Hey, listen!

An Exploration of Game Music as Provider of Information

A **Master's Thesis** by
Jacob Lynggaard Olsen & Mads Maretty Sønderup

1

Abstract

Games as an art form continuously evolves and reaches new heights. Game music plays an important part in most modern game experiences, often responsible for setting the mood of a scene, or providing clues about what kind of gameplay is expected (eg. combat music). Especially for game experiences that are linear or hand-crafted, it is often desired to guide the players' choices about where to go and what to do, but this seldom happens through the use of music.

This thesis explores the qualities of music as a medium for information, investigating whether music can effectively be used to guide player decisions in a real game scenario, with the following problem statement: "How do players perform in a game scenario when information necessary to successfully progress is provided solely through associative music?"

Through a highly iterative approach with multiple production and test phases, a game was developed to act as an industry-grade testing platform for the problem statement. A total of 8 tests were conducted, of which 5 were preliminary and guided the continuous production cycle of the game product. The final iteration was tested over 3 separate tests, of which one was published online to the world, another was conducted locally - both with control groups. Lastly, a bonus test was conducted to check for differences in performance when player were explicitly told to listen for the music - which is otherwise besides the motivation for the study.

Unfortunately, no significant performance improvements was found between the groups that received musical hints, and those who didn't, indicating that players are simply not used to receiving valuable information through music.

Contents

2 Introduction

Making games is hard. Making good games is even harder. Games as an art form continuously evolves and reaches new heights - only barely limited by technology. They are no longer just simple arcade games like space invaders, but can be deep, well-crafted cinematic stories, or vast open worlds to explore. Game music plays an important part in most of these cases, often responsible for setting the mood of a scene, or providing clues about what kind of gameplay is expected through, for instance, dramatic music for fighting scenes, and calm music when there is no imminent danger (Lynggaard Olsen et al., 2015).

Especially for stories and experiences that are hand-crafted (ie. not procedurally generated) it is often desired to guide the players' choices about where to go and what to do, so as to have the player remain in a state of *Flow* (Csikszentmihalyi, 1991). A vast array of techniques exist for this, including clever lighting: glowing exit signs or spotlights above the door to the next level; brightly colored objects in the game world that naturally draws the attention of the player (Mirror's Edge, (DICE, 2008)); or a trail of collectibles leading the player to the next interesting place. Surprisingly, audio is seldom used to inform the players decision making processes. It is most often found in a more supporting, rather than informing, role.

There are of course exceptions to this claim. Avid players of games like Counter Strike (Universal, 1999) prioritize good quality headphones with surround sound drivers in order to increase their spatial awareness. In this way, it is possible to hear the footsteps of nearby allies or enemies even when they are not rendered on the screen, giving the players an edge in their decision making process. While there are certainly more exceptions, it can however be noted that these usually regard spatial reasoning through sound effects like footsteps, or other diegetic sources from within the game world. It is rarely seen that the music is used for anything other than signaling the current state of the game (e.g. dramatic combat music signals to the player that there are aggressive enemies nearby), placing music in a merely supportive role.

As described by Jason Graves "the popularity of music for games increase along with our capabilities as composers" (Hoover, 2009), but why is it that music is almost never used as a source of information to guide the player? This thesis explores the qualities of music as a medium for information, and aims to investigate whether it can effectively be used to guide player decisions in a real game scenario.

3

Research

This chapter briefly introduces game music as a field of expertise, while exploring the current state of the art of designing music in games. The affective and associative qualities of music are then investigated, leading to the problem statement. In the end, design requirements are set up for the following experiments that seek to answer the problem statement.

3.0.1 Powerful Qualities of Music

Soundtracks for both film and games are widely popular (Bribitzer-Stull, 2015; Phillips, 2014), most likely because of the subtle, yet powerful qualities that music in these media possess. It is safe to assume that most people have experienced the 'goosebumps' as a result of being emotionally touched by music; recognized a company jingle; imagined themselves on an exotic beach when hearing exotic and relaxed music, etc.

Wingstedt (2004) presents six categories of musical functionality:

- **Emotive Class**: emotionalize content and acting
- **Informative Class:** communication of meaning; communication of values; establishing recognition
- **Depictive Class:** describing settings; describing physical activity
- **Guiding Class:** attention guidance; mask (out) unwanted or weak elements
- **Temporal Class:** provide continuity; define structure and form
- **Rhetorical Class:** comment; make statements; judge

These categories attempt to collect the qualities or functions that music can have and act on in media. Although originally the list was made for film music, Berndt and Hartmann (2008) argue that they can be transferred to interactive media such as games. Berndt and Hartmann (2008) furthermore state that most of the classes are widely underrepresented in game music, where it is most often seen that music acts as a simple background element to the mood, or increasing the temporal perception of action scenes, although sometimes use of the *informative* functions can be seen. Jørgensen (2006) found that especially the *guiding* role of music can prove helpful in games, acting as a secondary medium of information to not clutter the screen with even more visuals, while Berndt and Hartmann (2008) state that emotive and informative musics can influence the player's decisions by accenting some associations and masking out others. In relation to the initial motivation for this study, the informative and temporal classes seem highly relevant. Green (2010) argues that when using music to express meaning, the audience becomes less questioning. It would appear that music has a way of bypassing the conscious processes of the listener. Eladhari et al. (2006) explains that trying to convince something specific through music can be challenging, as what the composer intended for the music might not come through unaltered to the listener, making it insufficient as a channel for very specific or detailed information. Musical meaning seems to sustain itself in ambiguity (Cross, 2005).

3.0.1.1 Subconcious Influence

While most visual and more direct auditory stimuli (like sound effects or dialogue) tends to be perceived more consciously, music is often perceived much more subconsciously (Berndt and Hartmann, 2008). While it might at first seem like one should then take care to only mediate important information through the more direct channels rather than channels that are perceived on a more subconscious level, the indirect perception can in truth be a strength to the dramaturgic effects of music. Berndt and Hartmann (2008) state the following:

> *"The more subconsciously music acts, the more it can condition the audience in a desired manner and stimulate their perceptual direction"*

This effect is quite interesting, as it becomes possible to bypass players' active thinking with subconscious influence. Kungel (2008) explains that this subconscious perception of music directly accesses the limbic system, which is why music can have such an effect on the emotions of the listeners.

Music and affect is an extensive topic in itself, which lies well beyond the scope of this study, however, it can be noted that there is a general consensus between researchers that music can indeed directly influence emotional response of listeners (Juslin and Västfjäll, 2008; Aljanaki et al., 2016; Gabrielsson, 2002), making it a powerful tool for storytelling in film and games. While most individuals are rather proficient at telling whether a piece of music sounds *sad* or *happy* (often related to minor and major mode music), Gomez and Danuser (2007) has conducted research investigating the influences of several musical features down to a more detailed level. It is also relevant to note that researchers (and practitioners) usually distinguish between *emotions* and *mood* (Batson et al., 1992; Scirea et al., 2014; Beedie et al., 2005). The main difference being that emotions are more focused and defined, where moods are generally more unfocused or blurred while lasting longer (Scirea et al., 2014).

3.0.2 Music in other Media

A medium that is well known for its use of music is the cinematic medium. More specifically, films are often praised for their clever use of music to both support moods and pacing, as well as providing information about non-visible elements like the intentions of characters on screen, or to communicate the "'unspoken' thoughts of a character or unseen implications of a situation" (Green, 2010).

3.0.2.1 Leitmotifs

A particularly useful tool for media composers is known as a *leitmotif* (a *leading motive*). Leitmotifs are musical ideas that are linked to "represent or symbolize a person, object, place, idea, state of mind, supernatural force or any other ingredient in a dramatic work" (Paulus, 2000). A common example of leitmotifs is 'The Imperial March', or more often referred to as 'Darth Vader's theme' from the massively popular *Star Wars* franchise (Lucas, 1977). This is arguably one of the most popular and commonly known leitmotifs, easily recognized by many. At this point it can be speculated that even people who have not seen the movies might recognize the theme and connect it to Darth Vader, since Darth Vader is a massive pop-culture phenomenon.

Another example is Ellie's Theme (Giacchino, 2009) from Pixar's 'UP', which is introduced in the opening scene as a happy theme that represents Carl and Ellies blossoming relationship. The theme is then used to great effect when Ellie becomes worryingly ill, where the musical theme has suddenly transformed emotional qualities into something sad or depressing. Bribitzer-Stull (2015) explains that such associative musical constructs (like leitmotifs, idée fixes, etc.) can collectively better be described as 'associative themes', where leitmotifs by themselves should be defined as a "recurrent musical idea which has been invested by its composer with semantic content".

However, leitmotifs have a prerequisite: a successful connection that firmly establishes the link between the music and that which it is a theme of (henceforth referred to as the *element*). In other words, the very first time the theme is heard in connection with the element, there is no immediate association, and some form of repetition is necessary to fully establish the connection between music and element. The associative qualities of leitmotifs may not always be as conscious as it probably is to many with The Imperial March. Sometimes it is used for subconscious or emotional effect where if the listener would later listen to the theme, it might not have clicked, whereas while watching, the brain can subconsciously register that this theme is recognizable. According to Bribitzer-Stull (2015), the leitmotif technique is useful for recalling things that are dimly remembered, starting a process in the listener of seeing what sense can be made in this new context. Therefore, by listening to a already known leitmotif, despite it's musical context might be slightly different, it seems possible for the listener not only to recall familiarity but also connect it to the known leitmotif.

However, Kramer (2002) mentions that a central problem with the associative qualities of leitmotifs, is that they operates "between autonomy and contingency" and are therefore not necessarily a intelligible message but rather chunks of ideas. The composer can therefore only show the idea to a listener in the hopes of that the listener will understand it. Cook and Everist (2001) even argues that music cannot have a unmediated meaningful effect on the listener, meaning with the leitmotif alone it is not necessarily possible to plant an idea in ones head. Despite these difficulties of conveying any information or associations with leitmotifs, Bribitzer-Stull (2015) if the composer has a good relation to the social praxis and musical culture there are often "piece-specific sweet spot" that will make the associative quality of leitmotifs possible.

One approach to establish musical associations is according to Bribitzer-Stull (2015) by repeating the leitmotif, so that the listener slowly associates what is happening with the music. Another strategy is something called operative overture, which in film is often described as the movement of when preparing to see the film and as such sharpening the focus in order to be more focused on listening. Lastly, delivering the theme at significant and memorable moments in the experience should also help establishing the association.

Bribitzer-Stull (2015) also mentions that association is meaningless without recollection, which refers to not only when remembering or associating a theme to a previous experience of the theme, but also reminding the listener of the elapsed time, and therefore manifests itself at a specific time in memory.

3.0.2.2 Speeding up the Association Phase

For bigger games with countless hours of playability, this is less of a problem, as the players would most likely spend a lot of hours in the game, thus having time to establish connections and associations with musical themes over time. However, smaller productions like the one of the present study has to consider methods that can speed up the association phase, as it is both impossible to produce such expansive experiences in such a short time, as well as to expect test participants to participate in a study where the testing phase can take several hours.

3.0.3 Obvious Instrumentation

Various games have over the years been using different techniques to accomplish association and also without using directly leitmotifs. World of Warcraft (Entertainment, 2004) uses human cultural preferences to characterize their zones, factions and cities, by adding textures and special instruments to their themes. For instance, the peaceful nature of night elves is characterized in their theme with a harp and flute along with several other classical orchestral instruments. Another example is the use of sitar and middle eastern instruments in the game Uncharted 3 (Dog, 2011) when navigating the desert. By using musical instruments that are already strongly connected to a certain culture, region or behaviour, the player will already have the foundations of understanding of its cultural association, thus aiding the association phase, as some elements are already given.

3.0.4 Musical Foreshadowing

In addition to the powerful associative qualities of leitmotifs, music can also be used for *foreshadowing*. Musical foreshadowing is by Boltz et al. (1991) described as the following:

> *"Foreshadowing prepares an audience for a critical, upcoming event by presenting music during a preceding scene where the future course of activities is relatively uncertain. By doing so, the audience is drawn into the film and invited to generate expectancies about what will happen next."*

Scirea et al. (2014) found in their research that musical cues for narrative foreshadowing results in an increased percieved consistency between the music and the game. They also found that false foreshadowing (eg. playing happy music when something sad is about to happen) enhanced the self-assessed enjoyment of the players, which is further backed by Boltz et al.

(1991) and Tan et al. (2007) who also claim increased memorability after the experience.

Foreshadowing is then a useful tool for providing narrative clues, causing the listeners to establish expectancies for what is soon to happen. As such, it can reasonably be argued that the musical foreshadowing brings some informative qualities as well - whether the information turns out to be true or false is then up to time to tell. Bribitzer-Stull (2015) explain that the purpose of foreshadowing or presentiment is to act as a sort of *"psychological or emotional preparation for what is to come, thus making the actual event a fulfillment of this preparation*, which again relates to narratives. However, the aim of this study is to successfully guide the players decisions via music, meaning that foreshadowing (at least in the narrative sense) becomes slightly irrelevant. Instead, inspiration can be taken in the sense that information can be provided by the music in advance, thus hopefully participating in guiding the players decisions.

3.1 How It All Fits Together

While music and its possibilities is the focal point of this study, it cannot be denied that in most multimodal experiences, visual stimuli is in front of the direct attention of the audience. Bribitzer-Stull (2015) lists in his book *Understanding the Leitmotif - From Wagner to Hollywood Film Music* a series of psychological effects that music can have when combined with sounds and visuals, based on work by Cohen (1990). This list is relevant to present as it can help summing up why it is reasonable to assume that music can be used as a channel of information to players. To sum up the contents of the list that are relevant to this study, a condensed and somewhat paraphrased version is presented here:

1. **Meaning and Associationism**: Things happen at the same time and, because of connotative overlap (as in metaphor), one can evoke the other even when it is absent (i.e., music can recall emotion of previous scene it was linked to).

2. **Structure and Organization**: When both auditory and visual stimuli exhibit similar "structures", as when they both change at the same time, facilitating information-processing tasks.

3. **Memory and Awareness**: Musical soundtracks are popular in part because the evoke memories of the movie, although such memories are often visual more than auditory.

While the list that Bribitzer-Stull (2015) presented included **Experimental Aesthetics**, it has been excluded from the list presented in this section, as it mostly regards music overshadowing content of information (as seen in educational documentaries) (Bribitzer-Stull, 2015).

It can be noted that the three items on the list all revolve around music strengthening the overall connection in between the experience. While Cohen (1990) found that people are much more likely to recall if a scene was included in a movie or not, rather than telling if a specific music clip was part of a film or not, the claim in point 1. is very interesting in relation to the current study, as it further supports the idea that the associative qualities of music can be used to effectively make people remember previous experiences - similar to what is expressed in point 3. Point 2. pertains the design suggestion that if the music contextually changes with the setting, associations start forming, which is relevant in context with the associative music strategy.

