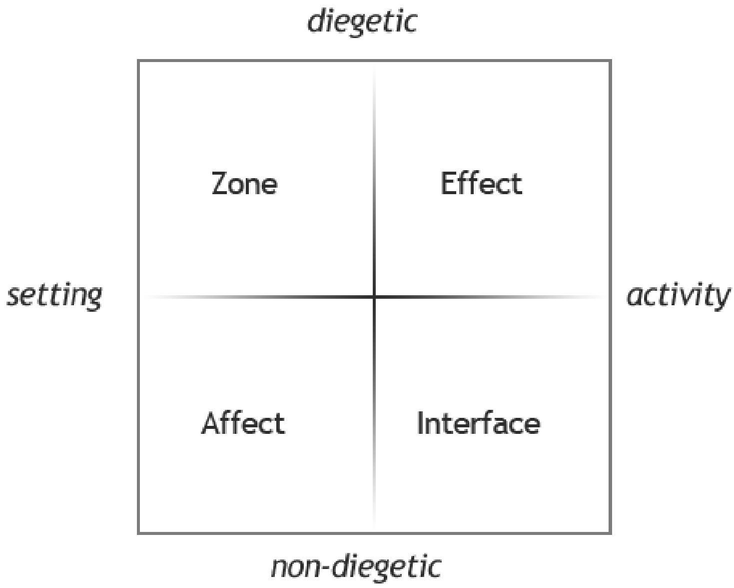


Figure 2.6: A diagram of the IEZA Framework [33]

In the effect domain, sounds are perceived as produced by (or attributed to) sources that exist within the game world (either on or off screen).

Effect-domain sounds are used regularly in games such as The Witcher 3: Wild Hunt [34] with rich virtual worlds but are also used in more "simple" games such as Tetris [35]. These sounds are "often designed to react to the player in a way that refers to sounds in the real world, and are often dynamically processed using techniques such as real-time volume changes, panning, filtering and acoustics." [1]

The zone domain contains sound sources that clearly originate from the diegesis and which are linked to the game's virtual environment. These sounds are used to give an auditory setting for the game world/environment, often by the use of an ambient/background layer. Examples of this are weather sounds, city sounds or general ambience. The Portal series in particular make heavy use of soundscapes [21] to provide ambience and set the mood of the game [36], [37]. The zone domain differs from effect largely in that these sounds are generally perceived as a single layer of sound instead of from separate sound sources.

The interface domain contains sounds from sound sources that exist outside the diegesis, showing activity in the non-diegetic game environment (e.g. menus and huds). This domain also includes representations abstract sources of activity (such as sounds that indicate that the desired action is not possible).

The affect domain is sound linked to the non-diegetic setting, usually background music (the affect domain is equivalent to non-diegetic sounds from Ekman [29]).

In terms of active gameplay, Huiberts [1] give two perspectives of the use of audio on player experience. One perspective is as a way to *optimise* gameplay, to help the player play and understand the game by giving useful gameplay information through audio. For example, footsteps of enemies in 3D action-oriented games such as Overwatch [9] or Call of Duty 4 [20]. The functionality of the audio is the most important aspect for optimisation, with the idea that giving the player more useful information makes the game more usable, therefore improving the overall experience. The other perspective given is as a way to *dynamise* gameplay, to make the experience more intense, thrilling and gripping. Horror games are a good example, games such as Silent Hill [38], Resident Evil [39] and the like use creepy/unsettling sound effects to make the game feel scarier and more intense. Sounds that dynamise gameplay are generally not functional in the game but are provided just to embellish the experience. It is important to note that the two perspectives are not mutually exclusive and are often very interwoven; sounds will often be designed to both dynamise and optimise at the same time. For example in Left 4 Dead 2 [40], the sound of a crying Witch around the corner works to optimise gameplay in informing the player of the nearby danger, but also dynamises gameplay by instilling a sense of dread that you have to deal with a dangerous enemy, increasing tension.

Viewing these terms with respect to the IEZA framework, optimising sounds are more applicable to the Activity side of the diagram, while dynamising sounds are more applicable to the Setting side. Huiberts states that audio from the Interface domain has the most impact on increasing optimisation, making the game more understandable and usable, audio from the other domains can also increase optimisation but the impact lessens as you travel anti-clockwise around the diagram. The converse is the case for dynamisation -the Affect domain has the most impact, with audio in the Affect domain focused almost entirely on enhancing experience rather than functionality. The impact of the other domains on increasing dynamisation decreases as you travel clockwise around the diagram from the Affect domain.

Figure 2.7 provides a visual representation of how the impact of audio on optimisation and dynamisation changes between the different domains.

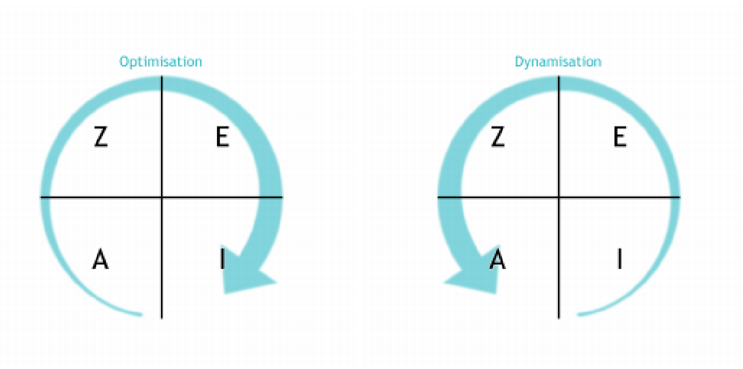


Figure 2.7: Figure to show increasing/decreasing optimisation and dynamisation with respect to the IEZA diagram [1].

The IEZA framework is a good categorisation system for game audio, drawing on concepts such as diegesis that categorise sounds based on their conceptual sources (in or outside of the game world), and mixing them with concepts on the function of the sounds and what they are designed to do, as described by the interdependence dimension. The framework marries these two concepts together in an intuitive way that can easily be used to analyse games in order to gain a clearer understanding on how they use audio.

2.3 Immersion in video games

As with most media, video games are experienced in a subjective way, and with the great diversity of experiences that games have to offer (storming the beaches of Normandy in Call of Duty 2 [41], building a galactic empire in Stellaris [42], or tending to a farm in Stardew Valley [43]), it is difficult to analyse and study player experience that these games offer. One concept that has "consistently proved itself to be an important component of experience players seek from games" [44] and is an important measure of player experience, is the concept of immersion.

Immersion is a term that is widely used colloquially by players and reviewers alike, often associated with the player feeling that they are 'in the game'. The term is used to mean the engagement or involvement a player feels as a result of playing a game. It is worth noting that the idea of being 'in the game' is not a concept of spatial location, but a cognitive player state.

Brown and Cairns [45] conducted a study into what players meant by the term immersion in order to better define it. They found that players were able to distinguish between three different levels of immersion. The first level -engagement- is where players simply invest the time and effort into playing the game. The second level -engrossment- is where players dedicate a lot of attention and emotional involvement into playing the game. The final level is total immersion, where players are completely involved with the game to the point which nothing else matters.

Players considered all three levels to be good features of a game, and it was noted that total immersion was generally a transient feeling that only generally occurs during the most intense periods of the game, while engagement and engrossment levels were experienced for most of the time playing.

It was shown by Hong [46] that the cognitive state of immersion is not exclusive to video games, an experience that is similar to immersion in video games can also be experienced when consuming other media such as books or films. Though a fundamental difference between games and other media is that players have agency within the game world.

Jennett et al. [47] built on Hong's work, devising a questionnaire designed to be a standardised way to measure immersion in games by adapting a questionnaire created by Hong, and incorporating concepts from Brown and Cairns. This resulted in the Immersive Experience Questionnaire (IEQ) which consists of 31 Likert scale questions that address the various subjective experiences that can collectively be called immersion. This questionnaire was validated in a large-scale survey and experiment. Performing factor analysis on the IEQ as part of the validation suggested that immersion has five constituent factors which can be divided into person factors (emotional involvement, cognitive involvement, and real world dissociation) and game factors (challenge, and control). It is important to note that the game factors are really the player perceptions of the games rather than intrinsic properties.

The IEQ is a widely used way of measuring immersion and is utilised in the study described in the following sections.

In terms of specific research into the effect of audio on immersion, Huiberts [1] presents one of the few analyses. Huiberts used a user survey [3] in which users could submit examples of game audio as either "pretty" or "ugly" examples of game audio. Users also submitted descriptions explaining how the audio was used in the game they submitted, and why they felt it was either "pretty" or "ugly". From this it was noted that most players acknowledged that audio can enhance immersion. Huiberts also presented that audio can hinder and reduce immersion when poorly designed and implemented, and noted ways in which audio could enhance and disrupt three types of immersion: Sensory, Challenge-based, and Imaginative, and linked it all into the IEZA framework. While this further categorisation system can be viewed as useful, it was entirely based on data from user surveys rather than any formal study and thus must be treated carefully.

2.4 Summary

Audio is often just as important a consideration in modern video games as the video is and, for some games it is not just used to generally enhance the experience, but it is an important modality and a key part of gameplay. An analysis of this subset of games resulted in three key ways in which they use audio: rhythm, pitch, and spatiality. Bearing in mind these three categories can help when analysing games, helping to improve our understanding about game audio.

A review of work done in the area of analysing for game audio is given, introducing the concept of diegesis that has been borrowed from films as well as considering the sources of sounds and the functionality sound provides in games. Various topologies have been detailed which incorporate these different concepts to produce frameworks for analysing audio and utilising concepts of interactivity and multiple players. The IEZA framework is discussed which takes many of these prior concepts and topologies and creates a framework and categorisation system which can be used to discuss and analyse audio in games.

Finally, a summary of the work done into defining and understanding the concept of immersion in video games is provided, detailing the three levels of immersion and the work done in creating a way of measuring immersion which resulted in the IEQ. Some of the small amount of work that has been done looking specifically into the effect of audio on immersion is also detailed.

3 Creating the Game

In order to perform a study into the effect of audio on immersion in video games, a game had to be created to act as the testbed for the experiment. It was important to create a bespoke game for the study to ensure strict control over all aspects of the game and to allow us to tailor the game precisely to the needs of the study.

3.1 Requirements

The first step of development was creating a list of requirements for the game based on the requirements of the study. It was important that any game designed would be able to meet these requirements so that it would work effectively when used in the study.

The requirements that were created are listed below:

1. The game must be playable equally well with equivalent gameplay, under all modes: visual-only, audio-only, and 'both'.
2. The game must be simple to learn and play.
    a) The game must have simple controls.
    b) The game must have a short but comprehensive tutorial for all modes.
3. The game's audio and graphics must be simple.
4. The game must have five minutes of interesting gameplay.
5. The game must be able to capture player performance data.
6. The game must run on the laptop that will be used for the study.

The most important and restricting requirement that the game had to meet was the requirement that defined the gameplay modes (the video-only, audio-only, and 'both' modes) as these modes made up the three experimental conditions. These three modes were necessarily equivalent in gameplay and playability to prevent confounding the independent variable. The video-only mode would only include video and no audio, the audio-only mode would only include audio and no video (i.e. a black screen), and the 'both' mode would be akin to a standard game with both video and audio present.

This requirement was the most restrictive in terms of the types of games that could be designed as the subset of games that work equally well both video-only, and audio-only is quite small.

The requirement on the simplicity of audio and graphics was another part of this mode equivalence. If either one of the graphics or audio was more complex than the other this would confound the results for the experimental conditions.

The other requirements were defined to ensure that the game would work well in the environment of the study. E.g. The requirements for the game being relatively short, and easy for participants to pick up were added to ensure the duration of the experiment for each participant was reasonably short which would work to reduce the amount of time required for the study as a whole.

Capturing performance data from the participants playing the game was deemed a useful feature to allow for the analysis of player performance as part of the study and was therefore defined as a game requirement.