So it seems that there is a lot of pointers to the idea that associative music might be useful for providing musical clues with the aim of guiding the actions of the player. It was discovered that clever sound design can provide valuable information to the player that might otherwise be confusing to present visually, and since the presented research shows plenty of mentions of the informative and guiding qualities of music, it can be hypothesized that music might be able to fulfill the same role. One of the strengths of music is that it can bypass the conscious processes, and instead directly influence the listener, suggesting that it might not be apparent to the listener if music in fact did influence them in any way. However for the associative qualities of music to work, it is necessary to successfully establish the connection between theme and element, requiring a certain learning phase.

3.1.1 Problem Statement

As a conclusion to the research, this section narrows the initial area of interest down to a more defined scope. Since the presented research suggests that music can in fact be used to influence the decisions of the listener, it is pertinent to set up a scenario to either confirm or reject the hypothesis that music can indeed be used for this purpose. As such, a more defined problem statement can be presented:

> ***"How do players perform in a game scenario when information necessary to successfully progress is provided solely through associative music?"***

In order to test this, it is necessary to also establish a control group, in which the information is not given, as otherwise it would be impossible to tell whether the music was the game element that had an influence on their decisions.

However, it is difficult to discuss *performance* without defining the terms that make a good or bad performance. While some games have points and high scores, other games are merely storytelling experiences, in which "good performance" is harder to define. In relation to this study, since the music will provide information that aids progression, it makes sense to look at performance as how much effort goes into progressing. There are many different ways to measure this, and the exact measures that are taken will be described thoroughly in chapter **6.**

Based on the research, it is hypothesized that the group that receives musical information (musical hints) will perform better than the control group. While it might seem trivial to compare a group that receives information versus a group that does not receive same information, the special nature of this study pertains such a set up. Because of musics usual place as a supportive background element in games, it is hypothesized that players will not expect valuable information to come from this medium. However, because of the potential subconscious influence of music, perhaps the performance of players in the musical hint group will indeed perform better than the control group.

3.1.2 Design Requirements

With a more defined problem statement in place, it is also possible to set up a list of requirements for the design of a test that can seek to answer the problem statement.

- There must be at least two groups:
  - A **musical hint group** that receives information through music, valuable to the progression of the game.
  - A **control group** that plays an identical experience, except for the fact that they will **not** receive the musical hint.
- The test must include training phases for the associative music.
- The test must repeat the musical constructs.

3.2

Technical Research

Games were quite primitive in the birth of the industry. Simple, blocky, yet interactive graphics displayed by the limited technology of those times. Game sound even more so. The engineering side of video games has also undergone significant changes. As video games are constantly increasing to new heights of accomplishments so is the expectations of game audio. In the early days, and still to be observed in many indie game productions, game audio responsibility often ends up "in the hands of a solitary figure wearing many hats" (Kastbauer, 2017). However, there has still been a increasing need for creating diverse and varied soundscapes along with flexible and dynamic music systems.

Video games without any audio is nowadays a rare thing, but it is to many a secondary element of games and is at times seen as just something you wrap on afterwards. However, some studios have produced extraordinary results with audio and music design, and with their achievements they inspire studios that would usually not even expose themselves to the process of audio design. This section presents some of the most notable audio design achievements.

A game known for it's highly detailed approach to sound design is the danish Playdead's game - *Inside*. Their approach to sound is much more focused on the how the systems work together, rather than having a catchy theme. *Martin Stig Andersen* (the Lead Sound Designer) describes their well composed achievements as like playing in a band and "its not only because of the game or the music, it's because of the way they works together" (Andersen, 2016). He also mentions how the game "listens", not to be confused with microphone input, but rather listening to the actions in game and having a nice collaboration between the sound engine and the game engine by with callbacks. One example is the respawn of the player which is not necessarily restarting music and ambiences, but rather executes a death stinger-like sequence, followed by a call from the music to the game engine, that its ready to respawn, so music and game feels within the same space and not just a slap on. Another example is an elevator that needs to be timed with a repetitive explosion, where the "speed of the elevator is actually adjusted to hit the beat" (Andersen, 2016). Other notable mentions is that the boy in Inside also has rhythmic breathing, different breath conditions depending on if you're about to jump or push (Schmid, 2016), and the list goes on.

Overwatch a recent product made by *Blizzard*, is known for having a priority system that handles what sounds are important to the player, by asking questions such as "Who is close by","Who is looking at me", etc. to generate a threat level of what is important for the listener to pay attention to. This way they've accomplished to create a dynamic system mapped to questions and behaviour rather than e.g. simple ducking systems on sounds vaguely classified as "important".

Last of Us, a game developed by Naughty Dog, uses a similar technique to generate direct audio cues on whether the player is in danger of being discovered or seen. When the player is about to get discovered, a low noise rumbling is slowly intensified indicating such and spoiling that you might be in danger, like a sort of spider-man sense.

Threat levels can also be used in a musical context like *Rise of the Tomb Raider* who procedurally generated percussive instrument rhythms depending on enemy AI awareness and player actions (Jack Grillo, 2015). As the music is linked up using MIDI [1] all the system needs is the sample data from the composer and the music can be completely controlled by behaviours and systems instead of being fixed layers, etc.

3.3 Music in Games

Unlike traditional linear media such as films, in video games it is inherently uncertain *when* a particular thing happens that would pertain a change or response in the music. While stories in interactive media can indeed be linear, the interactive nature of games mean that the players are in control of when the story will progress. The player can usually choose not to start the next chain of dialogue, or stay exploring in an area for longer than expected. Furthermore, an interesting phenomenon can be observed, further distinguishing film music from game music. In films, the music is not heard by actors in the film itself (assuming it is non-diegetic music), while in games the player is an actor, which means that even though the game music might be classified as non-diegetic (background music, source of emission is not inside the game world), it can still influence the game world, as the player might choose to act upon what music is playing (Berndt and Hartmann, 2008).

3.3.1 Responsiveness

In games, feedback is crucial. Without feedback or reaction, there is no interaction, and without interaction it is by definition not a game. Similar to how music carefully complements a scene in films, so is that result often desired by game music designers. However, while in films every action and

[1] Musical Instrument Digital Interface: A industry standard digital node format.

event is known with precision down to a single frame, that is not the case for games. This becomes a complicated matter if it is desired for the music to react to the actions of the player with the purpose of providing feedback, since music is a medium that needs time to evolve and transition from one thing to another.

Berndt and Theisel (2008) support this claim:

> "In the interactive context latency is another matter of importance and a fundamental conflict arises: musical change processes need a certain time (the longer the smoother). On the other hand, big latencies are disadvantageous for interactivity, hence, unwanted."

While Berndt and Theisel (2008) claim that the longer time a musical segment has to transition, the smoother it will be (a claim that can definitely be supported), it is arguably possible to achieve smooth transitions through other means, and without spending too long.

3.3.2 Game Music Methods

The inherent interactivity of games requires music for this medium to be designed with this uncertainty and highly variable timeline in mind. Add in that the musical reactions need to happen without too long latency, and it is quickly apparent that constructing dynamic music for games requires a lot of engineering. While there are many methods to aid composers designing music for such dynamic environments, two of the most common basic strategies are *horizontal re-sequencing* and *vertical remixing* (Phillips, 2014).

3.3.2.1 Horizontal Re-sequencing

Horizontal re-sequencing is the concept of rearranging modular pieces of music dynamically, similar to the Musical Dice Game - a game where rolls of dice determines which measure of music is played in which order, creating new music pieces.

In practice, this technique requires the composer to split up music pieces in smaller segments that can be dynamically rearranged and interchanged, often backed by a logical system that takes musical decisions based on information from the game. The illustration seen in figure 3.1 attempts to visualize how a music piece is split up into several parts, each with several variations that can be selected randomly.

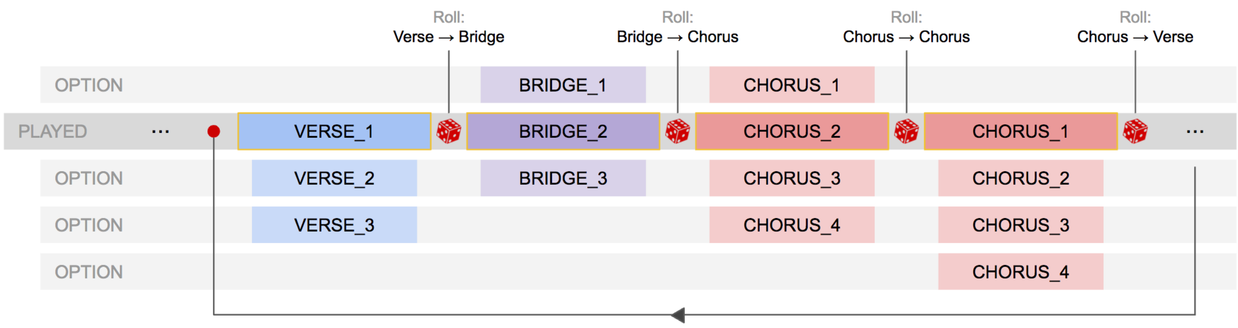


Figure 3.1: Illustration of the concept of Horizontal Resequencing.

It is typical that the music does not have to follow a conventional structure such as the one depicted in figure 3.1. Instead, the music system might select to go from bridge to verse, or verse directly to chorus, or even repeat the verse another time. Horizontal re-sequencing is often seen on a higher level as well - following finite state machine (FSM) logic along with the game code. For instance, the system might switch states from *Exploration* to *Combat*, while in the combat state, there might be a sub-FSM that switches between, for instance, phases of a boss-fight.

3.3.2.2 Vertical Remixing

Vertical remixing is the concept of adjusting the levels of each part of a musical track at runtime. This concept is illustrated in figure 3.2, where a theoretical music track is exported in several layers, consisting of drums, guitar, choir and synths.

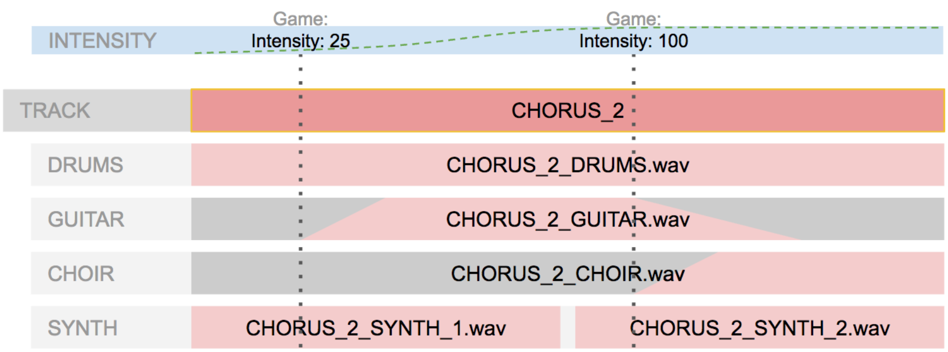


Figure 3.2: Illustration of the concept of Vertical Remixing.

With vertical remixing, the mix of the track is usually adjusted during gameplay, based on parameters controlled by game logic. A common example is an *Intensity* parameter that reflects how intense the situation of the player currently is. As the intensity rises, the music system can then enable or disable, as well as adjust the volume of individual layers in each music piece, in order to suit the current events in the game.

3.3.2.3 The Best of Both Worlds

In most practical scenarios, both horizontal and vertical remixing is used in combination, or to different extents. Both techniques have each their pros and cons, and allow different types of controls and responsiveness in context with the game. Both allow for extensive variability to reduce repetitiveness - and issue that stems from resource constraints. While a scene in a game might be designed to take around 15 minutes to complete, the player might be struggling with a certain puzzle, or just explore every little detail of the world, causing his/her playthrough to last 45 minutes instead. As such, music is usually composed with either both or one of these methods in mind, in order to allow much more variability than otherwise. However, since it is impossible, or at the very least too resource intensive, to produce an infinite amount of music, it is impossible to be exempt from having to loop segments. While there is a general consensus that looping music segments can lead to music that is repetitive, leading to boredom (Eladhari et al., 2006; Livingstone and Brown, 2005), clever usage of both vertical remixing and horizontal resequencing can combat this, adding much more variety than a simple static loop. This is important to diminish listening fatigue, as well as to prevent the music from becoming an obtrusive element of the game experience (Berndt et al., 2006).

3.3.2.4 Other Game Music Methods

Another approach to achieving highly dynamic music in games, and with infinite variability, is through procedural generation. Procedural Content Generation (PCG) has many facets, both useful and limiting. The main advantage is that the music system can reach incredible heights of dynamicity and variability. This flexibility has so far led some studies in the field of dynamic music in games to using PCG. As an example, Scirea et al. (2014) explain in their research about musical foreshadowing of narrative game experiences:

> "[About doing their study without procedural music generation] ... *we would have had to compose a number of different scores, record a performance of them and then insert the resulting music pieces in our game. As you can see, this would have been a time consuming and complicated effort, considering that*

> *the events that trigger the music changes aren't scripted, but triggered by the player.*"

While it is certainly an interesting and perfectly valid approach, one of the main disadvantages with procedurally generated music is arguably that it can still lack the personality and quality of careful and creative human craftsmanship, which is why some still hold on to the human composition element (Berndt et al., 2006), and for the sake of this study, so will we.

3.3.2.5 Industry Tools

As most game development studios that implement dynamic game music value the quality of bespoke music design for their productions, a lot of effort is put into achieving a homemade, yet flexible music system. It is usually necessary to employ certain tools that allow music designers who are not experienced back-end programmers to be able to bring out their visions of music design in the game experience.

While some studios spend resources assigning programmers to sit with music designers and construct bespoke software, it has become increasingly popular to license middle-ware solutions such as FMOD Studio (Fireflight Technologies, 2015), Elias (AB, 2015) or Wwise (*Audiokinetic Wwise*, 2015). Such tools can aid music designers in implementing music in a flexible way that can be specifically tailored for the desired effect in-game. In the case of the current study, and because of collaboration with Audiokinetic, Wwise will be the sound engine of choice.

4 Methods

This section presents the motivations behind the study, while defining the target group and discussing the implications of performing the current study on the chosen platform. It will then describe the two test phases and the general procedures.

4.1 About the Study

This academic research has as a primary goal to end up with results that can be used in the game industry by audio professionals and music designers in order to produce better, smarter and innovative game music systems and implementations, hopefully yielding a better overall experience for the end consumer. As such it is with great emphasis that we perform tests in an environment that is comparable with industry productions, and not merely a 'sterile' isolated academic environment that is deprived of otherwise standard or common game elements. In other words, even though this study focuses on providing information through the medium of music, it is deemed vital to the viability and validity of the results that the problem is investigated in an environment that includes 'the entire package' such as coherent graphics, visual effects, game mechanics and aesthetics that connect the game world as an entirety.

However, it is known from experience with previous research by the authors how high uncertainty and variability there can be in performing tests in such an environment. Furthermore, the scarcity of research on the particular problem of this study pertains an initially explorative approach, which is why most tests will obtain a lot of qualitative data. However, there are certain metrics useful for future analysis that can be captured as numbers and events, allowing (to some extent) quantitative data analysis. All captured data will be thoroughly explained and presented in section 6.0.3.

4.1.1 Target Group

As this study focuses on a problem that relates to design tools useful for general game industry productions, the target group is also shared between this study and the industry: gamers. In other words, we are interested in gathering data from any *gamer* or individual who have played and/or plays games. While there arguably might be notable differences between player-populations of different game genres, the explorative nature of this study means that performances of players from any group are relevant to be able to further target and pinpoint future topics of more in-depth research and exploration.

4.1.2 Game development

Even though games in the end are considered a singular product, behind the scenes it requires extensive planning and work from many widely different fields of expertise. Video games rarely just fulfill one purpose but attempts to please players over a variety of methods like creating interesting movement, fierce combat or inspiring soundscapes, appealing to many different player types. This diverse collection of systems bring along unique challenges, as in the end, most systems usually need to react or converse with one another, optimally reaching what's described by Kastbauer (2017) as "passing values back and forth in an endless stream of dynamic interplay". On those terms, game development requires extensive engineering in several fields, as well as extensive optimization and testing to reach a product that is reliable, performant and most importantly engaging. It is on that basis that game development is tackled using a highly iterative production method, and the game production that serves as a testing platform for this study (WPAG, described in chapter **5**) is no exception.

4.2 General Structure of Tests

The iterative production method pertains an iterative testing structure. As the product reaches its first testable state, the preliminary testing phase can begin, with each test leading to a new iteration of production, until a satisfactory level of quality for the production has been met, after which the final testing phase can begin. Figure 4.1 aims to visualize this method. It should be noted that throughout the initial production phase, constant internal testing is performed to evaluate stability, performance and other measures that affect the enjoyment of the players.

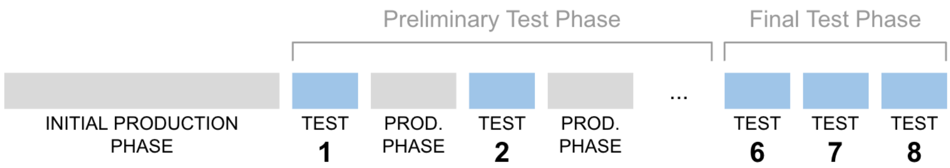


Figure 4.1: General project structure. An iterative approach with preliminary testing between production phases.

4.2.1 Preliminary Test Phase

The purpose of this phase is twofold - to get initial ideas about the feasibility of the current study, as well as guiding further development of the product that acts as the platform for testing the problem statement, aiming to reach a state where the product itself will result in minimum possible biases for the testing.

From each test in the preliminary test phase, a list of new design requirements will be established internally, including a list of bugs, glitches and other issues that need to be fixed before the next iteration. As these lists can assume great lengths, they will not be presented in this report, rather, condensed versions of changes to the game- or test design will be presented where relevant in the test sections of chapter **7. Testing**.

4.2.2 Final Test Phase

The aim of the final test phase is to finally test the problem statement, investigating whether music can successfully be used as the sole source of information in the form of musical hints in an otherwise equal-chance randomized quest.

The final test phase will consist of several tests, presented in the following list:

- **Online test**: This test will not have observation and interview data, however in-game metrics and questionnaire data will still be captured for evaluation. A disadvantage of the online test is that it is impossible to ensure that the participants will follow the given requirements (it is necessary for them to as a minimum wear a headset or use speakers while playing the game), however we try to ensure this through reminders and questions in the post-playthrough questionnaire.

  This test will be subdivided into 2 groups: one with musical hints, and a control group without musical hints.

- **Local test**: This test will be identical to the Online test, except since it is conducted locally we can enforce the participation requirements, as well as observe the participants' playthroughs, and follow through with a questionnaire. This means we will end up with the same data as in the online test, with the addition of observation- and interview notes.

  This test will be subdivided into 2 groups: one with musical hints, and a control group without musical hints.

- **Local Musical Hint Only test (Bonus)**: Although the results of this study targets game developers seeking to use music as an integrated gameplay tool for player guidance, it is interesting to see if there are notable differences in the results of the test, if the participants are directly informed of the fact that music is essential in this study. As such, it has been decided to conduct this bonus test, which will be identical to the Local test, except for the difference of the extra info the participants will receive.

5
Design

As described in section **4. Methods**, the format of this research is highly iterative, meaning that there was a lot of steps to get to the final design , however, this section will center around final build and the steps of getting to the final product will be integrated as part of the testing iteration sections.

5.1 About the Game

As it has been decided to perform this research in an environment that is realistically comparable to actual game industry productions, it is necessary to design a game experience that feels complete to a certain extent. This is where *The Wwise Project Adventure Game* comes to aid.

5.1.1 The Wwise Project Adventure Game

The Wwise Project Adventure Game (henceforth abbreviated WPAG, image in figure 5.1) is the product of the authors' internship at Audiokinetic, creators of the game audio middleware solution Wwise (*Audiokinetic Wwise*, 2015). It is a third-person adventure game where the player plays as a young adventuress that helps out the local *Wwizard*[1] in a small town, cleansing the world of an evil plague corrupting the local flora. A simple, linear storyline takes the player through several different areas, facing different types of enemies while collecting items on a quest-by-quest basis.

[1] *Wwizard* is a wizard. It is a play on words, since *Wwise* (Wave Works Interactive Sound Engine) is spelled with two w's.

Figure 5.1: Promotional image for WPAG, displaying the entire cast of characters and enemies.

The game is made in the widely popular game engine Unity (*Unity3D*, 2015), while audio implementations are handled through the bespoke Wwise/Unity integration package. The graphical style is best described as stylized, or 'cartoonish', or low-poly[2], which is a popular art direction for many indie games.

The goal of WPAG is for users to have a full game-like environment to experiment or learn audio design, audio implementation and programming hooks[3] in game systems to communicate with Wwise.

5.1.2 Game Design

WPAG is quite generic by design, as the focus of the project is on users experimenting with audio implementations using Wwise in connection with Unity. As such, the game features rather basic or conventional mechanics that are well-known from many other games. The core game loop consists mainly of exploring a 3D environment by moving around the player character, attacking enemies, opening doors and collecting items.

[2]Low-poly: although it refers to the low amount of polygons, it is more commonly used to refer to the stylized look of simple/characterised geometry.

[3]*Hook*: A term commonly used by audio professionals to describe a line or segment of code that 'connects' the middleware to the game systems, e.g. playing a sound when a gun is fired.

5.1.3 Controls

The game implements conventional third-person controls that should be familiar to anyone who have played third- or first-person games on a computer before. Keys W, A, S, and D are used for forwards, backwards and sideways movement, while mouse movement moves the camera in spherical coordinates around the focus target (the head of the player character, figure 5.2). The character movement is always relative to the direction the camera is looking in, rather than being relative to the global coordinate system of the game world. This means that regardless of which direction the player is currently looking at, pressing 'W' (forward) will always make the player character move in that direction.

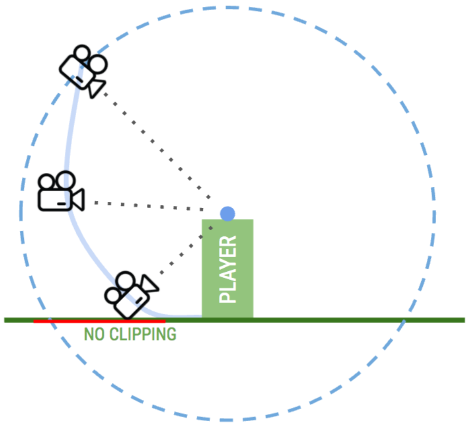


Figure 5.2: 2D illustration of spherical camera coordinates that revolve around the target point. Note that the camera radius decreases as the height of the camera lowers to look upwards in order to provide a more pleasing viewing experience.

For in-game objects like doors and gates, or items that can be collected, pressing 'E' will either unlock, open or pick up the object, depending on what makes sense for the object in question. All objects that can be interacted with will display a yellow outline when the player has moved within interaction range that signifies what pressing 'E' will result in (figure 5.3).

Figure 5.3: When the player is within range of objects that can be interacted with, they display a yellow outline and sometimes a text that describes how to interact and what the interaction will be.

Clicking the left mouse button will cause the player character to swing the equipped weapon, which is useful for deflecting enemy projectiles, or combating vicious monsters. This is the primary method that the player uses to fight the enemies that are scattered throughout the game world.

5.2 Reusing, Re-purposing and Reconstructing

While some elements can be directly reused from WPAG, the special nature of this research does require special development as an extension to WPAG. The elements that can be reused include the following items:

- The player character and its control systems for movement, camera movement and combat.
- Basic enemy Artificial Intelligence (AI).
- 3D models and animations, including *The Adventuress*, *The Wwizard* (both featured in figure 5.1), enemy models, and environmental art assets such as trees and stones.

The original WPAG was designed for a different purpose, so while some of the things listed could be reused, almost everything was built from scratch. Certain productivity tools like task management systems, version control repositories and various custom helper-classes for coding was also used as a basis for starting the new project. Additionally, because of the collaboration with Audiokinetic, any Wwise-related issues could be handled quickly by experts.

5.3 Test Design

To facilitate the test as described in section 6, a special setup was constructed in the game world of WPAG. Since we're dealing with three areas and three enemy types that each have their own unique musical theme, it was decided to make sure the player plays through two training phases before the actual test begins. After completing the training phases, the real test about communicating information solely through music begins, followed by the more direct auditory foreshadowing test.

5.3.1 The Beginning

Before everything begins, the players will find themselves in the Wwizard's hut in the village (figure 5.5), a couple of meters away from just outside where the Wwizard is idling about. In order to go talk to the Wwizard to initiate the experience, it is necessary for the player to use the mouse to move around the camera and locate the exit from the hut, as well as move from inside the hut to the location of the Wwizard (directly visible from the spawn[4] location). In other words, players will have to learn the basic control scheme in order to initiate the experience.

5.3.2 Training Phase 1 (Areas)

After talking to the Wwizard, the player is sent on a quest to gather three objects of importance - each located in different areas. This objects include:

- The Stick of Truth (figure 5.7), located in the forest
- The Potion of Fiery Wrath (figure 5.9), located in the volcanic area
- The Ancient Amulet (figure 5.11), located in the desert

Along with these objects, the player is also asked to gather some *essence* (figure 5.12) from each area.

This design forces the player to spend some time in each area, all the while the music for that area will be playing. This means that we know that

[4] *Spawn*: The initial location/placement of an object.

the player will have been exposed to the music of each area for a certain amount of time.

When all items are gathered, the player is prompted to return the items to the Wwizard, leading to the next quest.

5.3.3 Training Phase 2 (Enemies)

Once the player returns the collected items to the Wwizard, a cinematic cutscene is initiated that shows evil meteors landing in each area that the player has just visited. The Wwizard explains that each meteor contains dangerous enemies of different types, and that the player will need training before taking on the enemies.

The enemy presentation were initially static illustrations of the creatures while listening to *Enemy music themes*, but were changed to a new region, where the player would simply be teleported to. The so-called *Dream Level* would feature the Wwizard presenting each enemy, along with its weaknesses, in a sequential manner (random order between participants - similar to how players could enter each area in their own chosen order). As previously described, there are three types of enemies: One nature-based, one water-based and one fire-based. Each element has its natural counter, visualised in figure (figure 5.15), and enemies based on a certain element can only be killed by a weapon enchanted with the element that counters that type (e.g. Fire-based enemies can only be killed with the water enchantment).

For each enemy, the player needs to grab the correct enchantment as a counter, and successfully defeat a total of three enemies of that type. Meanwhile, the music of that enemy plays continuously.

5.3.4 Test Phase 1 (Information Through Music)

Once the player has returned from the dream world, the Wwizard explains that one of the meteors that previously landed contains a core holding an evil essence that is responsible for the invasion in the first place. He also explains that it is impossible to know which meteor contains the core that holds the essence. The player is then told to go find and retrieve the evil essence to rid the world of the enemies.

This is the crucial point of the study, where the only hint that is given to guide the player is through the music that now plays two themes simultaneously: The theme of the area where the correct core is and the theme of the enemy that is located in that area.

The player hereby ventures to a region of own choice but has no prior knowledge or have received any knowledge about where to go, besides those receiving a hint through the music. As the player kills all enemies in a region, a meteor will break in that same region with a core inside. The core element corresponds to the element type of the creatures, and as the player hits the core with the same enchantment used to combat the enemies, the core vanishes and, if the player has choosen the correct region, the evil essence is reveal. When picking up the essence, the player is informed to return to the wwizard.

5.3.5 Test Phase 2 (Auditory Hint about Container Contents

As the player have completed what seems to be the objective of the game, the Wwizard asks for another favor. The Wwizard tells the player he has a reward, but apparently he has forgot what container he hid it in (while talking to the Wwizard containers have spawned around the entire village area). To get a somewhat decent chance of getting bad containers randomly, the player will be required to find at least 9 coins. However, the containers also contain explosions, which will provide 90% damage to the player. The 90% have specifically been designed to give the player the impression that it's dangerous, but not kill the player unless the player kills another container before regenerating health again.

5.4 The New World

The version of the game that is used for the present study takes place in a relatively small 3D world with a total of 5 zones, of which 1 is exclusively available through following the quest line.

5.4.1 The Regions

There are 4 different types of environment in the game world, split into 4 separate regions connected by long dirt roads (figure 5.4).

Figure 5.4: Orthographic top-down view of the WPAG world map.

The three objective regions (Forest, Desert and Volcanic) are located with approximately equal distance from the village, which is where the player starts the entire experience. This way, one region does not attract the players more than other regions because of shorter distance (gamers have from experience often been observed to follow the path of least resistance), meanwhile giving the music system ample time to transition between region themes and the ambient music track. The contents of each region is conveniently hidden behind clever environmental design in order to not spoil what is going on by just looking in the general direction of the region from for instance the Wwizard house in the village. However, it is always possible to tell what kind of environment is located at the end of each path, as props from each region is placed at the beginning of the road (eg. palm tress by the road to the desert area).