3.2 Prototyping

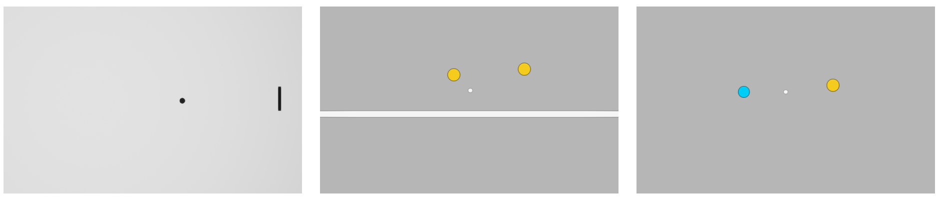
There were three initial ideas for games that would meet the stated requirements and so it was decided that a prototype should be created for each concept to help evaluate which design should be carried forward to full development.

These three ideas were given working names of Bat Pong, Bounce Game and Simon:

Bat Pong was loosely based on the classic game Pong [48] but was designed to be single-player with only one paddle and the goal to keep the ball "in the air" for as long as possible. In order for the game to be playable in an audio-only mode, both the paddle and the ball constantly made a noise (generated by a sine wave), the pitch of which was based on their vertical position i.e. the pitch would increase as the vertical position of the ball/paddle increased. The paddle sound was panned slightly right and the ball sound slightly left so that the player could distinguish between the sound sources. The volume of the ball would increase as it moved towards the paddle and decrease as it moved away from the paddle to give the player spatial information.

Bounce Game was based on one of the sub-games in Rhythm Heaven Fever [13] called Double Date. The concept was that balls would come from either side of the screen, bouncing in-time with the background music. The aim was for the player to press a key/button in time with the music as the ball reached the centre of the screen. In terms of audio, timings were indicated by the rhythm the balls made when bouncing, bouncing once a beat for first three beats and reaching the centre on the fourth.

Simon was a simple call and response style game where a rhythm would be shown/played to the player through the movement of a blue circle and audio being played. The player would have to repeat that rhythm back in-time by pressing the keys/buttons in time. A metronome click would play during gameplay and would also be represented visually by a pulsing white circle.



Bat Pong   Bounce Game   Simon

Figure 3.1: Screenshots of the three prototypes

From the initial prototype, Bat Pong proved a very difficult game to play in the audio-only mode. The easiest way to play in this mode was to follow the vertical position of the ball with the paddle by attempting to constantly keep the two pitches the same. This made the game become a purely pitch-matching game when audio-only, as the distance information wasn't helpful/necessary and meant that the game was not that fun or engaging. In order for the game not to be too difficult in the audio-only mode, the ball speed had to be very low and the paddle speed at least as fast as the ball so that the paddle could follow the ball's vertical position. The size of the paddle also had to be made quite large in order to account for difficulty in matching pitches. While necessary when audio-only, these alterations made the game uninteresting and easy in the video-only and 'both' modes, as part of the challenge of pong comes from the fast-moving ball and slow-moving paddle that forces the player to predict where the ball will move to. It was decided to not continue developing the Bat Pong game due to the difficulties in making all modes equally fun and engaging.

The Simon game worked quite well and was straightforward enough to play across all three modalities. However it was decided not to continue developing the Simon game as it would be difficult to extend the gameplay in a way that would remain interesting for multiple minutes. The gameplay was also similar but inferior to the gameplay provided by Bounce Game.

The Bounce Game had interesting gameplay and a good amount of challenge for first-time players, as well as a clear path for extending the prototype, making it the prototype that was chosen to develop for the study.

As part of progressing from the initial prototype, some modifications were made to the core gameplay of Bounce Game. In the prototype, balls entered from both the left and right sides of the screen with ball audio spatialised (the sound from balls on the left was panned slightly left etc.). In order to simplify the experience, balls were made to only enter from the left of the screen, and spatialisation was removed. The gameplay of the prototype could get boring quite quickly, so with a view to game interest and taking inspiration from Rhythm Heaven Fever, additional types of balls were added. These balls would bounce with different, interesting rhythms so that not all balls would bounce in a simple on-beat rhythm. Rhythm Heaven Fever was used as a reference point for many of the design decisions made in the development of Bounce Game due to it providing very good gameplay which transferred well into audio-only modes.

The next step was to tidy up the prototype code for further development (in particular removing code that related only to the Simon game whose prototype used the same code-base as the Bounce Game).

3.3 Bounce Game Development

An agile-style [49] development method that was somewhat similar to scrum [50] was used during development, with short 1-week sprints followed by analysis, testing and feedback before the deciding on the focus of the next sprint.

The game was created using the Unity engine (Version 2017.2.0) [51] with code written in C# as Unity allows for short prototyping periods and quick iteration cycles. Visual Studio 2015 [52] was the IDE used during development. Using Unity and C# over another engine such as Unreal and C++ (which was used for the Bat Pong prototype) removed the overhead for things such as memory management and improved the overall development workflow. Git [53] was used for version control as well as for maintaining an external code repository on GitHub [54].

3.3.1 Timing & Bytecode

As a rythm-based game, the core issue to tackle in the development of Bounce Game was that of timing; actions needed to happen at a steady rate, precisely in-time with the music and the beat. It was decided to use Bytecode [55] and a data-driven approach to run the timing system, so that the various timings and events in the game could be easily changed and adjusted.

```
FLASH,    SPAWNSINGLEBOUNCE,    BALLBOUNCEINFOUR
```

Figure 3.2: An example line of the bytecode csv file.

The bytecode data is supplied using a csv file where each line contains a comma-separated list of bytecode commands that should all be executed on the same beat/sub-beat. A decision had to be made early on as to what level of granularity to use. Initially the system would execute a line of bytecode once every beat (i.e. every crotchet/quarter note) but this didn't offer enough granularity, as we wanted to be able to run commands on the off-beat (halfway between two beats) so that balls could bounce with off-beat timings. For this reason it was decided to increase the granularity of the timing system by reading a line of the csv four times every beat (i.e. every semiquaver/sixteenth note). This allowed for much more complex and interesting rhythms to be added to the game and allowed for more fine control of the sequence of events.

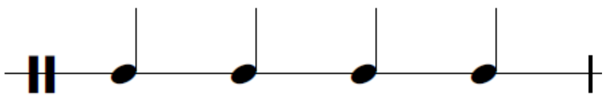


Figure 3.3: Initial command execution regularity

Figure 3.4: Final command execution regularity

The system for reading and executing bytecode with the correct timings was implemented thusly: To ensure the beat timings did not fall out of time with the accompanying music, the DSP time from the system's audio driver was used for all timing calculations. During gameplay a busy-wait loop runs on a separate thread to the main game loop, waiting until it is time for the next line of bytecode to be executed (i.e. the next beat/sub-beat). It then reads the next line of bytecode, executing each command sequentially. Once it has executed all the bytecode in the current line, the thread again busy-waits, first calculating when the next beat is and therefore when the next line of bytecode should be executed using the beats per minute value of the music that it is supplied with. Whilst the game is running, nearly all actions are performed using bytecode commands. Some examples of the operations these commands perform are: causing a ball to spawn, causing a ball to bounce and causing a sound to play.

A full list of the bytecode commands are included in appendix A.1.1

Three different types of balls were made which all bounced with different rhythms/timings. The first type of ball bounces twice, once before reaching the paddle and once at the paddle, with the two bounces being one beat apart (one crotchet/quarter note). The second type of ball bounces three times, it differs from the first type in that it bounces an extra time one semiquaver(sixteenth note) before it reaches the paddle. The third type of ball is quite different from the other two. Its three bounces are regularly spaced one quaver(eighth note) apart meaning that in order for the final bounce to happen on-beat, the first bounce must be off-beat. The precise rhythms of these three balls are shown in figure 3.5.

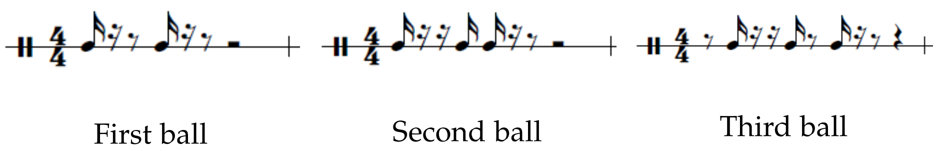


Figure 3.5: Rhythms for the three Balls

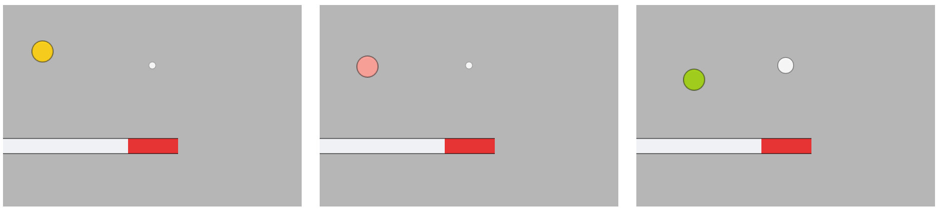
An extremely important aspect of rhythm games is the timing system relating to player input. The rhythmic and timing capabilities of humans are far less precise and less accurate than computers, and it is important to consider this when creating a game where timing is so critical to gameplay. Due to how the bytecode system in Bounce Game functions, timing events are extremely precise, expecting player input to be be entered exactly on-time, with at least the precision of a single frame (1/60th of a second at 60 fps). This is clearly not a reasonable expectation of the player. To combat this, nearly all rhythm-based games [10], [11], [13], [56] use a timing window which extends for a short time before and after the desired input timing. If the player enters their input during this window, it will be counted as correctly in-time, which helps to account for the inherent limits in human timing capabilities. Some games use multiple windows of different sizes for different degrees of timing accuracy. In Bounce Game, the timing window was set as extending 0.124 seconds either side of the specified timing to account for players making slight timing mistakes. The size of the timing window was important; too long and the game would be too easy, too short and the game would be too hard. The window size used for Bounce Game was arrived at through extensive play-testing to find a size of window that felt fair when playing.

3.3.2 Design

It was important that the different types of balls were easily distinguishable both visually and audibly both in terms of what constitutes good game design, but also to satisfy the requirement of equivalent gameplay.

For the visual side of things, different colours were used for the different ball types. A yellow (`#FFD41EFF`) for the first type, a pink (`#FFA59DFF`) for the second type, and a green for the third type (`#A7D41EFF`). These colours were carefully chosen to be distinguishable to those with red-green colour blindness.

For the audio aspect, each type of ball also made a different sound when bouncing. Auditory icons were used, with the first type using the sound of a football, the second type using the sound of a basketball and the third type using the sound of a volleyball. These three sound clips were all audibly distinct enough that the player would be able to know which type of ball was bouncing purely via sound.



First ball | Second ball | Third ball

Figure 3.6: Screenshots of the three ball types to show the colour differences.

Initially, nothing would happen when the ball reached the centre of the screen and was hit/missed by the player, this gave poor feedback to the player and wasn't very exciting. So a move was made from the abstract goal of pressing a key when the balls reached the centre of the screen, to a new goal of triggering a paddle to hit the balls as they passed over the it. This was a simpler concept to convey to players and gave a better visual indicator as to whether the player had entered their input at the correct time (the paddle would hit the ball, sending it upwards, otherwise the ball would fall off the platform).

Success and failure sounds were added in order to give audio-only players the same information. These sounds were designed as earcons, with the success sound designed to feel happy and the failure sound more sad. Another acoustic addition to the game was an "approach" noise. The players that had visuals had an advantage in the sense of visual spatiality, being able to see where the ball was in relation to the paddle/centre of the screen. This gave them spatial information that the audio-only players did not have. To combat this and to avoid actually spatialising the audio of the balls, an approach noise was added. The approach noise was a reversed sample of a crash cymbal hit and plays from the beat before the ball reaches the centre, ending just as the ball reaches the centre. This helped to give a good indicator of the spatial position of the ball in relation to the centre of the screen. The rising pitch and volume of the sound provides a good audio-only analogue for the spatial information of the visuals.