Practically, the three types of environment was decided from the options of conventional characteristics in terms of instrumentation, with the aim of having as few overlaps as possible.

5.4.1.1 The Village

The village (figure 5.5 is where the player spawns. It is also the location of the Wwizard who is the quest-giver in charge of progressing the game, as well as his magnificent Wwizard Tower that is so tall that it acts as a beacon. It is a calm place without danger, as a trigger was set up to ensure that no hostile creatures could ever enter this zone. Perhaps more importantly, this zone acts as a sort of auditory calibration, as the village music is especially ambient and stripped from characteristic instruments, giving a sense of separation from the music themes of the objective regions. The music for each region is further described in section 5.5.2.

Figure 5.5: A screenshot of the village. Not a lot of houses for a village.

5.4.1.2 The Forest

The forest is a lush, serene region garnished by trees and waterfall (figure 5.6. In it, a series of raised platforms connected by wooden bridges, leading to *The Stick of Truth* (figure 5.7, which is a quest item in the first quest.

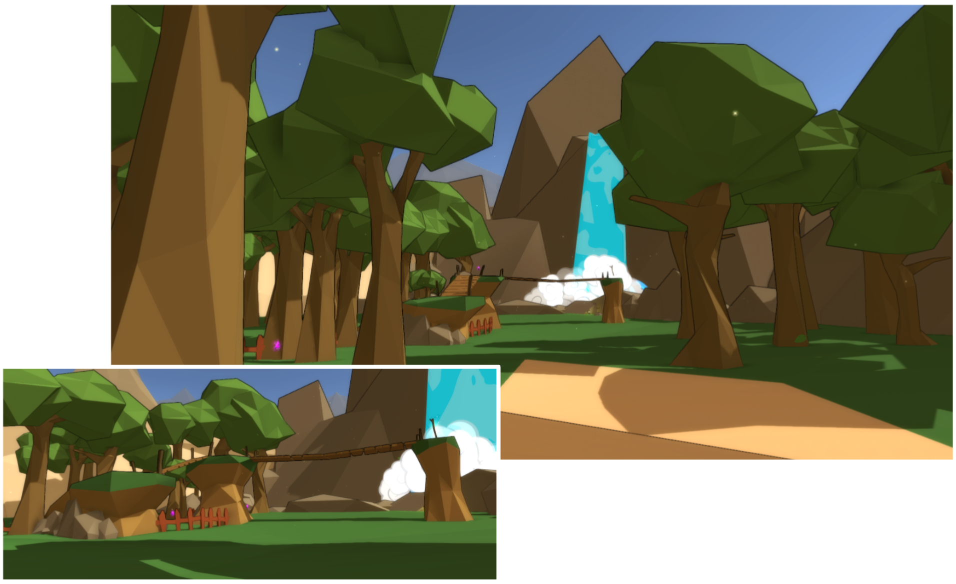


Figure 5.6: Screenshots of the forest.

However, in order to climb the first platform, the player must find a key located near the waterfall. This key can then unlock a gate blocking the path into the deep forest.

Figure 5.7: A screenshot of The Stick of Truth (quest item).

5.4.1.3 The Volcano

The volcanic area (figure 5.8) is designed to be a frightning place. With inspiration from dry, volcanic areas in active erruption, this region looks dangerous with its flowing lava, and perhaps more noticable, flying fiery rock boulders.

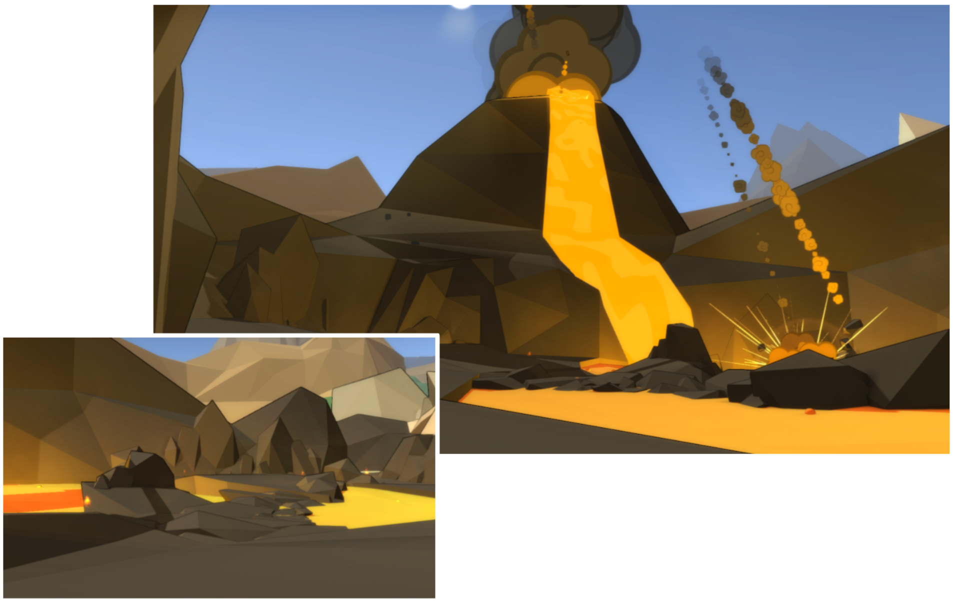


Figure 5.8: Screenshots of the volcanic area.

Beneath the volcano, an entrance is hidden, with a chamber keeping the Potion of Fiery Wrath (figure 5.9) on an old table, which is also an item in the first quest. The key to unlock the door to the chamber is found on a big stone platform in the middle of a lava lake, surrounded by lava, connected by a stone bridge.

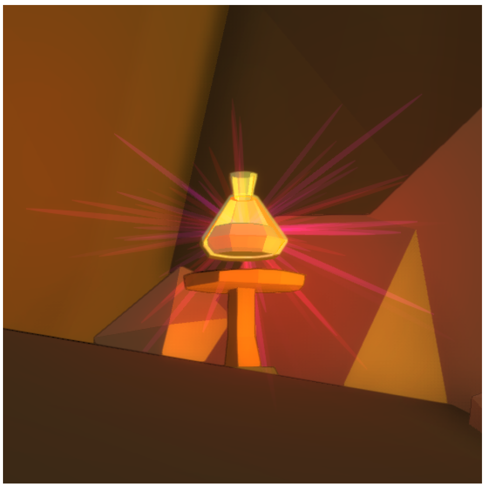


Figure 5.9: A screenshot of The Potion of Fiery Wrath (quest item).

5.4.1.4 The Desert

The desert (figure 5.10) is a bit larger than the other zones, to give the sense of a traversing through a daunting, dry desert. Additionally, a particle system was set up to simulate a mild desert storm, at times making it more difficult to see.

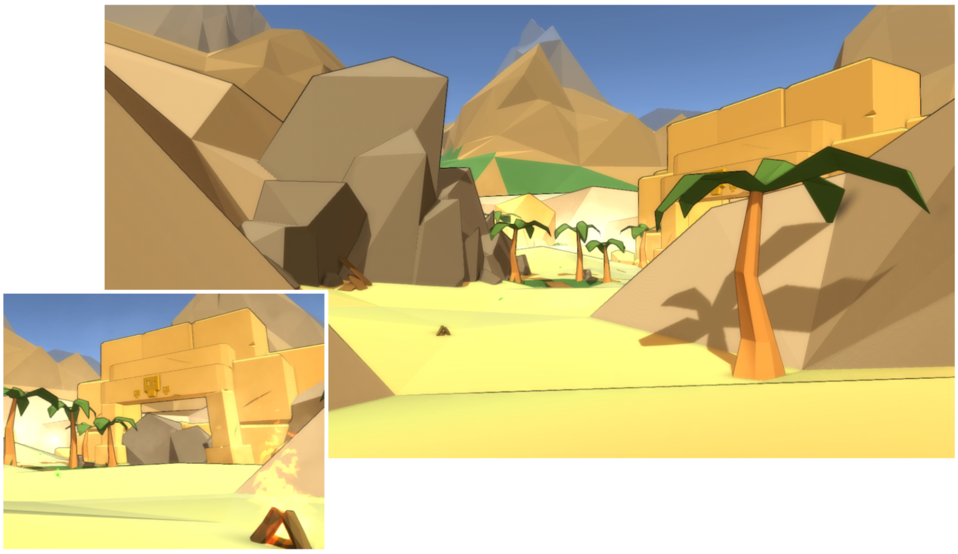


Figure 5.10: Screenshots of the desert.

A huge desert temple is found deep in the desert, hiding the Ancient Amulet (figure 5.11), which is also a quest item in the first quest.

Figure 5.11: A screenshot of the Ancient Amulet (quest item).

5.4.2 The Story

Since this project does not focus on narrative intelligibility, the story is quite simple. Upon first contact with the Wwizard, the player is told that he senses an evil presence is about to arrive, and that they must act quickly. The player is then sent on a mission to gather some essences (figure 5.12) and items (figures 5.7, 5.9 and 5.11) for the Wwizard in preparation.

Figure 5.12: A screenshot of the Ancient Amulet (quest item).

Suddenly 3 meteors drop from the sky, each landing in separate regions. The player is told that the evil monsters have arrived, and that the Wwizard knows how to defeat them. He then teleports him and the player to a dream world where the player is trained in combat against each enemy, using a sword enchanted with each monster's elemental weakness (figure 5.13).

Figure 5.13: A screenshot of the Ancient Amulet (quest item).

After returning from the dream, the player is then sent out to find and retrieve evil essence from one of the cores, which the Wwizard then destroys, cleansing the lands of the evil plague.

5.4.3 Enemies

Each enemy (figure 5.14) was designed to be different in a number of ways. Firstly, each enemy acts differently. The Spitter is stationary, but periodically spits deadly projectiles at the player. The Crawler is slow, dragging itself around in an almost rhythmic fashion. If the Crawler gets close enough to bite, it is quite deadly. Lastly, the Evil Head flies around fearlessly, can be hard to hit, but also doesn't damage as much as the other enemies.

Figure 5.14: The Spitter (left), Crawler (middle) and Evil Head (right).

Each enemy is based on an element; The Spitter is water-based, the Crawler is nature-based and the Evil Head is fire-based. As such, each enemy also has its natural counter (figure 5.15). The system is totally generic, and can be found in many other games, such as Pokémon (Nintendo, 1997).

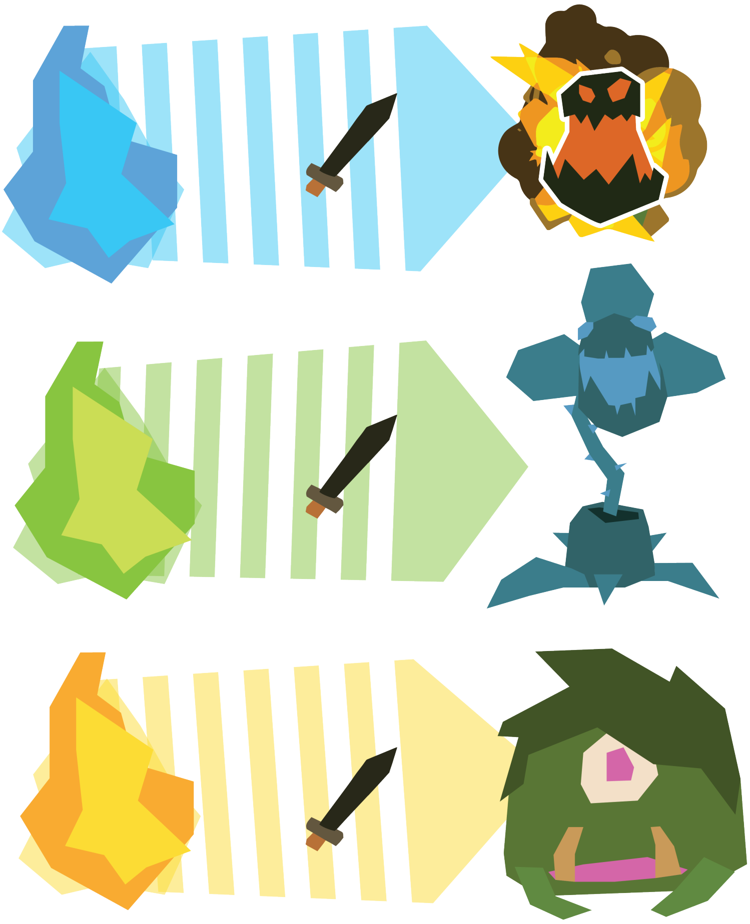


Figure 5.15: A chart illustrating each enemy's weakness.

The basic idea is that the system plays on intuitiveness; ideas that hold true in the real world as well:

- Water beats fire, as water is used to put out fires.
- Fire beats nature, as nature primarily consists of flammable objects.
- Nature beats water, as nature thrives on water.

5.4.4 Miscellaneous

Since the final test is also to be conducted online, some extra precautions were implemented to ensure that the basic test participation requirements were met. Namely, before starting the game, an intro screen (figure 5.16) tells the user to either wear headphones, or use good quality speakers.

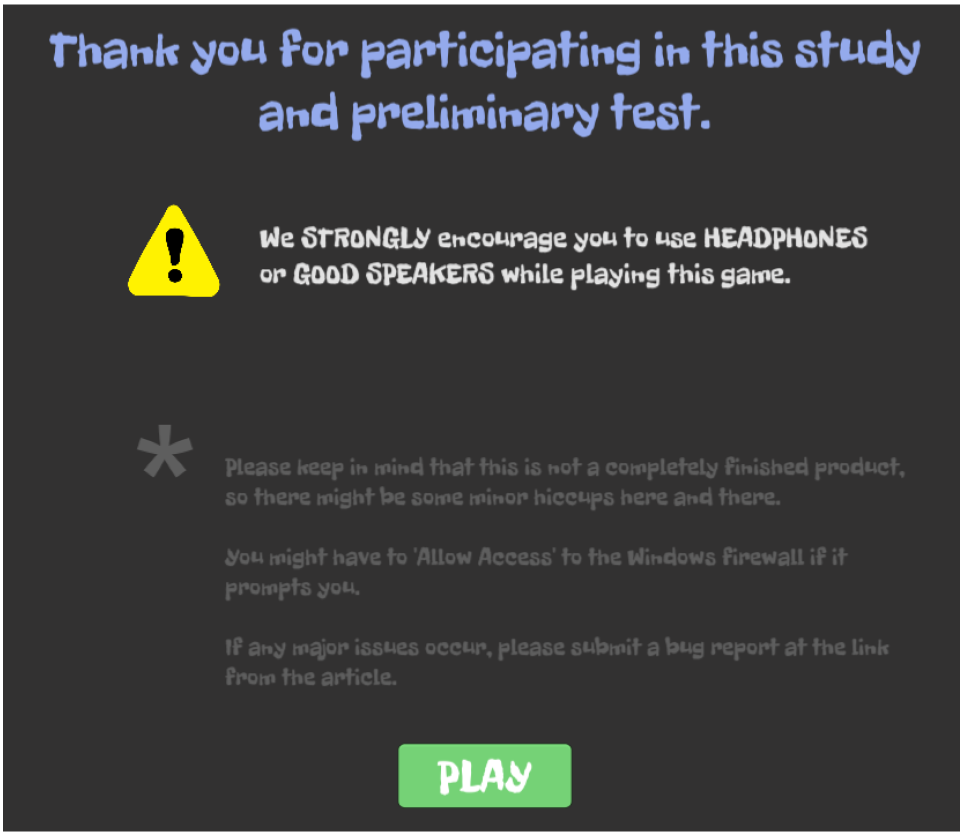


Figure 5.16: The intro screen that is shown before starting the game.

Additionally, a menu system was set up, allowing the player to change mouse sensitivity, change graphics settings (in case the game was booted with settings that were too demanding for the players computer) as well as master volume slider (clamped at a minimum of 10%).

5.5 Audio

Of all the systems and files in WPAG, audio is without doubt the element with the largest storage consumption. With a fairly large collection of sound effects (SFX) and numerous amounts of music layers, the sound banks take around 40% of unity project, not even including the Wwise project. The SFX features ambiences, impact sounds, and even menu click sounds, but all despite voice lines have been mixed and adjusted to be less intrusive, to pave way for the music. This means that there is no ducking on the music to make sure the participants are not hindered from listening to the music while hearing voice lines or while assaulting enemies.

All musical content were created in the digital audio workstations (DAW) Pro Tools (Avid, 2017) and Logic (Apple, 2017).

5.5.1 Audio Systems & Engineering

Audio design in video games are usually a lengthy processing using a variety of software to enable it in-game. The first step is recording and editing. The WPAG is assisted by a Canadian Audio Recordist Team called *Bay Area Sound* (BAS), providing audio assets (recorded .wav files), which is then edited and engineered by the WPAG game developers. The BAS team therefore provides WPAG with a library of sounds ready to be enabled, but in many cases it has been necessary to implement additional content in order to prototype, try out variations and get the most faithful soundscape in-game. Next step is enabling the sounds in the WPAG game. While most modern game engines come with an integrated audio engine, many choose to use an external engine like *Wwise* to e.g. make it more ideal to audio designers or extend the variety and quality of audio designer tools. Wwise is not only a obvious choice to WPAG because of the collaboration with *Audiokinetic*, but one of the leading Audio Engine Middle-ware in video game industry, providing a particularly high stepping stone for controlling the sound and music design. As this project investigates topics that is directly applicable in the video game industry, music should be of likewise high quality.

5.5.2 Music design

Since the game requires the player to move in and out through all areas, the music is going to switch quite a lot. In order to make such transitions smooth and natural sounding, it is necessary to develop a plan that can ensure that transitions will not sound awkward or draw too much attention. Furthermore, while each area is different, they are all located in the same world, pertaining a certain coherence between them.

It is recommended to listen to the music, found via appendix 10. Although we will try our best to describe the elements that each music piece consists of, it is definitely easier to have a listen yourself.

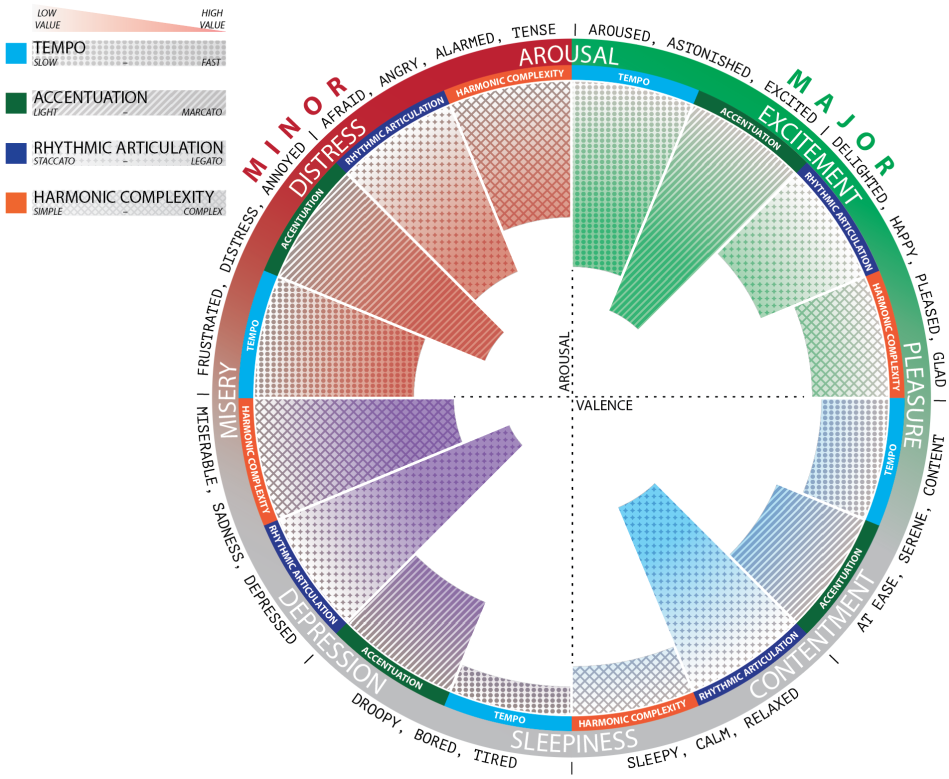


Figure 5.17: The affective composing model (Lynggaard Olsen and Maretty Sønderup, 2015)

To aid the music production process, inspiration was taken from conventions established through several years of experience with game music, as well as the *Affective Composing Model* (ACM, figure 5.17) - a tool by Lynggaard Olsen and Maretty Sønderup (2015) to aid composers in creating music with a certain affect. It should be mentioned that since this study does not focus on either conveying or inducing a certain affect, the ACM was merely used as an inspirational tool.

Following discussions in chapter **3. Research**, it has also been decided to use a mix of the horizontal and vertical approach. It is necessary for this study to have themes that do not vary beyond recognition (thus diminishing the associative qualities of the leitmotivic material), however, repetitive and predictable music is still unwanted. As such, it has been decided for each musical theme to compose a selection of alternative instrumentation and slight variations that contain the same qualities, while allowing to be interchanged dynamically in the music system - something Wingstedt (2004) classifies as *elastic scoring*. This process is described further in section 5.5.2.5.

5.5.2.1 Common Musical Features

A general structure has been established, basing all the music tracks on the same harmonic progression (figure 5.18).

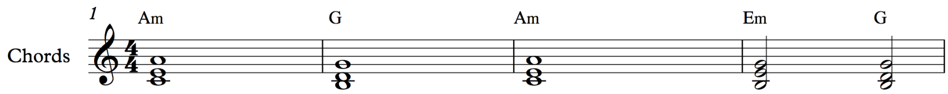


Figure 5.18: The common chord structure that is shared between all music pieces.

With this progression in common, it is possible to establish a sort of coherence between all themes, while still having each theme be different. The main advantage is that since it is known that all music pieces has this progression in common, it is possible to combine area and enemy themes knowing that they will not clash harmonically.

5.5.2.2 Area Themes

Along with each area looking distinctly different, each area has its own theme. In this section, snippets from each theme will be presented, along with a small discussion about instrumentation choices.

Forest There are arguably two parts to the associative parts of the forest theme. The theme is relatively calm, with a pizzicato string section forming the base of the composition (figure 5.19).

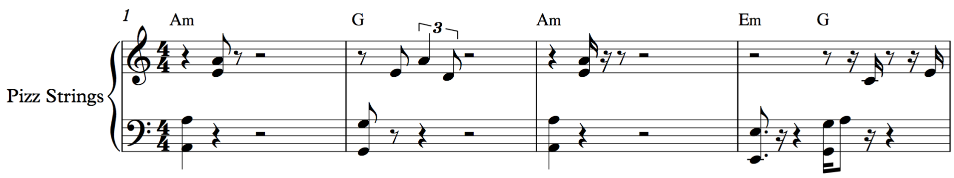


Figure 5.19: Pizzicato strings from the Forest theme.

The pizzicato strings are arguably one of the most defining elements of the theme, partly because of the fact that there will always be a variation of pizzicato strings playing. On top of that, a flute section plays what is seen in figure 5.20.

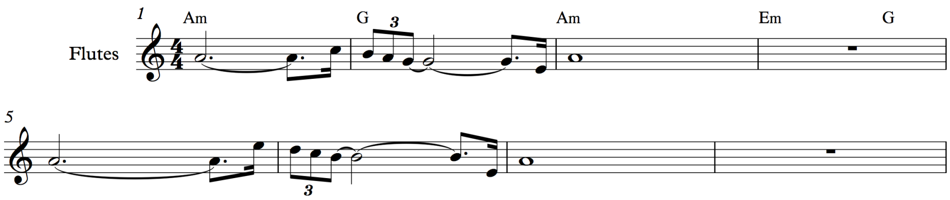


Figure 5.20: Flute melody from the Forest theme.

Additional instruments in the forest theme includes a harp playing an arpeggio pattern, a mellow tuba, a clarinet accompanying the pizzicato strings and an almost angelic female voice singing a calm complimentary melody.

Volcanic This theme was designed to be stressful, or at the least ominous and unnerving. At its base, it has rumbling orchestral bass-drums that produce a constant rumbling low-end, accompanied by a harsh, almost distorted brass section playing deep, 'movie trailer-like' tones (similar to the horns heard in the trailer for the hollywood hit Inception, directed by Christopher Nolan (*Inception Cinematic Trailer*, n.d.)). On top of that, a divisi string section plays an ostinato melody that participates in the intense atmosphere of the music piece (5.21).

Adding to the chaos, a church organ plays heavy chords along with an ominous choir, that are both periodically detuned up to a an octave (in a descending fashion), giving an unnatural feeling that something is not right.

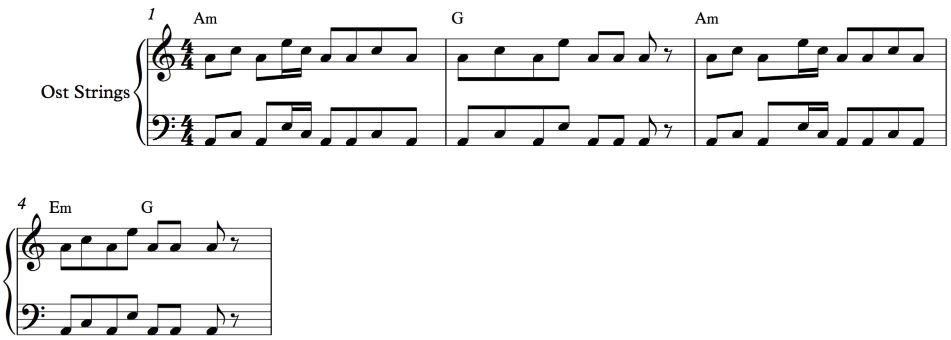


Figure 5.21: Dramatic string ostinato from the Volcanic theme.

Desert Similar to the harp and choir of the forest theme, the desert theme also draws from a lot of conventions established in many other games. A prominent element of this theme is the middle eastern sitar that plays a mystic melody (figure 5.22)

Figure 5.22: Mystic sitar melody from the Desert theme.

Another element of the desert theme is the constant, dry sound of a detuned shakuhachi playing drone tones to create a mystic and dry ambience. Additionally, now and then more mellow, yet majestic tubas play a melody that relates to the huge desert temple found in the desert (figure 5.23).

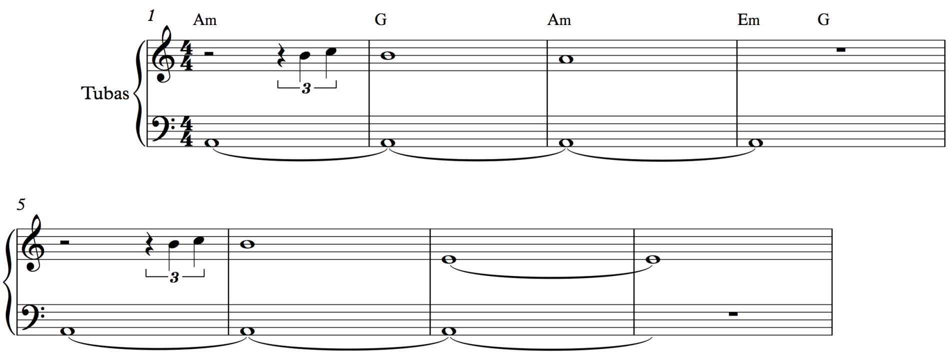


Figure 5.23: Majestic tuba melody from the Desert theme.

5.5.2.3 Enemy Themes

Just as the areas, each enemy has its own unique theme. However, certain things are in common between the themes. Since the enemy themes are meant to be played on top of the area themes, they occupy more or less the same auditory space in terms of instrumentation. Each enemy theme has a *Base Melody* and a *Melody* part, of which the base melody is played by orchestral double bass, celli, and piano. The melody part is usually a clarinet, tuba or flute section. All enemy themes also have a variation of a tamburine and shaker combo, adding percussive elements.

Since it is not possible to draw from conventions for the enemies, as they are all unique, and enemies typically don't have music associated with them in the same way that area types do, aside from the musical sections, each enemy theme also includes sound effects from the respective enemy. The Crawler theme has growls and the movement sound of the enemy (rattled leaves and sticks), the Evil Head theme has growls, bite sounds and the constant hovering sound that it makes when flying around, and the Spitter theme has growls and the sound that it makes when firing a spit bullet.

Crawler The Crawler theme reflects to a certain extent the movement pattern of the Crawlers. The base melody can be seen in figure 5.24. The base melody can be described as "sneaky" or "clumsy", which fits the visual appearance of the monster. It also draws slight inspiration from the infamous *Jaws* shark theme, except in a more rhythmically consistent way.

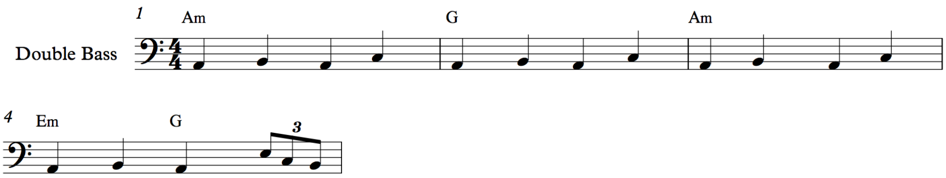


Figure 5.24: Base-melody of the Crawler theme. Note: the notes are played staccato.