The game had a few distinct states which made up the game flow, a diagram of which is shown in figure 3.7. The game opens to the menu screen (figure 3.10) which allows the game mode to be selected (either audio-only, video-only, and both). Once a game mode is selected the game progresses to the start screen which holds the game until the start button is pressed. The menu screen should not be seen by the participants and so the start screen was added so that the game mode could be selected by the experimenter before the participant entered the room. When the start button is pressed, the game moves into the tutorial and then into the game. Once the game is finished, the game state transitions back to the menu screen.

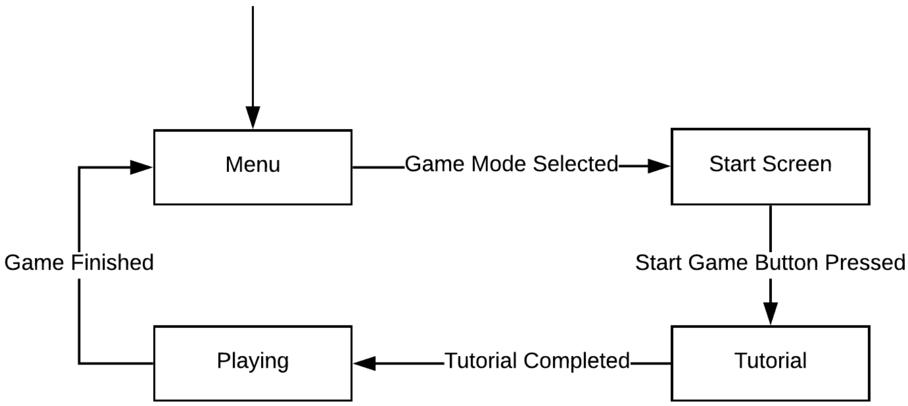


Figure 3.7: High-level game state diagram

3.3.3 Tutorial

An important part of the game was the tutorial as one of the game requirements was that the game must have a short, comprehensive tutorial for all modalities so that participants are able to understand and play the game reasonably well the first time that they play it.

Tutorial systems in video games are difficult to implement well [57], [58] and in a way that is satisfying to the player, evidenced by the wide range of different varieties of tutorial and the research done into creating better tutorials [59], [60]. Some games such as Civilisation IV [61] use large amounts of text to explain the various aspects and concepts of the game to the player before they begin playing. This method can cause issues for some games as many players don't want to read large amounts of text in order to understand the game though for complex strategy games such as the Civilisation games it can be necessary to give this much information to get the player up to speed quickly before they start playing.

Other games such as Hyper Light Drifter [62] almost entirely avoid using text, instead encouraging the player to learn by sending them through carefully crafted gameplay situations that teach the game's mechanics to the player as they progress through them. As this type of tutorial can include enjoyable gameplay, players will often find these forms of tutorials more fun to play [63], however it can be possible for players to get stuck during the tutorial if they don't understand a particular concept. This type of tutorial will often take longer than a text tutorial as they are less condensed, allowing the player to make mistakes and figure mechanics out on their own.

Tutorials often contain a mix of gameplay and text, teaching the player about a mechanic through text and then allowing them to practice/try out the mechanic in gameplay. This makes the tutorial more interesting for the player as they get to play the game but also keeps the tutorial relatively short through concise textual explanations. The Rhythm Heaven [13] games use this form of tutorial, explaining the concept to the player and then giving them a chance to practice the mechanic before moving on and it is this style of tutorial that is used in the Bounce Game.

A difficulty particular to Bounce Game was in creating a tutorial experience that was equally effective and essentially equivalent for each of the three different modalities. Creating a single, standardised tutorial using both video and audio to be used for all three modalities would

Figure 3.8: Civilisation IV uses a great deal of text split into many text boxes in its tutorials due its many complex mechanics.

Figure 3.9: One of the few pieces of text in Hyper Light Drifter, teaching the player they can shoot switches through the use of a gameplay situation.

have been ideal from a development standpoint but due to the nature of the study, would have spoiled the experimental condition of participants experiencing only one of the three modes. For this reason three similar tutorials were created, one for each mode. The aim of which was that by the end of the tutorial section, each participant would be equally able to play the game, irrespective of which tutorial they experienced.

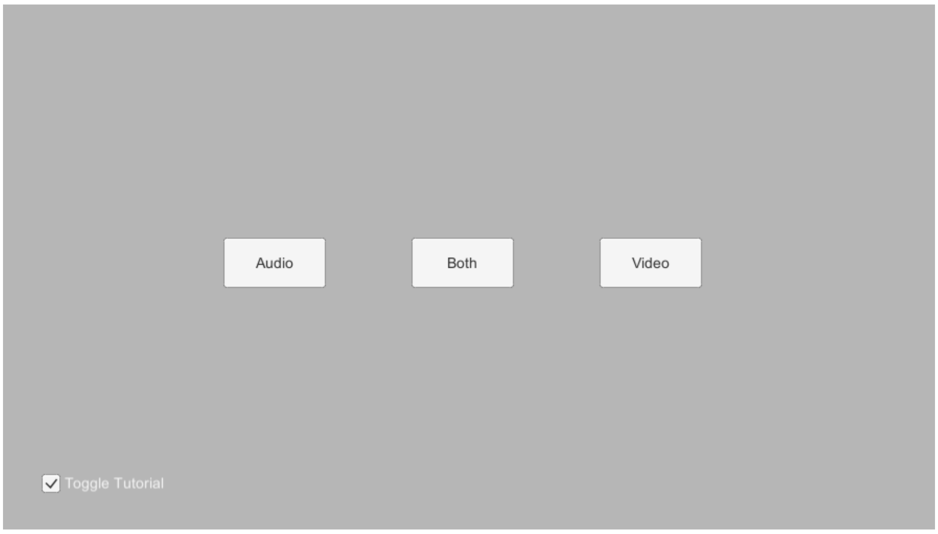


Figure 3.10: The mode select screen. Which allowed the tutorial to be disabled and the game mode to be selected.

The overall concept for each tutorial was the same: Firstly, an overall explanation of the game and its goal (to hit the ball with the correct timing). Following this was an explanation of the first type of ball and how it bounces. The player then practices hitting the first type of ball at the correct time as the game regularly spawns this type of ball. In order for the player to continue with the tutorial, they must correctly hit the ball three times in a row. This explanation-practice segment is repeated for the other two types of ball before the tutorial ends and the main game begins. This tutorial means that by the time the player begins playing the main game, they already have experience with the three types of ball and the overall gameplay of the game. This allows the player to play the game effectively on their first time which in turn allows meaningful data to be gathered from a this single play session, reducing the amount of time the study takes to perform.

The video-only and 'both' modes use essentially the same tutorial, using traditional text boxes to display the text information. The player progresses through the text boxes using the spacebar allowing them to read at their own pace. There are only small differences between these two tutorials with the key difference being that all references to audio are excluded from the video-only tutorial. The audio-only tutorial was necessarily different from the others as written text could not be used. Instead spoken word audio was used to convey the tutorial information to the player. Additionally, the concept of the game had to be conveyed in a more conceptual way, painting a mental picture for the task in the mind of the player through the spoken word. Another addition to the audio tutorial was a short section explaining the various game sounds by playing the sound clip along with a spoken explanation of why/when it occurs. It was necessary for this to be done before the player experienced gameplay due to the more abstract nature of the game under the audio-only condition, so that they had a mental framework from which to work from when playing the game.

The text used for the different tutorials is provided in the appendix A.2.

The tutorial was implemented using a set of bytecode that was specific to the tutorial. This allowed the order and contents of the tutorials to be easily adjusted by simply editing the data files which meant that creating a tutorial just consisted of creating the bytecode file and loading it in. For the gameplay sections of the tutorial, the bytecode system for the main game was used, provided with a different bytecode file. This modular system made the development of these different tutorials quite straightforward.

The bytecode for the tutorial system is included in appendix A.1.2.

3.3.4 Logging

Another one of the requirements of the game was that it must capture player performance data. Data is recorded for each practice section of the tutorial as well as the main game itself. The primary data that's recorded is whether the player hit or missed it each ball (with respect to the timing window), along with precisely how early or late they were from the perfect timing. The type of the ball and the beat number the ball spawned on is also recorded. From this data, the hit-percentage is calculated and recorded. The final metric recorded is the length of time taken to complete each section of the game. E.g. the amount of

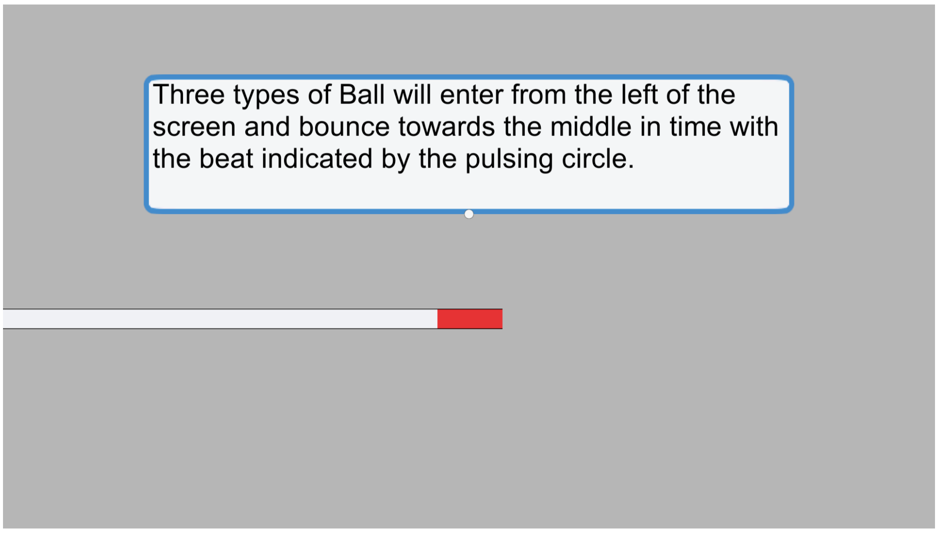


Figure 3.11: Screenshot of the game's tutorial. Tutorial text was displayed in the textbox.

time taken to complete each section of the tutorial. These metrics were decided on as part of the experimental design of the study.

The actual process of logging is quite simple:

A Statistics Manager class was created to be a central class which contains several methods that other areas of the codebase call to notify the statistics manager of events that have taken place. Example methods are "Start Game" and "Ball Spawned" which cause the Statistics Manager to perform certain actions. This system causes a new log file to be created when the game starts in which the data is written as events take place in the game. For each section of gameplay (i.e. a single tutorial practice or the main gameplay section), the timing data gathered is stored in a list using a helper "Timing Statistic" class to store individual data points. Once the section of gameplay has ended, all the timing information recorded for that section is written into the log file so that it can be accessed later.

An example log file is included in appendix A.3 to indicate the format of the logs generated.

4 The Study

The purpose of this study is to explore the effect of audio on immersion in video games. In the following sections, the methodology used for the study will be detailed, the results that were gathered displayed, and the results discussed.

4.1 Hypothesis

The hypothesis that will be tested in this study is the assumption that audio has an effect on immersion in video games, and also that when deprived of a visual interface, players are still able to become immersed in games.

4.2 Design

The experiment measured player engagement and immersion by using the Immersive Experience Questionnaire (IEQ) [47]. The independent variable that was changed was the presence of audio/video in the game played by participants. This variable was divided into three experimental conditions, audio-only (the game is played with audio and no video), video-only (the game is played with video and no audio), and 'both' (the game is played with both audio and video), which acted as the control and the "standard" way games are usually played. The dependent variable of player experience was measured by the IEQ score.

It was important that the game had equivalent gameplay with the same playability across the three experimental conditions to avoid introducing a confounding variable.

A between-participants design was used for this study.

4.3 Participants

A total of 45 participants were gathered through opportunity sampling, fifteen per condition. Out of the participants, 37 identified as male, 7 as female, and one chose not to answer. Participants were all aged 18 and over, with 41 participants in the 18-22 range, and 4 in the 23-27 range.

Participants were asked to report their prior experience in playing rhythm games. 8 reported no prior experience, 34 reported some, and 3 reported a lot of prior experience playing rhythm games.

Participants were also asked whether they regularly play a musical instrument and to rate their musical ability if they do. 18 participants said that they regularly play a musical instrument, with 3 classing themselves as beginners, 10 as intermediate and 5 as advanced.

These demographics were reasonably well split across the three experimental conditions aside from musical ability, in which the audio-only condition leant towards slightly higher ability than the other two conditions.

4.4 Materials

The game was run on a Dell Inspiron Laptop with the following specifications:

- Intel i5-8250U CPU @ 1.60GHZ
- 8GB Ram
- NVIDIA GeForce 940MX Graphics Card
- Windows 10 64-bit Operating System
- 17-inch 1080p 60Hz screen
- Realtek inbuilt audio

All participants wore headphones whilst playing the game, irrespective of what mode they played under, for consistencies sake and also to help reduce any external sound. The headphones used were Audio-Technica ATH-M50X headphones. The laptop was placed on a table which the participants would sit in front of. The laptop screen brightness was set to maximum and volume to 16/100. The game was run in "low" quality mode so that it would run smoothly without any stuttering. The laptop was plugged into power at all times during the experiment.

Online questionnaires were used for the demographic and immersive experience questionnaires using Qualtrics [64]. Participants completed these on the same laptop on which they played the game.

The demographic survey included questions on past video game experience (in particular rhythm-based games), as well as musical experience/ability. These two areas were of interest as they were deemed likely confounding factors.

Other materials used include an informed consent form, separate information sheets (one for each experimental condition) and an experimenter's notes sheet, which helped ensure consistent procedure was applied across the participants.

All of these documents including the questionnaires can be found in appendix B.

4.5 Procedure

Before the participant entered the, room the correct game mode for the experimental condition was pre-selected so that the participants would not see the game mode menu.

Participants would be given the information sheet for the experimental condition they would be operating under. The sheet provided a description of the experiment and what they would be required to do, it also specified how their data would be used and that they could withdraw at any time. Participants would then be allowed to ask questions pertaining to the information they had read and then were asked to fill in the consent form before proceeding.

Next participants were asked to complete the demographic survey before putting on the headphones and beginning the game. They would first play through the tutorial so they could learn how to play, and then they would play the main game itself. Once the participant had finished playing the game they would be asked to complete the immersive experience questionnaire. At the end of the experiment, participants were debriefed and were asked to comment on their experience of playing the game. The entire experiment generally took around 10 minutes to complete per participant.

4.6 Pilot

The experiment was piloted on three Computer Science students, spread across the three experimental conditions. A few minor changes were made to the game as a result of feedback from the pilot participants:

The mouse cursor was set to be hidden during gameplay, and the colour of the second type of ball was changed, from its initial orange to the current pink colour, due to comments that the colour was too similar to the yellow of the first type of ball. The laptop brightness, volume and the game quality settings were decided on during the course of these pilots. Data from the pilots were not included in the final results due to changes made during and after the pilots.

4.7 Results

The Immersive Experience Questionnaire [47] was answered by participants after playing the game and was used to obtain values for immersion.

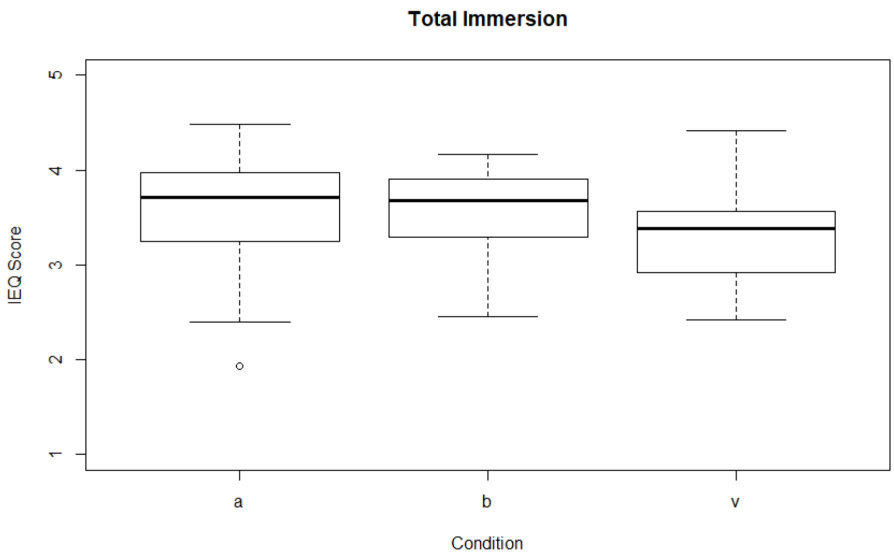


Figure 4.1: Boxplot of Total Immersion scores across the three experimental conditions (audio-only, video-only, and both).

	Audio-Only	Both	Video-Only
Mean	3.54	3.56	3.24
SD	0.69	0.48	0.53

Table 4.1: Descriptive statistics of IEQ scores across the experimental conditions.

4.7.1 Total Immersion

Figure 4.1 shows a boxplot of the values for total immersion across the experimental conditions in which scores for the audio-only and 'both' conditions appear slightly higher than for video-only. A one-way ANOVA showed no significant difference in immersion rating between the conditions ($F(2,42)=1.469$, $p=0.242$), and a calculation of eta-squared for effect size gave $\eta^2=0.065$. Table 4.1 shows descriptive statistics for IEQ score across the three conditions.

The IEQ can be broken up into several subcomponents to glean ratings for various factors that commonly contribute to the feeling of immersion.

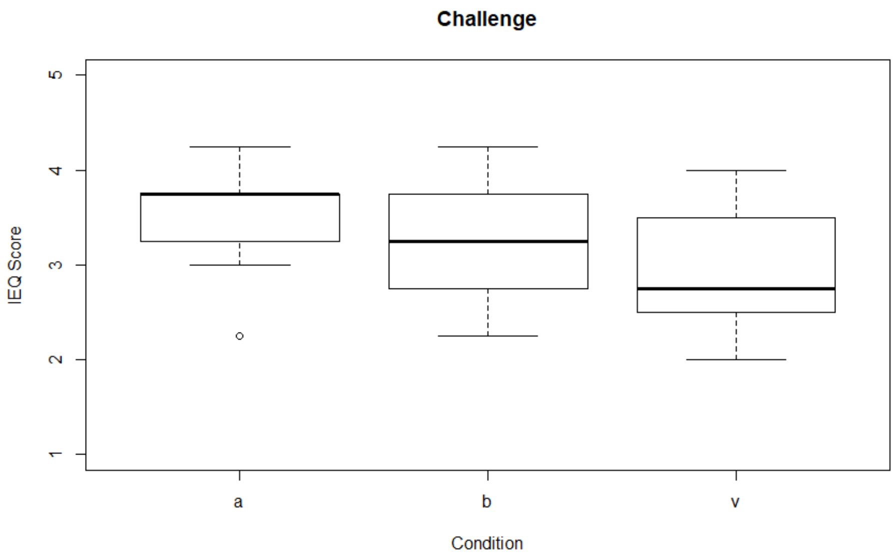


Figure 4.2: Boxplot of Challenge scores across the three experimental conditions.

	Audio-Only	Both	Video-Only
Mean	3.53	3.22	2.98
SD	0.52	0.60	0.64

Table 4.2: Descriptive statistics of Challenge scores across the experimental conditions.

4.7.2 Challenge

Table 4.2 shows descriptive statistics for the challenge score across the experimental conditions. Figure 4.2 shows a boxplot of these scores. The challenge score for the audio-only condition is markedly higher than the other two conditions. A one-way ANOVA showed a statistically significant difference in challenge score between the conditions ($F(2,42)=3.277$, $p=0.048$), and a calculation of eta-squared for effect size gave $\eta^2=0.135$.

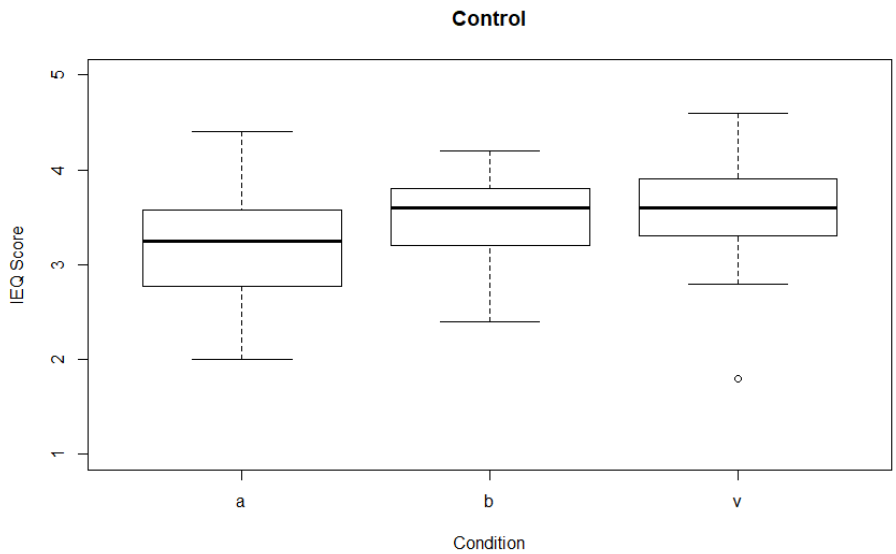


Figure 4.3: Boxplot of Control scores across the three experimental conditions.

	Audio-Only	Both	Video-Only
Mean	3.18	3.45	3.47
SD	0.67	0.53	0.66

Table 4.3: Descriptive statistics of Control scores across the experimental conditions.

4.7.3 Control

Table 4.3 shows descriptive statistics for the control score across the experimental conditions. Figure 4.3 shows a boxplot of these scores. The control scores for the video-only and 'both' conditions are higher than score for the audio-only condition. A one-way ANOVA showed no significant difference in control score between the conditions ($F(2,42)=1.005$, $p=0.375$), and a calculation of eta-squared for effect size gave $\eta^2=0.046$.

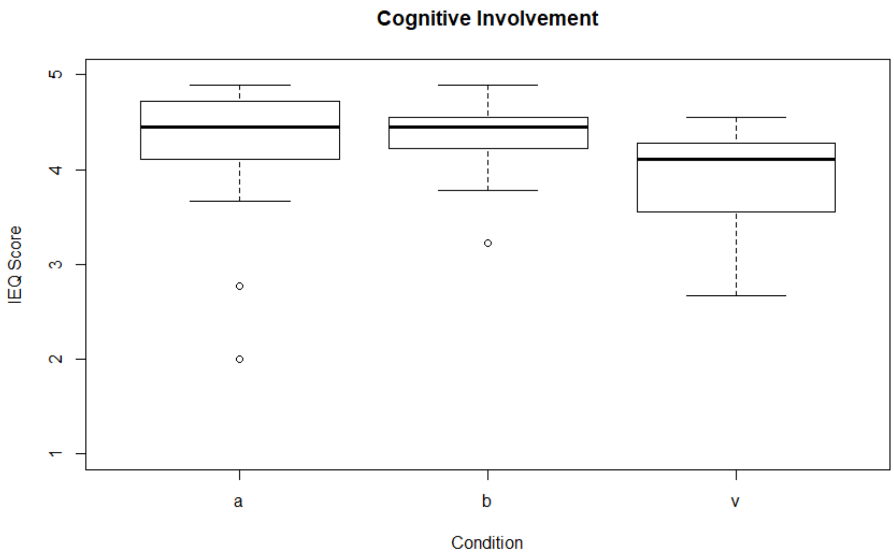


Figure 4.4: Boxplot of Cognitive Involvement scores across the three experimental conditions.