On top of the base melody, a clarinet plays a melody (figure 5.25) with a couple of 'awkward' tones, giving a sense of imperfection, again, fitting the visual appearance of the enemy.

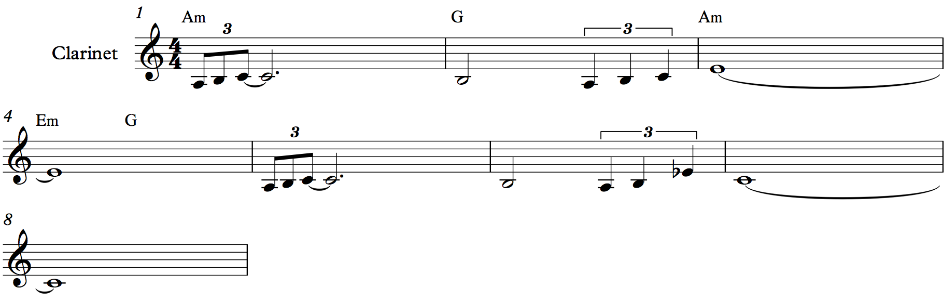


Figure 5.25: The clarinet melody from the Crawler theme.

Spitter Similar to the other enemies, the Spitter has a strong base melody that is responsible for most of the associative qualities. This base melody is seen in figure 5.26. It is a bit more driven and playful than that of the Crawler, which fits the increased awareness that the player needs to have in order to successfully dodge the spit-projectiles of the enemy.

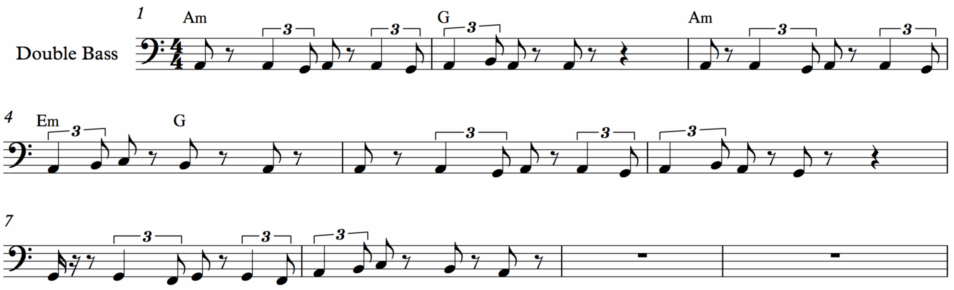


Figure 5.26: The base-melody of the Spitter theme. Note: the notes are played staccato.

Similarly, the melody on top of the musical base is more melodic and exciting (figure 5.27), played by a flute section. This melody can be considered quite catching, and again stresses the increased attention required around this enemy.

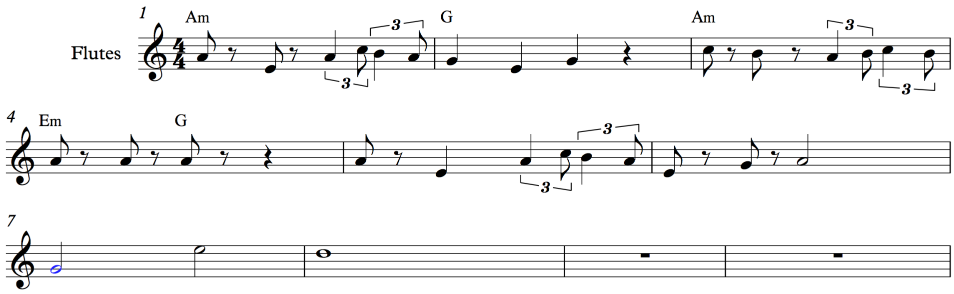


Figure 5.27: The flute melody of the Spitter theme.

Evil Head The base melody of these fierce, flying monsters is even more "hasty" than the rest, as they are also the fastest enemies. The base melody consists of something reminiscent of a guitar riff, or perhaps a hectic piano riff. The melody kan be seen in figure 5.28.

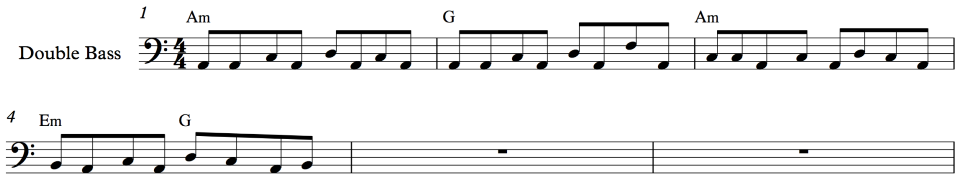


Figure 5.28: The base-melody of the Evil Head theme. Note: the notes are played staccato.

On top of the base melody, a legato clarinet adds an ominous and dangerous-sounding melody that contrasts the otherwise fast-paced base melody.

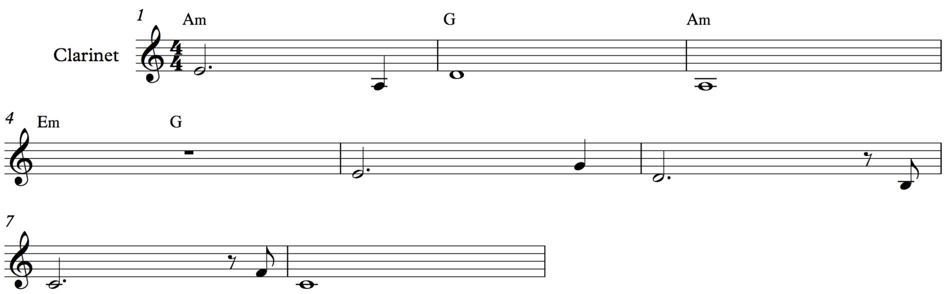


Figure 5.29: The clarinet melody of the Evil Head theme. A variation exists, where a tuba plays the melody instead.

5.5.2.4 Ambient Music

The ambient music consisted of a very simplified melody. It was designed specifically to convey as little info as possible, therefore it was very ambiguous and mellow, accompanied by a slow rhythmic bass (score seen in figure 5.30).

Figure 5.30: The bass line of the ambient music sequence.

5.5.2.5 Implementation in Wwise

All the music has been exported in stems, meaning that each individual instrument track has been exported as a separate file (.wav). These files are then combined in Wwise as a single music segment, meaning they will all play at the same time, even though the might have different durations of intro- and outro segments. Wwise allows the user to set up what is known as *pre-entry* and *post-exit* markers that define what part of the files are the actual "body" of the loop (illustrated in figure 5.31).

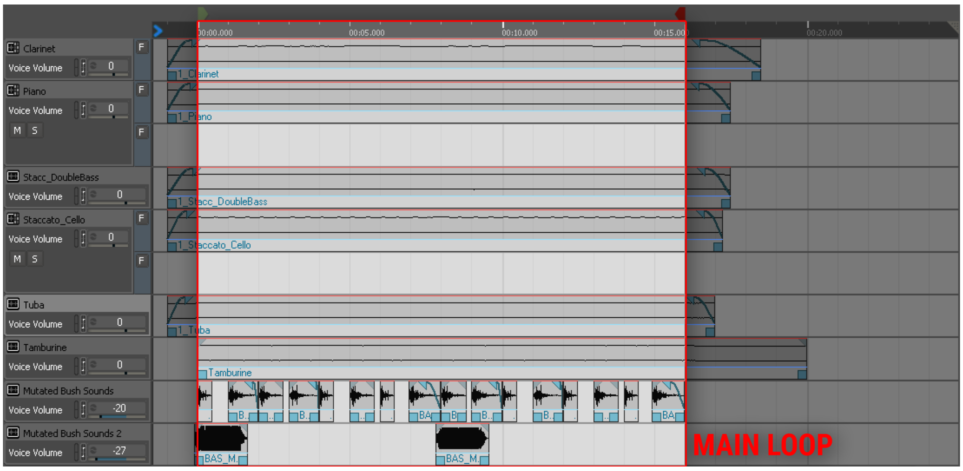


Figure 5.31: The highlighted main loop section of the music. The extents go from the pre-entry marker (green) to the post-exit (red) marker.

This "body" or main loop section is the defined extents of the loop. The two markers are indeed a powerful tool, as it is possible to keep eg. reverb tails without merging them with the beginning of the clip, allowing for greater flexibility for transitioning between many different tracks. An illustration of the looping method is shown in figure 5.32.

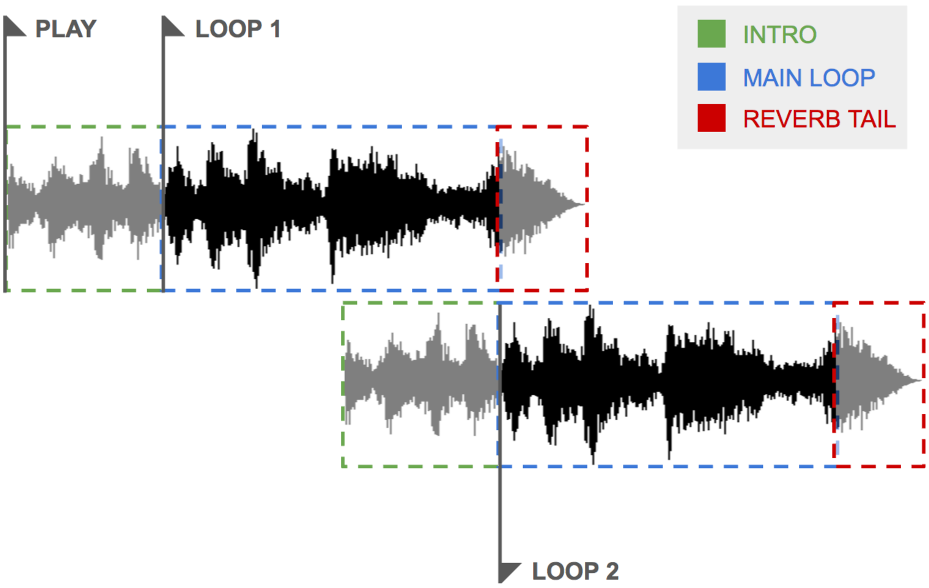


Figure 5.32: A simplified illustration of how Wwise handles looping tracks.

There are various advantages of setting the music up this way. One advantage is the ability to cut down each track to the minimum required length, saving space in the sound banks. Another, and perhaps the largest advantage is the fact that each instrument track can have several sub-tracks, that can be set up to play either randomly or sequentially. These subtracks can be used to store different variations of the same melody, or even silence (in the form of empty sub-tracks). An example of this is seen in figure 5.33, where the cello has a single empty sub-track, which means that each time the segment loops, there is a 50% chance that the melody will play.

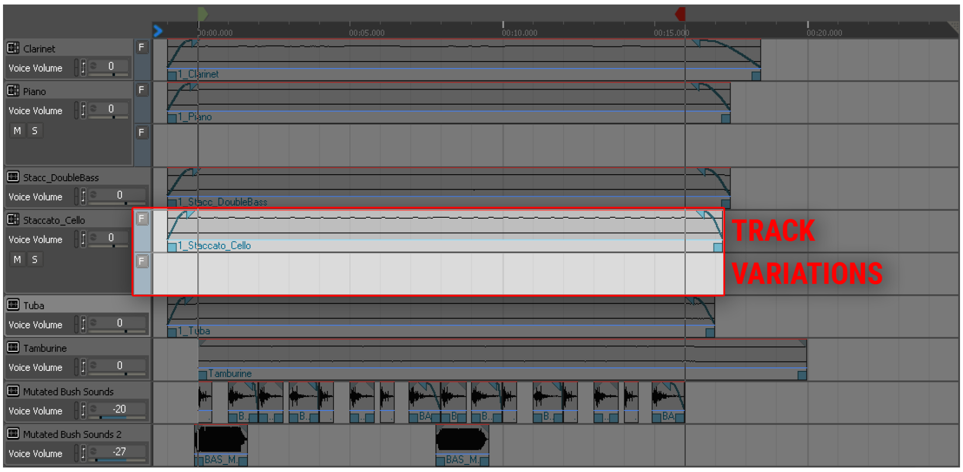


Figure 5.33: An example of clever use of random selection of sub-tracks in Wwise. This track has a 50/50 percent chance play either the melody file, or silence.

In order to switch between the different themes, a FSM system was set up in Wwise (figure 5.34), allowing the game logic to easily communicate with the music system.

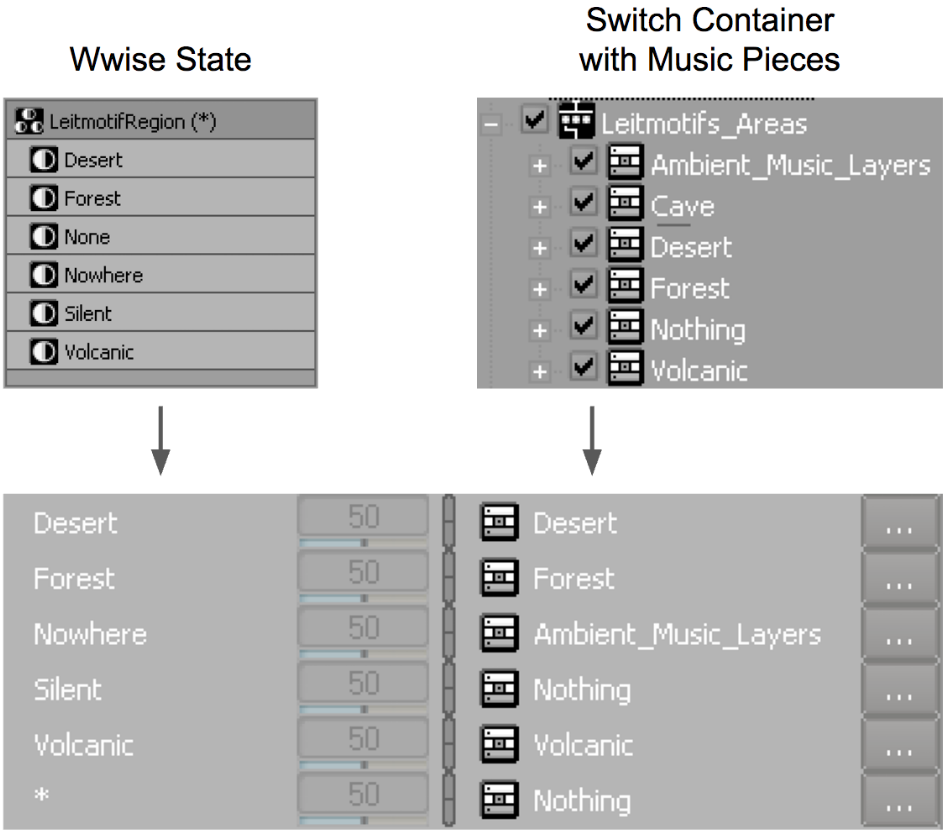


Figure 5.34: Visualisation of how each music object is connected to a State object in Wwise that the game code can communicate with.