	Audio-Only	Both	Video-Only
Mean	4.19	4.34	3.88
SD	0.82	0.41	0.56

Table 4.4: Descriptive statistics of Cognitive Involvement scores across the experimental conditions.

4.7.4 Cognitive Involvement

Table 4.4 shows descriptive statistics for the cognitive involvement score across the experimental conditions. Figure 4.4 shows a boxplot of these scores. The cognitive involvement scores for the audio-only and 'both' conditions are higher than the score for the video-only condition. A one-way ANOVA showed no significant difference in cognitive involvement score between the conditions ($F(2,42)=2.117$, $p=0.133$), and a calculation of eta-squared for effect size gave η^2=0.092.

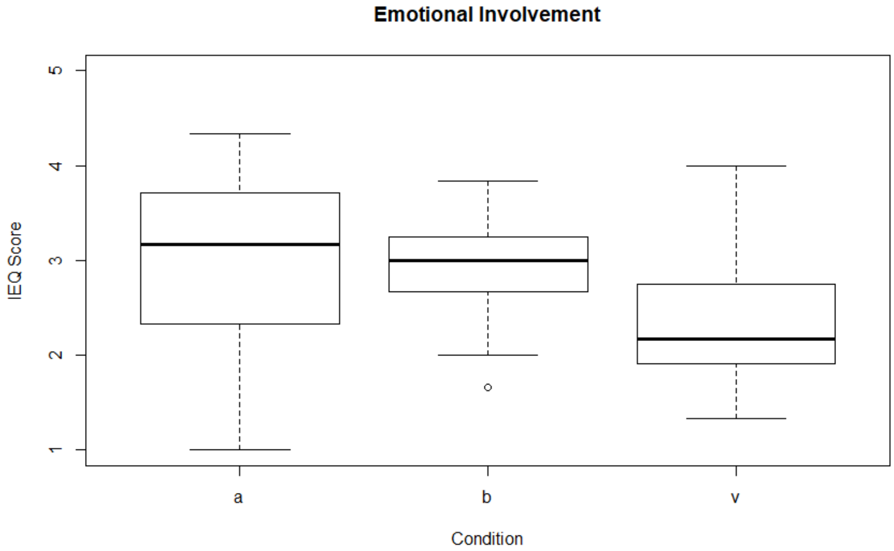


Figure 4.5: Boxplot of Emotional Involvement scores across the three experimental conditions.

	Audio-Only	Both	Video-Only
Mean	3.01	2.89	2.34
SD	0.98	0.55	0.73

Table 4.5: Descriptive statistics of Emotional Involvement scores across the experimental conditions.

4.7.5 Emotional Involvement

Table 4.5 shows descriptive statistics for the emotional involvement score across the experimental conditions. Figure 4.5 shows a boxplot of these scores. The emotional involvement scores for the audio-only and 'both' conditions are higher than the score for the video-only condition, with a wider spread for the audio-only condition. A one-way ANOVA showed no significant difference in emotional involvement score between the conditions (F(2,42)=3.124, p=0.054), and a calculation of eta-squared for effect size gave η^2=0.129.

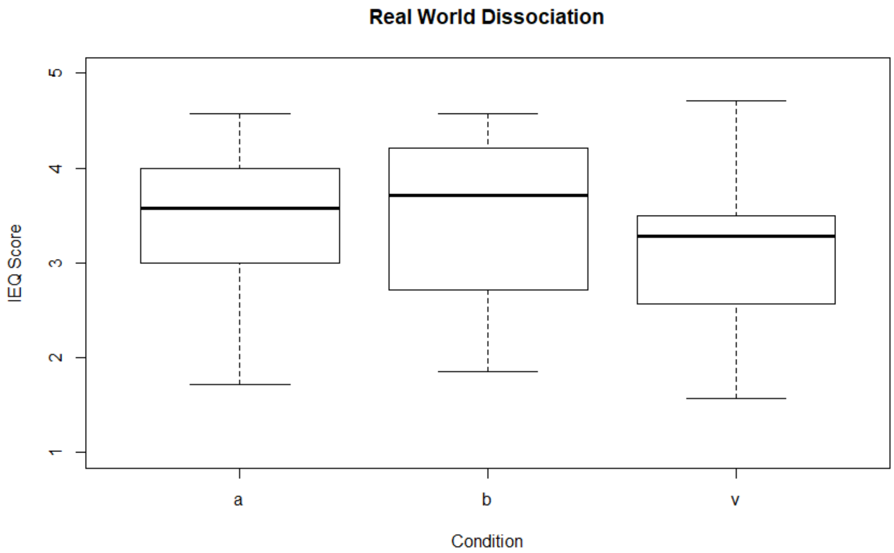


Figure 4.6: Box plot of Real World Dissociation scores across the three experimental conditions.

	Audio-Only	**Both**	**Video-Only**
Mean	3.42	3.42	3.17
SD	0.81	0.90	0.85

Table 4.6: Descriptive statistics of Real World Dissociation scores across the experimental conditions.

4.7.6 Real World Dissociation

Table 4.6 shows descriptive statistics for the real world dissociation score across the experimental conditions. Figure 4.6 shows a box plot of these scores. The real world dissociation scores for the 'both' and audio-only conditions are the same, and higher than the score for the video-only condition. A one-way ANOVA showed no significant difference in real world dissociation score between the conditions ($F(2,42)=0.418$, $p=0.661$), and a calculation of eta-squared for effect size gave $\eta^2=0.020$.

4.7.7 Player Performance

Player performance data was gathered from each participant. This data primarily included the percentage of balls that the participant hit successfully. A summary of this data is included below:

Table 4.7 shows descriptive statistics for the player performance score across the experimental conditions. Figure 4.7 shows a box plot of these scores. The player performance for the video-only and 'both' conditions are higher than for the audio-only condition. A one-way ANOVA showed a significant difference in real world dissociation score between the conditions ($F(2,42)=22.4$, $p<0.001$), and a calculation of eta-squared for effect size gave $\eta^2=0.516$.

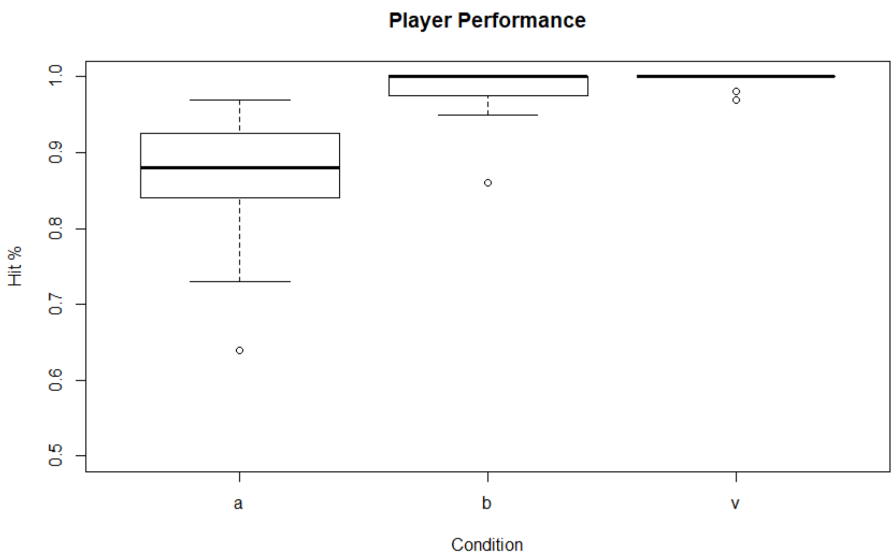


Figure 4.7: Box plot of Player Performance across the three experimental conditions.

	Audio-Only	Both	Video-Only
Mean	86.3%	98.0%	99.5%
SD	9.32%	3.69%	1.01%

Table 4.7: Descriptive statistics of Player Performance across the experimental conditions.

4.8 Discussion

There are multiple metrics by which we can analyse the results of the study, but the key result is the score for total immersion gathered by the IEQ. From this we see that audio certainly has some effect upon the amount of immersion that participants experienced. Though not a statistically significant difference, the mean total immersion scores for the two conditions that include audio were certainly higher than for the video-only condition, and with an effect size of roughly 6.5%, audio clearly has *some* effect on immersion. It is clear to see that it is possible for players to become immersed in an audio-only game, indeed for the game used in this study, the participants found the audio-only version to be more immersive than the video-only version.

Breaking the IEQ total immersion score down into its constituent factors gave us further insight.

The factors of most interest are those of Challenge, and Emotional Involvement. There was a statistically significant difference between the challenge scores, with the two conditions that utilised audio perceived as more challenging (having higher mean scores for challenge). The audio-only condition had a higher mean challenge score than the other two which is not unexpected as there are very few video games which are exclusively audio, and do not include any video/visual elements. Therefore it is likely that most participants had much less experience with this type of game and thus found it much more challenging. This increased challenge felt by participants in conditions that utilise audio is borne out in the player performance metrics in which the participants that played audio-only had drastically worse performances than the other two conditions. The variation in player performance scores between conditions is statistically significant with a very large effect size.

There are several possible causes for this wide gap in performance, it could be due to participants being unfamiliar and uncomfortable with a purely audio-based experience that made them find the game more difficult, or that without the concrete game-state information that visuals provide, they found it harder to understand the game-state, leading to more mistakes due to a mismatch in the believed game-state versus the actual game-state. Another possibility is that these audio-only participants are having to do more mental work in interpreting, understanding, and visualising what is happening in the game due to being limited to only audio information. This increased mental workload could reduce player ability, as less mental resources can be spent on ensuring the correct timing.

In support of these latter two hypotheses are the metrics gathered for control and cognitive involvement. Whilst neither metrics are statistically significant, control scores are higher for the two conditions with video than for the audio-only condition, indicating that players felt a greater degree of control when they had access to video. It makes sense that this increased control would result in fewer mistakes from these players. The cognitive involvement scores show a higher involvement in the conditions that used audio than the video-only condition, though there isn't much difference between the 'both', and audio-only conditions.

Another interesting part of the player performance results that likely affected challenge scores is that the video-only participants had an extremely high average performance score of 99.5%. This metric corresponds with the anecdotal evidence gathered by asking participants how they found the game during the debrief portion of the experiment. Most video-only participants said that they found the game quite boring and straightforward, it seems this is likely due to the low level of challenge they experienced, corroborated by their performance statistics and challenge scores.

It is similarly interesting that the in the average performance under the video-only condition was higher than the average performance for the 'both' condition. Players performed worse when they had access to both audio and video than when they only had video. One would assume that the additional information channel provided by audio would only help players but this does not appear to be the case. One possible cause for this discrepancy is that the two sources of sensory input are interfering with each other, distorting the player's perception of time [65] in a similar way to the stopped-clock illusion(chronostasis) [66], [67]. It could also simply be that participants were distracted by the music and focused on listening to the audio instead of focusing on the game. Either way the difference is quite small and may just be due to random variation, but it does point towards an interesting possible avenue of research.

There is a clear difference in mean emotional involvement scores across the conditions; the conditions using audio have much higher average scores. It seems likely that the cause of this higher level of emotional involvement with the game in the audio conditions is related to how challenging the players found the game, with increased challenge making the participants become more emotionally involved and care more about the performing well. The differences between emotional involvement across conditions are marginally statistically significant, and have a reasonable effect size of 12.9% so the presence of audio clearly has an impact on emotional involvement.

Across the different factors of immersion there is consistently a clear difference between the mean scores of the video-only condition and the 'both' condition. Excluding the control factor, in all other cases, the 'both' condition was markedly higher than the video-only condition, indicating that audio does indeed have the effect of increasing immersion as was posited. The data obtained for total immersion was not statistically significant which may be due to the fact that the control factor was working opposite to the other effects, therefore reducing the overall significance of total immersion. It is an interesting result that the control was felt to be worse in the conditions with audio than in the condition without, something that might merit further research.

For the other section of the hypothesis that was concerned with players being able to become immersed when deprived of a visual interface, we can see that for the audio-only and 'both' conditions, the mean scores are either more or less equivalent, or higher for the audio-only condition for almost all factors (excluding control in particular). This indicates that not only are players indeed able to become immersed in the game when deprived of a visual interface, but for some factors the lack of the visual interface resulted in a more immersive experience.

4.9 Evaluation of Bounce Game

Agile development was used as the development methodology which worked well with the chosen game engine. Unity allowed for fast feature implementation and iteration which worked well with the sprint-based system of agile development and scrum.

In terms of requirements, the game met nearly all of them. The one requirement that wasn't fully met was requirement 4, that the game must have five minutes of interesting gameplay. Excluding the gameplay that was part of the tutorial, the game provided around two and a half minutes of interesting gameplay due to the length of the music that was used for the game being roughly that long. Extending the length of the music that was used (the song from Rhythm Heaven: Fever's Double Date game) would have been difficult and creating an entirely new song to use would have been an even larger amount of work. Anecdotal evidence from speaking with participants indicated that most did not find the length of the gameplay problematic, and most felt that the length was reasonable. It seems unlikely that this length of gameplay had a negative effect on immersion as a study performed by Denisova and Cairns [68] showed no correlation between immersion and length of game for similarly short gameplay experiences. Still, it is important to note that the game did not fully satisfy this requirement.

The game successfully met all other requirements, controls were simple (just the spacebar during gameplay and the trackpad for menus) and the tutorial worked well in educating players on how to play the game -all players understood how the play the game at the completion of the tutorial and were able to play adequately. The game also successfully captured performance data for both main gameplay and the tutorial practice gameplay.

The game ran on the laptop properly once the quality setting was set to low. On medium and higher graphical quality settings the game would sometimes stutter which made it very difficult to play as the timing kept changing due to fluctuating lag causing the game to run slower-than-intended. This did not affect the results in any way as this issue was solved in the pilot stages of the study. It seems likely that the cause of this lag was some of the additional graphical effects that Unity performs under higher quality levels, though these were entirely unnecessary for the simple graphics of the game. It is also possible that the busy-wait loop used in the bytecode thread could have also contributed.

The use of bytecode for controlling the gameplay along with the audio driver-based timing system worked extremely well. The bytecode allowed for very quick changes to be made to the game without requiring recompilation. The biggest issue with the bytecode was in editing the bytecode csv files themselves, which could be very difficult as it was hard to translate the desired timings of things into bytecode effectively which could waste time. On a larger project, or with a greater amount of time, it would be prudent to create a tool which would assist in editing the bytecode, possibly allowing a visual interface to be used to edit the file.

Further improvements that could be made to the game include: increasing performance by removing the busy-wait loop used in the bytecode thread and replacing it with an event-based system. Similarly, the logging system could be improved by utilising an event-based/message-passing system, thereby reducing the current tight coupling between the Statistics Manager and several other classes. Making these adjustments would however have required quite a large amount of time and effort, which was better spent in other areas of development. In addition, the visual appeal of the game could be increased by improving the menus which are currently very basic and unappealing, though only the start screen was seen by players.

A difficult part of developing the game was working with the limitations imposed by the audio-only mode requirement. Most conventional game design wisdom treats audio as supplementary and it proved quite difficult to transfer many of the traditional game design tricks into this audio-only space. For example, there is a concept of "juice" in game design [69], which encompasses parts of games that are added to make them more satisfying and fun to play by increasing player-feedback. Examples of juice are things like floating damage numbers that appear when you deal damage to an enemy or adding screenshake [70] when the player takes damage, these techniques make the game more fun and engaging. However most of these techniques are purely video or purely audio which means it is very hard to add these and maintain parity between the various game modes.

Even considering adding simple hit-streak or score counters, which are very common forms of juice was problematic. It would be easy in video modes to include text on screen displaying the current score/hit-streak, but for the audio-only mode it would require either the score to be given through audio using spoken word whenever the value changed, or by using some set of abstract sound effects to communicate to the player. Either of these options would only add audio clutter to the game, which players would likely find confusing and distracting.

The jack-of-all-trades approach that the game necessarily had to have to preserve parity between the different modes, meant that the quality of the game under each individual mode was much weaker than it could have otherwise been, which no doubt reducing the possible level of immersion for the game. However working with these restrictions was important in ensuring that the game worked very well in the study for which it was designed.

4.10 Experimental Validity and Limitations

4.10.1 Internal Validity

Participants were gathered through opportunity sampling of students from several areas of the university campus. Due to this, there was no way of controlling previous experience with video games. Similarly, there was no way of controlling musical experience as this was discovered as part of the demographic survey which had to occur after participants had already been assigned to a particular condition. Both of these limitations may have affected results, though any large variances in these areas between conditions are noted in the participants section.

As mentioned during the evaluation of Bounce Game, one of the requirements was that the game must provide a comparable experience between the different modes. This requirement was met, meaning the gameplay and the overall game experience was indeed comparable between the modes, with the only difference being whether an audio/video interface was present. Due to this comparable experience, the study was indeed fit to address the research questions on the effects of audio on immersion.

4.10.2 External Validity

Due to opportunity sampling, participants were all students aged between 18 and 27 with most participants playing video games at least multiple times a week. This limits the generalisability of the study as it represents a reasonably small portion of the population. Rhythm games -the genre of game created for the study- are a relatively niche genre of game that are often quite dissimilar to most other games. The use of a rhythm game may slightly reduce the generalisability of the study. The gameplay is also quite simplified in the study game as part of ensuring parity among the different modes which means that the game may not be that representative of games in general.

4.10.3 Ecological Validity

Participants used a laptop that was placed on a desk in front of them. They used the spacebar primarily to interact with the game. In real-world situations, many players would customise their gaming setup, often with larger monitors and, in particular for the style of game used in the study, might often use a controller to interact with the game instead of the keyboard. Even so the the setup of the experiment was generally similar to one that might be found in a real-life situation.

Most popular games are much further developed than the one used in this study, with more juice and features and generally providing a tighter, more polished experience. Also the large majority of games are not audio-only/video-only. Due to this, the ecological validity of the study is slightly reduced due to the game used not being particularly generalisable to a real-world situation.

5 Conclusions

5.1 Contributions to Literature

This report contributes a study which examines the effect of audio on immersion, including how to develop a rhythm-based game that may be used in such a study. It contributes the results of the study, finding that that addition of audio slightly increases immersion, as well as the knowledge that it is possible for players to become immersed in audio-only games. This report also contributes an analysis of how audio is used in various types of games where audio is an important modality.

5.2 Further Work

This study is initial exploratory research into the effect of audio on video game immersion and there are many different avenues of research that can be taken from this point.

While from this study it would seem that audio does have an effect on immersion and that audio-only games can be immersive, results on the extent of this effect are lacking. Future work could perform studies to better ascertain the extent of the effect of audio on immersion, possibly exploring different genres/styles of game as the extent of audio's effect seems likely to differ between genres (horror games versus role-playing games for example). Further research could also be performed into the area of audio-only games in order to find out what can be done in order to improve immersion in these games. The area of audio research with respect to immersion in video games is still very much in early stages and any well-performed future work would likely improve our understanding and overall confidence in our knowledge.

The results of study also hint at audio having an effect on the difficulty and amount of challenge a game is perceived to have; and while some work has been performed in this area previously (particularly into player performance [71]), there is still a good amount of untrodden ground. Similarly, there was an interesting result in the perception of control that players felt (control was perceived as higher for the video-only condition) which was surprising. It seems that the addition of audio reduced the control that players felt they had which may also relate to the challenge and performance metrics. Further work could investigate these effects and determine what factors affect challenge (e.g. genre of game, style of sound used etc.) which may help in better explaining the results gathered in this study.

Further work could also be done in improving and extending the game. Gameplay could be improved with more interesting types of balls and more complex rhythms. Better animations and audio could be added to improve game-feel, and juice could be added to various areas including how the paddle hits the balls etc. The game length could be extended and some form of difficulty system added, possibly multiple levels/stages with more difficult rhythms and timings or increased speeds and smaller input windows for example. An interesting scoring system could be added to give more player feedback which may make players feel more involved. Menus could be improved as well, making the first impression of the game more positive. These additions may increase the amount of immersion players experience when playing the game and could possibly be used to gain a further insight into the effect of audio on games.

5.3 Summary

The aim of the study performed was to discover what effect audio has on immersion in video games and if players can become immersed in a game when deprived of a visual-interface. A game with three distinct modes was developed in order to investigate this question and used as part of a study to gather ratings for immersion of these three modes. The results of the study indicated that audio appears to have the effect of increasing immersion and that it is possible for players to become immersed in an audio-only game. This study was primarily exploratory as little previous research has not been performed on the effect of audio before. As such, future work should be undertaken to better understand both the effect and extent of audio on immersion.

6 References

6.1 Bibliography

[1] S. Huiberts, 'Captivating sound the role of audio for immersion in computer games', PhD thesis, University of Portsmouth, 2010.

[2] R. van Tol and S. Huiberts, 'Ieza: A framework for game audio', *Retrieved January*, vol. 21, p. 2013, 2008.

[3] R. van Tol and S. Huiberts. (2018). PrettyUglyGameSoundStudy, [Online]. Available: http://prettyuglygamesoundstudy.com/ (visited on 27/04/2018).

[4] R. Stevens and D. Raybould, *The game audio tutorial: A practical guide to creating and implementing sound and music for interactive games*. Taylor & Francis, 2013.

[5] J. Schell, *The Art of Game Design: A book of lenses*. AK Peters/CRC Press, 2014, p. 351.

[12] D. Bruno. (2009). Battle music in mother 3, [Online]. Available: https://danbruno.net/writing/mother3/ (visited on 18/04/2018).

[15] B. Hillier. (2016). The witness: How to solve those damn jungle puzzles, [Online]. Available: https://www.vg247.com/2016/01/31/the-witness-how-to-solve-those-damn-jungle-puzzles/ (visited on 18/04/2018).

[16] lmsouza-br. (2013). Reddit (r/thewitness) "audio puzzles", Reddit, r/TheWitness, [Online]. Available: https://www.reddit.com/r/TheWitness/comments/5kgww1/audio_puzzles/ (visited on 18/04/2018).

[18] S. Lawlor and T. Neumann. (2016). Overwatch - the elusive goal: Play by sound, GDC, [Online]. Available: https://www.gdcvault.com/play/1023317/Overwatch-The-Elusive-Goal-Play (visited on 18/04/2018).

[21] (2017). Valve developer community: Soundscape, Valve Software, [Online]. Available: https://developer.valvesoftware.com/wiki/Soundscape (visited on 18/04/2018).

[22] W. W. Gaver, 'Auditory icons: Using sound in computer interfaces', *Human-computer interaction*, vol. 2, no. 2, pp. 167–177, 1986.

[23] M. Grimshaw and G. Schott, 'A conceptual framework for the analysis of first-person shooter audio and its potential use for game engines', *International Journal of Computer Games Technology*, vol. 2008, p. 5, 2008.

[24] W. Staab. (2014). Auditory icons, earcons, and speech, [Online]. Available: http://hearinghealthmatters.org/waynesworld/2014/auditory-icons-earcons-speech/ (visited on 17/04/2018).

[26] J. Friberg and D. Gärdenfors, 'Audio games: New perspectives on game audio', in *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology*, ACM, 2004, pp. 148–154.

[27] M. Chion, *Audio-vision: sound on screen*. Columbia University Press, 1994.

[28] P. Schaeffer, *Traité des objets musicaux*. Le Seuil, 2016.

[29] I. Ekman, 'Meaningful noise: Understanding sound effects in computer games', *Proc. Digital Arts and Cultures*, 2005.