For handling transitions, a transition matrix editor is used to define default transitions, as well as specific specialised cases between certain music objects. As the music was written to be able to seamlessly fade between segments, it was only necessary to set up a couple of general transitions that fade over approximately 3 seconds. Furthermore, as this study requires the enemy themes and area themes to play concurrently, it was necessary to set up a second music system. As such, two music systems run simultaneously, which has brought with it some unique challenges. One of the biggest returning issues was making sure that the two systems were always synced. Since all themes share the same chord progression (as described in section 5.5.2.1, figure 5.18), if one theme was not in sync with the other, a horrifying cachophony would bother the listener.

5.5.3 Containers

The goal of the last mission in WPAG was to find, destroy crates or barrels, and bring back coins to the *Wwizard*. These containers were models of either barrels or crates configured with a destructible script that would allow the player to destroy them. Besides aesthetics differences, the contains would behave similarly and have equal chance of drops. When a player would attack one of these containers, the container would receive damage and, if below it's health-border, be destroyed. On destruction the crates would either drop coins or an explosion object. The explosion would only reduce the player by 90% life, so the player would not die only if charging for the next barrel.

Figure 5.35: Trigger zones of crates.

5.5.3.1 Container Polarities

The containers were also equipped with a polarity script which would, on entering a trigger, measure the the distance from the player to the container. This container would then be added to the *StatusManager* which would take an average of all nearby container distances and polarities, and send it to the Wwise bank using a global Real-Time Parameter Control (RPTC). An concept example is draw in figure 5.36 of how the distance and polarity is collectively controlling the volume.

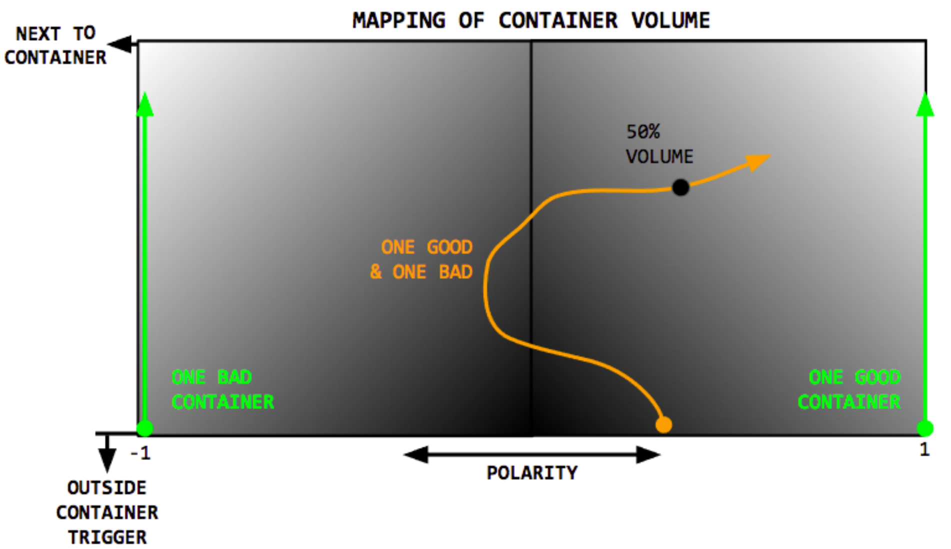


Figure 5.36: Concept drawing of volume intensity dependent on polarity and distance.

Major & Minor The polarity axes (-1) and (+1) is represented by two different music layers accompanying the ambience theme. The good (+1) and bad (-1) is composed by a voila A(2) with a major third, and with a minor third, respectively, as the ambient music is composed in A. The layers are designed with reference the influence of major and minor compositions described by (Gomez and Danuser, 2007).

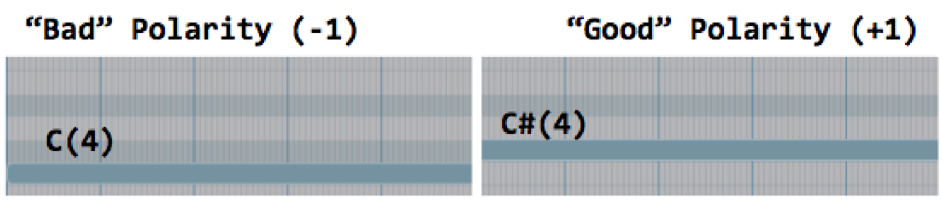


Figure 5.37

5.5.4 Other Audio Implementations

5.5.4.1 Footstep Analysis

A common method in games to trigger a certain footstep sound, is by either making trigger zones or by tagging certain player collisions, to keep track of whether the player is walking on sand, grass, etc. However, since the WPAG project is rather large project and required something more dynamic a special foostep material analyser was designed. The system analyses the polygon below the players foot at a current point in the animation, and compares the material to a list of materials to see what sound it should trigger. Each material in the material list has a footstep switch state, which also means there can be a variety of different materials for one sound, it's all a matter of what sound would represent the material most authentically. However, it is not necessarily possible to just take a material from a polygon since a polygon does not have any information only the vertices, and therefore the system looks and compares all the vertices to estimate the correct material as seen in figure 5.38

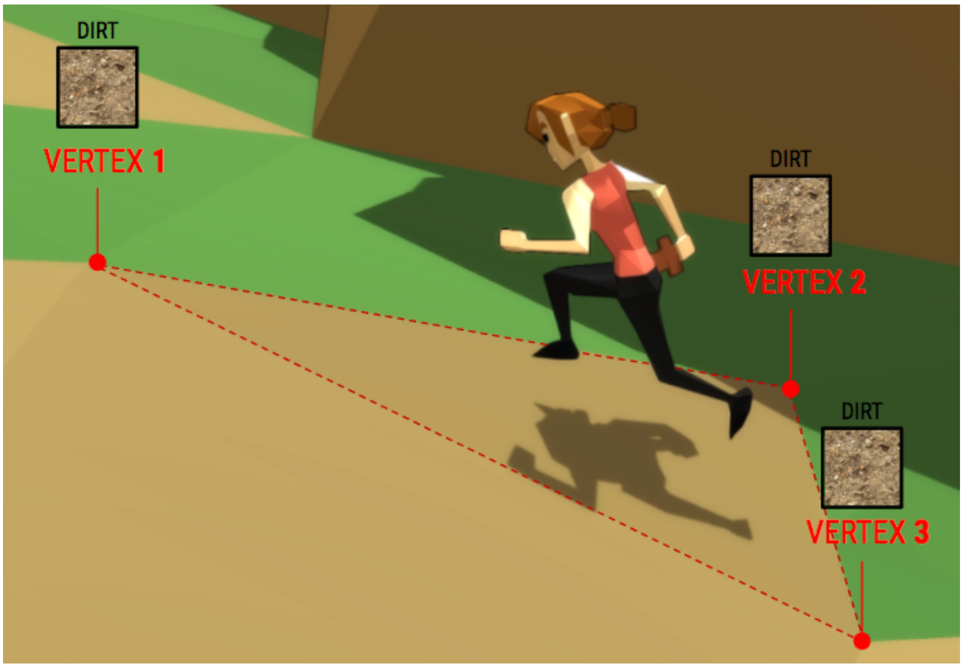


Figure 5.38: Visualization of how the footstep analyser finds the material by looking at each vertex.

5.5.4.2 Impact and Collisions

To support the combat interactions, a impact system was designed. First the weapon swings themselves were supported by a simple wind 'woosh' depending on weapon type. As the weapon swing is animating, a collider is enabled on it and everything it collides with triggers a impact analysis. The impact analysis retrieves the material of the collided object and again like the footstep material 5.5.4.1, when it matches a material it triggers the switch in Wwise and after plays the sound.

Figure 5.39: Impact switch group in Wwise.

5.5.4.3 Environmental Ambiences

The ambience in WPAG was designed using the Wwise components *AkAmbient* where the ambience sound sources are spread out on a set of positions as seen in figure 5.40. The ambiences were designed so the closer the player would be to the sound source, the more spread the sound signal would be.

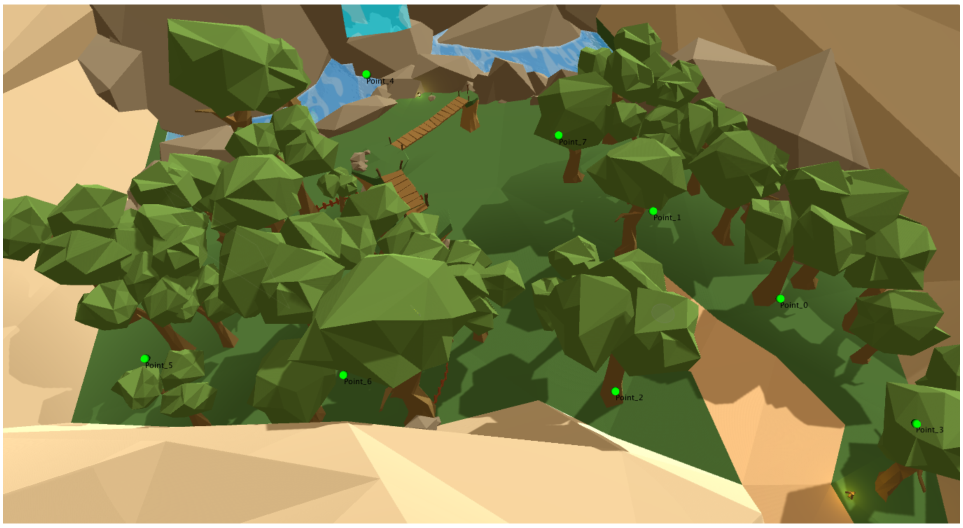


Figure 5.40: Ambience emitters spread out in the Forest Region represented as green dots.

Also a low pass filter were added on the sound while moving away from the sound source along with the decrease in volume as seen in figure 5.41.

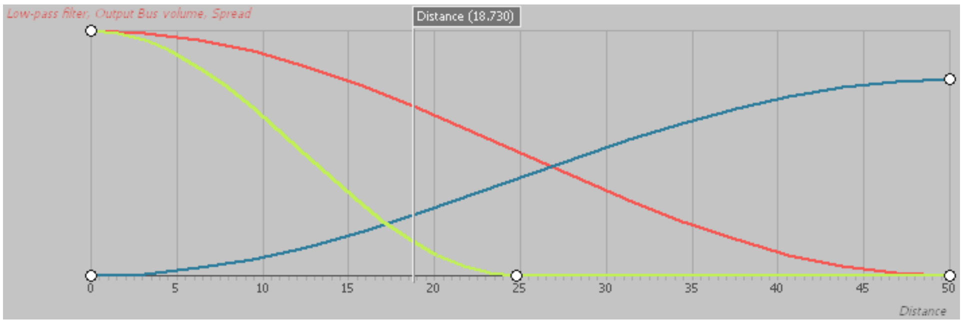


Figure 5.41: Attenuation curves in Wwise. Green is spread, red is volume, blue is low pass filtering.

5.5.5 Miscellaneous

As a final touch to the online export versions, an introduction screen was added before entering the game. This screen would describe what the participants were about to experience, but also that it would be strictly required to use headphones or a good set of speakers.

6 Test Methods

This section will present the general testing procedure of the final tests, along with information about each test location, and what data is gathered in each test. Additionally, this section presents a few problems that occured during testing, as well as the methods used to evaluate the gathered data.

6.0.1 General Test Procedure

To test the problem statement presented in section **3.1.1**, a test procedure with multiple phases was designed. While the procedure has changed slightly between test iterations, the one presented in table 6.1 is the procedure of the final tests.

Table 6.1: General test procedure and the goals for each phase.

Label	*In-game Objective*	*Goal*
Training Phase 1	Gather items in the three areas.	Exposing the subject to the music themes of each area.
Training Phase 2	Defeat the different enemy types with the correct enchantments as the Wwizard introduces them in the dream world.	Show each enemy along with their weakness and musical theme, while trying to connect the enchantments to the enemies.
Test Phase 1	Find and destroy the core that contains the evil essence, and bring it to the Wwizard	Giving away the solution via the music, see if subject goes to the right place with the right enchantment.
Test Phase 2	Collect coins in crates and barrels, while avoiding getting killed by explosions.	Foreshadowing the contents of the container (good/bad) via musical sound effect.
Questionnaire Phase 1	Answer questions about demographics, as well as the recent playthrough	Gather information about the thought process for finding the evil essence, and destroying containers with coins.
Questionnaire Phase 2	Listen to music, pick association from images 6.5.	See if participants can recognise or identify the music post-playthrough.

6.0.2 Testing Locations

Throughout the testing perod, multiple locations were used. This was due to various factors, including availability and going where there was opportunities to gather test subjects. All test setups had somewhat identical equipment setup, but the surroundings were at times different.

6.0.2.1 Graphics Lab

The *Graphics Lab* is a laboratory at *Aalborg University Copenhagen* where multiple 'workhorse' computers are set up. It is intended for performance-heavy programs, such as renderings of 3D- and light simulations, as well as machine learning algorithms. They are equipped with strong graphics cards, making them a good fit for an ultra smooth gaming experience. The laboratory is fairly small, and requires special access, so it is fairly secluded from noise and other distractions. It is, however, a room where other students can work on their projects, although not a common room for leisure or free time.

(a) (b)

Figure 6.1: Images of the setup in the Graphics Lab at Aalborg University Copenhagen.

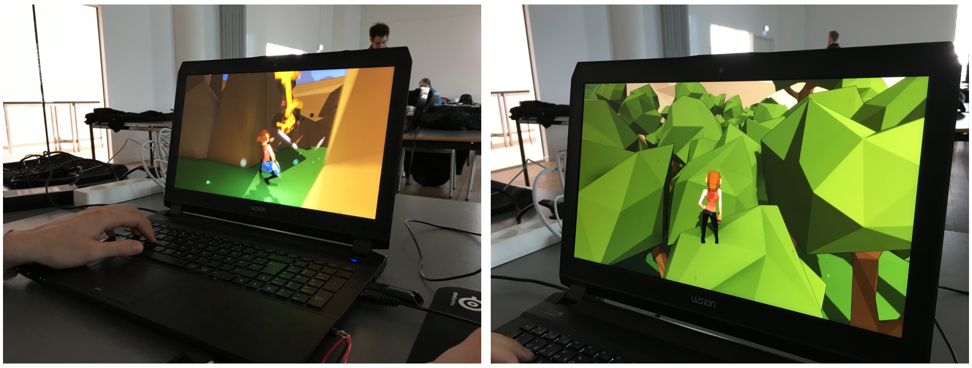
6.0.2.2 Aalborg University Public Areas

Other test locations were used in the main building of *Aalborg University Copenhagen*, close to the information desk and canteen.