[31] K. Collins *et al.*, 'An introduction to the participatory and non-linear aspects of video games audio', *Essays on sound and vision*, pp. 263–298, 2007.

[33] S. Huiberts. (2013). Ieza framework high resolution png, [Online]. Available: http://captivatingsound.com/by-request-ieza-framework-high-resolution-png/ (visited on 18/04/2018).

[36] (2018). List of portal soundscapes, Valve Software, [Online]. Available: https://developer.valvesoftware.com/wiki/List_of_Portal_soundscapes (visited on 18/04/2018).

[37] (2014). List of portal 2 soundscapes, Valve Software, [Online]. Available: https://developer.valvesoftware.com/wiki/List_of_Portal_2_Soundscapes (visited on 18/04/2018).

[44] P. Cairns, A. Cox and A. I. Nordin, 'Immersion in digital games: Review of gaming experience research', *Handbook of digital games*, vol. 1, p. 767, 2014.

[45] E. Brown and P. Cairns, 'A grounded investigation of game immersion', in *CHI'04 extended abstracts on Human factors in computing systems*, ACM, 2004, pp. 1297–1300.

[46] R. Hong, 'Immersion in reading and film as a function of personality', PhD thesis, B. Sc. Thesis, Dept. of Psychology, University College London, 2006.

[47] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs and A. Walton, 'Measuring and defining the experience of immersion in games', *International journal of human-computer studies*, vol. 66, no. 9, pp. 641–661, 2008.

[49] (2001). Principles behind the Agile Manifesto, Agile Alliance, [Online]. Available: http://agilemanifesto.org/principles.html (visited on 29/04/2018).

[50] (2018). What is Scrum?, Scrum.org, [Online]. Available: https://www.scrum.org/resources/what-is-scrum (visited on 29/04/2018).

[51] Unity Technologies. (2017). Unity engine (2017.2.0) release notes, [Online]. Available: https://unity3d.com/unity/whats-new/unity-2017.2.0 (visited on 15/04/2018).

[52] (2015). Welcome to Visual Studio 2015, Microsoft, [Online]. Available: https://msdn.microsoft.com/en-us/library/dd831853.aspx?f=255&MSPPError=-2147217396 (visited on 29/04/2018).

[53] (2018). Git, Software Freedom Conservancy, Inc., [Online]. Available: https://git-scm.com/ (visited on 29/04/2018).

[54] (2018). GitHub.com, GitHub, Inc., [Online]. Available: https://github.com/ (visited on 29/04/2018).

[55] R. Nystrom, *Game programming patterns*. Genever Benning, 2014.

[57] E. Adams. (2011). The designer's notebook: Eight ways to make a bad tutorial, [Online]. Available: https://www.gamasutra.com/view/feature/134774/the_designers_notebook_eight_.php (visited on 16/04/2018).

[58] J. Bycer. (2015). Examining organic tutorials, [Online]. Available: https://www.gamasutra.com/blogs/JoshBycer/20150702/247703/Examining_Organic_Tutorials.php (visited on 16/04/2018).

[59] N. Hedges. (2017). Video game tutorials: How do they teach?, [Online]. Available: https://www.gamasutra.com/blogs/NathanHedges/20171013/307378/Video_Game_Tutorials_How_Do_They_Teach.php (visited on 16/04/2018).

[60] (2015). Miyamoto on world 1-1: How nintendo made mario's most iconic level, Eurogamer, [Online]. Available: https://www.youtube.com/watch?v=zRGRJRUWafY&ab_channel=Eurogamer (visited on 16/04/2018).

[63] (16th Mar. 2011). Tutorials 101 - how to design a good game tutorial - extra credits, Extra Credits, [Online]. Available: https://www.youtube.com/watch?v=BCPcn-Q5nKE (visited on 22/04/2018).

[64] (2018). Qualtrics, Qualtrics, [Online]. Available: https://www.qualtrics.com/uk/ (visited on 18/04/2018).

[65] V. Van Wassenhove, D. V. Buonomano, S. Shimojo and L. Shams, 'Distortions of subjective time perception within and across senses', *PloS one*, vol. 3, no. 1, e1437, 2008.

[66] D. M. Eagleman, 'Human time perception and its illusions', *Current opinion in neurobiology*, vol. 18, no. 2, pp. 131–136, 2008.

[67] Z. Shi and D. Burr, 'Predictive coding of multisensory timing', *Current opinion in behavioral sciences*, vol. 8, pp. 200–206, 2016.

[68] A. Denisova and P. Cairns, 'Adaptation in digital games: The effect of challenge adjustment on player performance and experience', in *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, ACM, 2015, pp. 97–101.

[69] M. Jonasson and P. Purho. (2012). Juice it or lose it, GDC Europe, [Online]. Available: http://www.gdcvault.com/play/1016487/Juice-It-or-Lose (visited on 23/04/2018).

[70] J. W. Nijman. (2013). The art of screenshake, Vlambeer, [Online]. Available: https://www.youtube.com/watch?v=AJdEqssNZ-U.

[71] L. Varghese, S. R. Mathias, S. Bensussen, K. Chou, H. R. Goldberg, Y. Sun, R. Sekuler and B. G. Shinn-Cunningham, 'Bi-directional audiovisual influences on temporal modulation discrimination', *The Journal of the Acoustical Society of America*, vol. 141, no. 4, pp. 2474–2488, 2017.

6.2 Ludography

[6] *bit Generations: Soundvoyager*, Nintendo Game Boy Advance, Skip Ltd., 2006.

[7] *Papa Sangre*, Apple iOS, Somthin' Else, 2010.

[8] *Dota 2*, PC, Valve Corporation, 2013.

[9] *Overwatch*, PC, Blizzard Entertainment, 2016.

[10] *Guitar Hero*, Sony PlayStation 2, Harmonix, 2015.

[11] *Crypt of the NecroDancer*, PC, Brace Yourself Games, 2015.

[13] *Rhythm Heaven Fever*, Nintendo Wii, Nintendo, 2012.

[14] *The Witness*, PC, Thekla, Inc., 2016.

[17] *String Theory*, PC, Zenisport, 2016.

[19] *Find the Invisible Cow*, Scriptist, 2013. [Online]. Available: https://findtheinvisiblecow.com/.

[20] *Call of Duty 4: Modern Warfare*, PC, Infinity Ward, 2007.

[25] *Super Mario Bros.* Nintendo Entertainment System, Nintendo, 1985.

[30] *Bastion*, PC, Supergiant Games, 2011.

[32] *Super Mario Kart*, Super Nintendo Entertainment System, Nintendo, 1992.

[34] *The Witcher 3: Wild Hunt*, PC, CD Projekt Red, 2015.

[35] *Tetris*, Various, Various, 1984.

[38] *Silent Hill*, Sony Playstation, Konami, 1999.

[39] *Resident Evil*, Sony Playstation, Capcom, 1996.

[40] *Left 4 Dead 2*, PC, Valve Corporation, 2009.

[41] *Call of Duty 2*, PC, Activision, 2005.

[42] *Stellaris*, PC, Paradox Interactive, 2016.

[43] *Stardew valley*, PC, Chucklefish, 2016.

[48] *Pong*, Arcade, Atari, 1972.

[56] *pop'n music 20 fantasia*, Arcade, Konami, 2011.

[61] *Sid Meier's Civilization VI*, PC, Firaxis Games, 2016.

[62] *Hyper Light Drifter*, PC, Heart Machine, 2016.

A Bounce Game

A.1 Bytecode

A.1.1 Regular Gameplay

Command	Description
FLASH	Causes the pulsing metronome to flash
INPUT	Specifies that a player input is expected on this sub-beat
COUNTINBEAT	Makes the count-in sound effect play
GOTOSTART	Jumps to the start of the bytecode, used for testing
GOTOTWENTY	Jumps to line 20 of the bytecode, used for testing
GOTOTWENTYFOUR	Jumps to line 24 of the bytecode, used for testing
STARTMUSIC	Starts the music playing
BALLBOUNCEINFOUR	Causes a ball to bounce in 4 sub-beats
SPAWNSINGLEBOUNCE	Spawn the first type of ball
SPAWNDOUBLEBOUNCE	Spawn the second type of ball
SPAWNTRIPLEBOUNCE	Spawn the third type of ball
BALLBOUNCEINONE	Causes a ball to bounce in 1 sub-beat
BALLBOUNCEINTWO	Causes a ball to bounce in 2 sub-beats
BALLBOUNCEINTHREE	Causes a ball to bounce in 3 sub-beats
BALLNOISE	Plays a default ball noise
SINGLEBALLNOISE	Plays the first type of ball noise
DOUBLEBALLNOISE	Plays the second type of ball noise
TRIPLEBALLNOISE	Plays the third type of ball noise
ENDGAME	Finishes the game

Table A.1: List of bytecode commands for regular gameplay

A.1.2 Tutorial

Command	Description
NEXTMESSAGE	Causes the next message to occur (either text or audio)
NEXTPRACTICE	Causes the next practice session to occur
ENDPRACTICE	Ends the current practice session
ENDTUTORIAL	Ends the tutorial

Table A.2: List of bytecode commands for the tutorial

A.2 Tutorial text

A.2.1 Audio-Only

1. Balls will bounce in time with the music, towards a paddle which you control by pressing the space bar.

2. As a ball approaches the paddle you will hear this noise: <approach noise> just beforehand. This indicates when the ball is directly over the paddle.

3. Press space at the right time to hit the ball. If you are successful you will hear this sound: <success noise>. If you miss you'll hear this: <failure noise>

4. Each ball bounces slightly differently and makes a different sound.

5. Let's try practising with the first type of ball, it makes this <first ball noise> sound.

6. Successfully hit the ball three times in a row to move on to the next part of the tutorial.

7. Nicely done!

8. The next type of ball does an extra bounce just before reaching the paddle. It sounds like this: <second ball noise>.

9. Hit the ball three times in a row to move on to the next part of the tutorial.

10. Nice Job!
11. The final type of ball has a slightly different timing to the others and makes this sound: <third ball noise>.
12. Hit the ball three times in a row to move on to the next part of the tutorial.
13. Nice Job!
14. Now it's time for the real thing, good luck!

A.2.2 Video-Only

1. Three types of Ball will enter from the left of the screen and bounce towards the middle in time with the beat indicated by the pulsing circle.
2. Press the space bar to move the paddle when the ball reaches the middle of the screen in time with the beat.
3. Let's try this with the first, yellow ball. Hit it with the paddle three times in a row to continue to the next part of the tutorial.
4. Nice Job!
5. The next time of ball is the orange ball, this ball bounces quickly again just before it reaches the centre. Try hitting it three times in a row.
6. Nice Job!
7. The final ball is the green ball. This ball has a slightly different timing to the others. Try hitting it three times in a row.
8. Nice Job!
9. Now it's time for the real thing, good luck!

A.2.3 Both

1. Three types of Ball will enter from the left of the screen and bounce towards the middle in time with the music. The beat is indicated by the pulsing circle.

2. Press the space bar to move the paddle when the ball reaches the middle of the screen in time with the music.
3. Let's try this with the first, yellow ball. Hit it with the paddle three times in a row to continue to the next part of the tutorial.
4. Nice Job!
5. The next time of ball is the orange ball, this ball bounces quickly again just before it reaches the centre. Try hitting it three times in a row.
6. Nice Job!
7. The final ball is the green ball. This ball has a slightly different timing to the others. Try hitting it three times in a row.
8. Nice Job!
9. Now it's time for the real thing, good luck!

A.3 Logging

An example log file is included (figure A.3) below to give a sense as to the data that is recorded from a single participant.

```
24/04/2018 14:14:04:1842
Gamemode: AudioOnly

—Tutorial 1—
Start time: 24/04/2018 14:14:35:7601
#
Spawn Beat Number, Ball Type, Timing, InputHit
13,SINGLEBALL,-0.773333333333143,False,
21,SINGLEBALL,-0.837333333333050,False,
...
45,SINGLEBALL,-0.0106666666661123,True,
Total Balls:5
Total Hit: 3
Percentage Hit: 60%
#
End Time: 24/04/2018 14:14:46:3192
Time taken: 00:00:10.5591599

—Tutorial 2—
...

—Tutorial 3—
...

—Game Start—
Start time: 24/04/2018 14:15:33:8293
#
Spawn Beat Number, Ball Type, Timing, InputHit
13,SINGLEBALL,0.0748886666668716,True,
21,SINGLEBALL,-0.0959222999997194,True,
...
565,SINGLEBALL,-0.037334443326842,True,
575,TRIPLEBALL,0.0266321666732931,True,
Total Balls:59
Total Hit: 47
Percentage Hit: 80.0000000000000%
#
End Time: 24/04/2018 14:17:38:3875
Time taken: 00:02:04.5582042

Overall End Time: 24/04/2018 14:17:38:3910
Time taken: 00:03:35.0731239
```

Table A.3: Example log file including the three practice sections in the tutorial followed by the main gameplay.

B The Study

B.1 Study Documents

B.1.1 Audio-Only Information Sheet

Information Sheet for Immersion Project Study

Huw Talliss
Supervisor: Dr Paul Cairns
University of York

Overview

In this study, you will be asked to play an audio-only game then to fill in a questionnaire about you experience of playing the game. The game is an audio-only rhythm game based on the style of gameplay from the Rhythm Heaven series of games. There will be a short in-game tutorial before the game starts to help you to understand the game. You are not being assessed on your performance in the game but do try to play as best you can. The study should not take more than 10 minutes.

The game

The game is a rhythm game where conceptually balls bounce in time with the music, towards a paddle which you control. Your goal is to trigger the paddle as to hit the balls in time with the music.

All in-game interactions are performed through pressing the spacebar which, during gameplay will trigger the paddle to move, hitting any balls that have reached the paddle.

There are various audio cues to help you to play the game which will be explained to you during the tutorial

You will wear headphones while playing so that you may hear the game audio whilst playing and to help remove external distractions.

Questionnaire

After playing, you will be asked to fill out an online questionnaire of around 30 questions about your playing experience.

Questions

If you have any questions about the game or the study please ask them now but once the study has started, please keep your questions until the debrief at the end.

Withdrawing

You are free to withdraw from this study at any point without giving a reason. You are free to request for any data that you have supplied to be withdrawn/destroyed. You are free to refuse to answer or respond to any question that is asked of you.

Data

Any personal information you provide will be stored securely and separately from the data gathered from today's study and will only be used so that the data you have supplied may be destroyed upon request.

The study will be written up as a report for my final-year MEng Computer Science project. It may also be published in academic outlets such as journals, conferences or research books. In any case, the data will only be presented in a summary form and you will not be directly identifiable in any way.

B.1.2 Video-Only Information Sheet

Information Sheet for Immersion Project Study

Huw Talliss
Supervisor: Dr Paul Cairns
University of York

Overview

In this study, you will be asked to play a game with no audio then to fill in a questionnaire about you experience of playing the game. The game is a timing-based game based on the style of gameplay from the Rhythm Heaven series of games. There will be a short in-game tutorial before the game starts to help you to understand the game. You are not being assessed on your performance in the game but do try to play as best you can. The study should not take more than 10 minutes.

The game

The game is a timing-based game where balls bounce from the left of the screen, towards a paddle in centre-screen which you control. Your goal is to trigger the paddle as to hit the balls as they reach the centre of the screen.

All in-game interactions are performed through pressing the spacebar which, during gameplay will trigger the paddle to move and during the tutorials will advance the text boxes.

You will wear headphones while playing to help remove external distractions.

Questionnaire

After playing, you will be asked to fill out an online questionnaire of around 30 questions about your playing experience.

Questions

If you have any questions about the game or the study please ask them now but once the study has started, please keep your questions until the debrief at the end.

Withdrawing

You are free to withdraw from this study at any point without giving a reason. You are free to request for any data that you have supplied to be withdrawn/destroyed. You are free to refuse to answer or respond to any question that is asked of you.

Data

Any personal information you provide will be stored securely and separately from the data gathered from today's study and will only be used so that the data you have supplied may be destroyed upon request.

The study will be written up as a report for my final-year MEng Computer Science project. It may also be published in academic outlets such as journals, conferences or research books. In any case, the data will only be presented in a summary form and you will not be directly identifiable in any way.

B.1.3 Both Information Sheet

Information Sheet for Immersion Project Study

Huw Talliss
Supervisor: Dr Paul Cairns
University of York

Overview

In this study, you will be asked to play a game then to fill in a questionnaire about you experience of playing the game. The game is a rhythm game based on the style of gameplay from the Rhythm Heaven series of games. There will be a short in-game tutorial before the game starts to help you to understand the game. You are not being assessed on your performance in the game but do try to play as best you can. The study should not take more than 10 minutes.

The game

The game is a rhythm game where balls bounce from the left of the screen in time with the music, towards a paddle which you control. Your goal is to trigger the paddle as to hit the balls in time with the music, as they reach the centre of the screen.

All in-game interactions are performed through pressing the spacebar which, during gameplay will trigger the paddle to move and during the tutorials will advance the text boxes.

You will wear headphones while playing so that you may hear the game audio whilst playing and to help remove external distractions.

Questionnaire

After playing, you will be asked to fill out an online questionnaire of around 30 questions about your playing experience.

Questions

If you have any questions about the game or the study please ask them now but once the study has started, please keep your questions until the debrief at the end.

Withdrawing

You are free to withdraw from this study at any point without giving a reason. You are free to request for any data that you have supplied to be withdrawn/destroyed. You are free to refuse to answer or respond to any question that is asked of you.

Data

Any personal information you provide will be stored securely and separately from the data gathered from today's study and will only be used so that the data you have supplied may be destroyed upon request.

The study will be written up as a report for my final-year MEng Computer Science project. It may also be published in academic outlets such as journals, conferences or research books. In any case, the data will only be presented in a summary form and you will not be directly identifiable in any way.

B.1.4 Informed Consent Form

Informed Consent for Immersion Project Study

I, the undersigned, confirm that (please tick box as appropriate):

1. I have read and understood the information provided on the information sheet. ☐
2. I have been given the opportunity to ask questions about the study and my participation in the study. ☐
3. I voluntarily agree to participate in the study ☐
4. I can withdraw at any time without giving a reason and there is no penalty for withdrawing. ☐
5. The use of the data for research and publications has been explained to me. ☐
6. The confidentiality of the data has been explained, in particular that it will all be anonymised and I could not be identified by the data. ☐

Name:

Signature:

Name of researcher: Huw Talliss

Signature of researcher:

Date:

B.1.5 Experimenter's Notes

Experimenter's Notes for Immersion Project Study

Things to do

1. Ensure the screen is at full brightness and the volume is set to 16.
2. Ensure the game is running on "low" mode
3. Select the correct mode for the game and leave the game at the start screen.
4. Ensure they have read the information sheet for the correct study variation.
5. Ask them if they have any questions.
6. Ask them to sign the informed consent form.
7. Ask them to fill in the demographic survey.
   Note down the time they did this.
   `https://york.qualtrics.com/jfe/form/SV_2f40ZokkAqD9MkB`
8. Ask them if they have any questions.
9. Ask them to put on the headphones and ask them to click start when they are ready.
10. Once they have finished playing, ask them to fill out the immersion questionnaire at:
    `https://york.qualtrics.com/jfe/form/SV_81dZjsvwLuWO8Lj`
11. Debrief the participant.

1 Debrief

Ask if they have any questions.

Otherwise explain the study:

"This study is looking into the effect of audio on immersion in games. I'm running three different types: audio-only, video-only and both. The questionnaire you just filled in will be particularly helpful as well as some gameplay statistics gathered by when you played the game.

If you could not tell anyone about the details of the study as I'm still getting participants, that would be appreciated."

B.2 Study Questionnaires

B.2.1 Demographic Survey

Video Game Habits and Preferences

How often do you play video games on average?

- ○ Every day
- ○ Multiple times a week
- ○ Multiple times a month
- ○ Less than once a month

What platform do you regularly play video games on (Select all that apply)

- ☐ PC/Laptop
- ☐ Console
- ☐ Tablet
- ☐ Mobile
- ☐ Other (Please specify) ____________

What genres of video game do you enjoy the most?

____

What past experience do you have with rhythm-based games? (e.g. Guitar Hero, Tap Tap etc.)

- ○ None
- ○ Some
- ○ A Lot

Musical Experience

Do you regularly play a musical instrument?

- Yes
- No

What instrument(s) do you regularly play?

How would you rate your overall musical ability?

- Beginner
- Intermediate
- Advanced

Demographics

What is your age?

- 18-22
- 23-27
- 27-32
- 33 or older
- Prefer not to say

What is your gender?

- Male
- Female
- Other
- N/A
- Prefer not to say

B.2.2 IEQ Questionnaire

Please answer the following questions by selecting the option that best reflects your experience whilst playing the game.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree	No Response
The game had my full attention	O	O	O	O	O	O
I felt focused on the game	O	O	O	O	O	O
I put effort into playing the game	O	O	O	O	O	O
I tried my best	O	O	O	O	O	O
I lost track of time	O	O	O	O	O	O
I felt consciously aware of being in the real world whilst playing	O	O	O	O	O	O
I forgot about my everyday concerns	O	O	O	O	O	O

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree	No Response
I was very much aware of myself in my surroundings	O	O	O	O	O	O
I noticed events taking place around me	O	O	O	O	O	O
I felt the urge to stop playing and see what was happening around me	O	O	O	O	O	O
I felt like I was interacting with the game environment	O	O	O	O	O	O
I felt that I was separated from the real-world environment	O	O	O	O	O	O
The game was something that I was experiencing, rather than just doing	O	O	O	O	O	O
The sense of being in the game environment was stronger than the sense of being in the real world	O	O	O	O	O	O

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree	No Response
I found myself so involved that I was unaware I was using controls	O	O	O	O	O	O
I moved through the game according to my own will	O	O	O	O	O	O
I found the game challenging	O	O	O	O	O	O
There were times in the game in which I just wanted to give up	O	O	O	O	O	O
I felt motivated when playing the game	O	O	O	O	O	O
I found the game easy	O	O	O	O	O	O
I felt that I was making progress towards the end of the game	O	O	O	O	O	O

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree	No Response
I performed well in the game	O	O	O	O	O	O
I felt emotionally attached to the game	O	O	O	O	O	O
I was interested in seeing how the game's events would progress	O	O	O	O	O	O
I wanted to "win" the game	O	O	O	O	O	O
I felt in suspense about whether or not I would do well in the game	O	O	O	O	O	O
I found myself so involved that I wanted to speak to the game directly	O	O	O	O	O	O
I enjoyed the graphics and the imagery	O	O	O	O	O	O

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree	No Response
I enjoy playing the game	O	O	O	O	O	O
When I stop playing, I am disappointed that the game is over	O	O	O	O	O	O
I would like to play the game again	O	O	O	O	O	O

C Statement of Ethics

Do no harm - During the course of the experiment no participant was put in any physical danger, asked to do something illegal or asked to do something that is against their best interests. All questions participants had pertaining to the experiment were answered (any questions that could have affected the results of the experiment were answered at the end of the experiment).

Informed consent - All participants were given information about the purpose and contents of the experiment (including what they would be asked to do), to read before the experiment began. Participants were able to asks questions before the experiment began and were told that they could withdraw at any time and refuse to answer any question without giving a reason. Participants were asked to read and sign an informed consent form before taking part in the experiment.

Confidentiality of data - All information collected from participants was kept confidential in an anonymised format. Participants were informed of this before the experiment. Only my supervisor and I have access to the data in non-summary form.