The first setup was during a demo day reserved for students of *Medialogy* and *Sound and Music Computing* to showcase and present their current work and projects. The participants played on a high performance laptop (figure

6.2) and were mostly other students from different fields of education. As there around 6 projects presented, the noise and distractions were fairly low.

The second setup was close to the first, but further down a hallway that is a bit more public than the first location. However, participants were positioned in a corner with noise cancelling headphones to limit distractions and noise.



(a) (b)

Figure 6.2: Pictures of the portable setup that was sometime used around public areas at Aalborg University Copenhagen.

6.0.3 Captured Data

As with the test procedure, the collected data has changed slightly from what was captured in the preliminary test. The amount of in-game metrics have mostly expanded over time to include more back-up captures in case of problems, while a few non-used metrics were removed to diminish file sizes (which is an improvement since the data is saved on a server).

The final list of captured metrics is presented here:

- **General Playthrough Information**
  - **Control Group**: 'True' if the participant is in the control group, 'false' otherwise.
  - **Time Played**: Time (seconds) from when the application was executed to when the data was submitted by the program.
  - **Music State Recordings**: The current states of the music system, captured once every second.
- **Player Actions and Character Information**

- **Player Positions and Rotations**: Position and rotation data of the player character, captured with an interval of 1/4th of a second.
- **Camera Positions and Rotations**: Camera positions and rotations, captured with an interval of 1/4th of a second.
- **Death Information**: What killed the player, where did it happen, and when.
- **Attacked Objects**: A list of all attacked objects, each with their own timestamp.
- **Trigger Tracking**: Information about when the player entered triggers in the game world.

- **Quest/Mission Objective Information**
  - **Quests Skipped**: The quests skipped, so if the conductor would skip a participant to the next quest because of e.g. a bug, it would be recorded here.
  - **Quest Changes**: A list of quest names and their time-stamps of when the quest was initiated.
  - **Container Information**: Constant capture of container-related information like polarity, distance, etc.
  - **Area Configurations**: Information about the areas and what was the correct solution in Test Phase 1.
  - **Crate Destruction**: When and where a crate was destroyed.
  - **Enchantments**: Information about when an enchantment was picked up, and what type it was.
  - **Mission Objective Interactions**: Information about interactions with coffins (preliminary tests) or meteors (final tests), and whether that object was the quest objective.
  - **Coins**: A list of the collected coins, time-stamps for when they were picked up, etc.
- **Questionnaire Data**: The entire questionnaire object, including all questions, their texts, descriptions, options and the users answers seen in appendix 10.

Aside from the in-game metrics that are automatically captured and saved as an XML file, for all tests that were conducted locally, observations and interview notes were captured and saved for later analysis.

6.1 Evaluation Methods

In this section we present the methods for parsing and analysis of the captured data (listed in section 6.0.3).

6.1.1 Preparing Inconsistent Data Sources

As all the captured data were parsed in Matlab from an XML to more comprehensible graphs and matrices, some data sources caused some problems by being somewhat inconsistent. This sub-section will try to briefly explain how we attempted to fix the data to a more usable format.

6.1.1.1 Latency

As any computer rarely operate without any frame rate latency or under variable speed, a notable issue was recording data with a delay. Multiple data recordings like *Player Positions* and *Music States* were captured using delays.

6.1.1.2 Inter Sample Delays

One method to solve this was to measure their inter-sample frame rate delay using the Music States as they had also time-stamps.

As the Music States were suppose to capture every second, the additional time would then be subtracted. As the Player Positions were captured four times a second, the subtracted time could then be divided by four, and from that have a somewhat estimate of the player latency. The following formula represents the capture method:

$$\bar{D} = \frac{\sum_{s=1}^{n}(t_s - t_{s-1}) - i}{n}$$

The medians of all participant delays were then sampled in a histogram to show overall inter sample delay in figure 6.3.

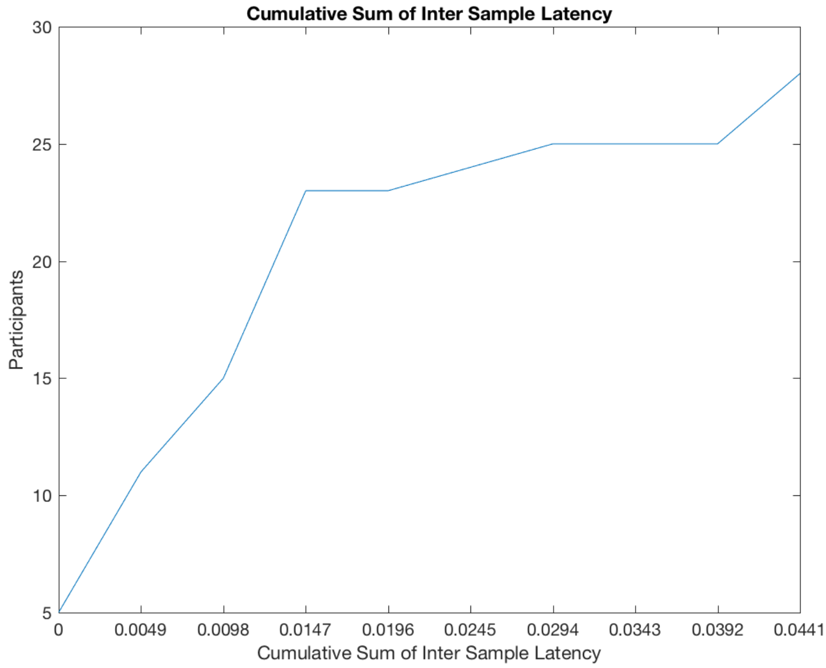


Figure 6.3: Cumulative sum plot of the inter sample latencies of all participants in test 7.2.4.

The steepness of the curve in figure 6.3 describes what amounts of latency cause the most delay in the data sets.

6.1.1.3 Music State Gaps

As the WPA game was not recording data in the background, another issue was when WPA was minimized and the game would not record both *Music States* and *Player Positions*. This would cause, in some cases, great gaps in the data gathering, resulting in e.g. unaligned time-stamps and presumed player position capture intervals.

However, as the *Music States* also contained time-stamps, this was used to count of the duration of each gap in the data recordings. Each timestamp recorded by other data variables (e.g. 'QuestChanges'), was then reduced by the gap durations. One of the more extreme examples is participant 19 in the 7.2.4 test, whom ended up submitting a time duration of around 46 hours. When looking closer at this participant it seems the participant minimized several times, possibly due to interruptions, but didn't seem to any wildly different player behaviour during gameplay compared to other participants in the same test.

6.1.1.4 Re-sizing Delay Captured Data

Despite filtering out the music gaps in section 6.1.1.3 and the latency, the player position was still slight inaccurate in size compared to the full gameplay duration. As the player position and complete play duration should ideally be of same length, a re-size method was used on the *Player Position* as described below.

6.1.1.5 Interpolation

After having parsed the player positions from the XML recording, the data would processed using an interpolation method seen in snippet 1. This process would depend on a linear curve of the player positions, and from this generate new positions in an array matching the size of the entire gameplay duration.

```
% Size of Play Position Array
PPsize = (size(PlayerPositionsX, 2)/4);
% Size of play duration, without gaps and scene loading
Dsize = ((PlayDurations(p)-theseGapsVal)-SceneLoad{p,2});
% Get ratio of which
durationRatio = Dsize / PPsize;
% Linear increasing array - size as Player Position Samples
x = 1:size(PlayerPositionsX,2);
% Player Position Sample Data
y = PlayerPositionsX;
% Linear increasing array of the requested size
xq = 1: (1/durationRatio): size(PlayerPositionsX,2);
% linear interpolation to new array size
yq = interp1(x,y,xq, 'linear');
% Store the new array in the older array variable
PlayerPositionsX = yq;
```

Listing 1: Snippet of interpolation method in Matlab.

This functionality also improved the accuracy of the trigger analysis in section 6.1.7. An example of the player positions interpolated can be seen in figure 6.4.

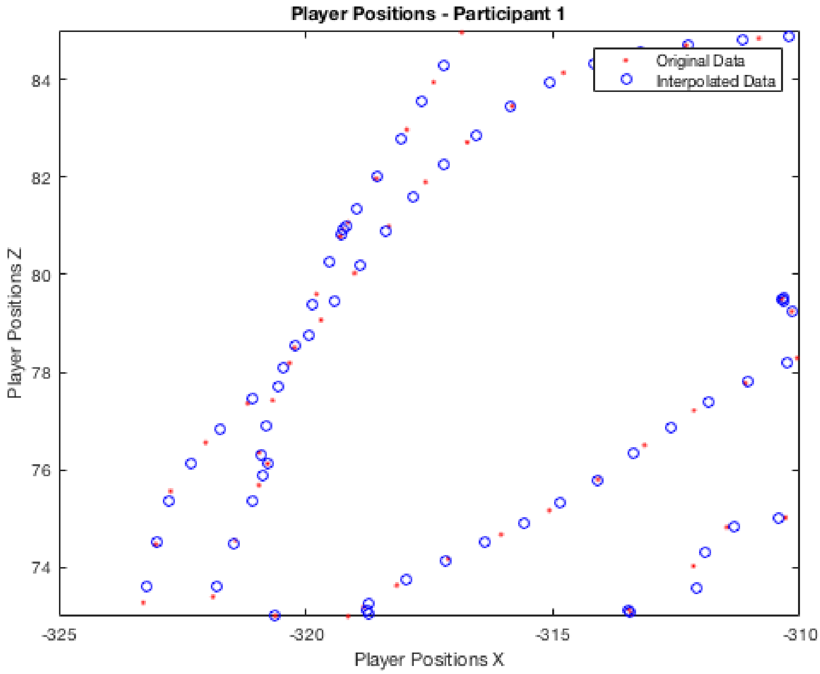


Figure 6.4: Example of player positions interpolated, from test 7.2.4.

As the interpolated points are more a prediction of what it might have been rather than actual position, this procedure does induce small artifacts. One example is when the player dies and respawns, the player position is set directly to the new position. But the interpolated data could easily think the player is moving from death position to respawn position and therefore predict the player being in position not visited. This could potentially mess up the trigger analysis in section 6.1.7, predicting the player to be in a trigger in between the death and respawn position. However, due to the many respawn positions (bonfires), this issue should occur very rarely.

6.1.2 Observations

While the participants were playing WPAG in the local tests (1-7) a conductor was observing their performance and behaviour. While writing observation notes, a few guidelines were used to aim for certain traits and behaviours. Any odd unforseen behaviour was, however, also noted during the test iterations. The study is closely investigating the influence of musical constructs, therefore it was highly important that the conductors would also be wearing a headset, as well as monitoring whether the participant had any embodied reactions on sound triggering. During gameplay there were two points where the conductor would have to pay additional attention. One being on the start of quest 3 of where to find the essence of skull. Here the conductor would pay special notice to what region the participant would be heading for, especially if the participant would be walking back and fourth without entering the actual triggers and hereby now revealing this in the in-game metrics. Secondly the conductor would have to pay special attention to the interaction with the containers, especially when the player entered the sphere of influence and would start to hear the music layer.

6.1.3 Questionnaire

In the initial testing protocol was based on an interview with the conductor answering a template questionnaire form. This way we could get feedback on the questions themselves. In the later iterations a questionnaire was designed directly inside the game, to make sure the responses would be linked to the in-game data. See figure 6.5 for the listener test choices.

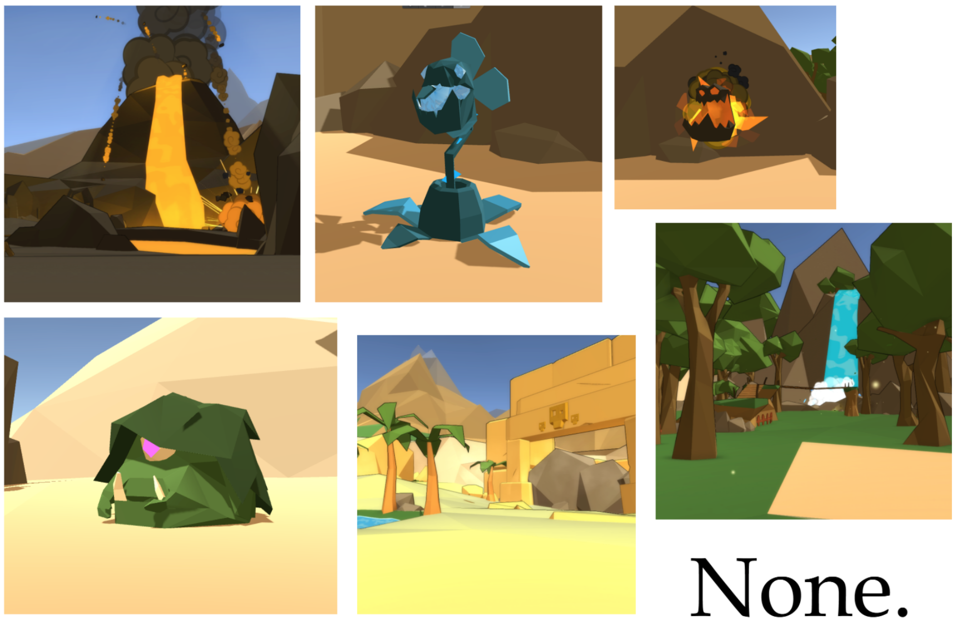


Figure 6.5: Questionnaire listener test images.

6.1.4 External Testing

In order to reach a lot of participants, it has been decided to invest effort into developing a system that allows testing abroad. Collaborations with Audiokinetic allows us to take advantage of the several hundred connections they have established as a product of their popular sound engine middleware solution.

To the fortune of this study, Audiokinetic published an article (Audiokinetic, 2017) revealing the WPAG project to the public while asking for people to participate as playtesters, meanwhile participating in the thesis project. This meant that it was necessary to to establish a system that could capture the relevant data remotely. This system is further described in chapter **5. General Design and Implementation**.

6.1.5 A Measure for Performance

Due to the comprehensive data collection and various parameters when rating the participants performance, a scoring system was designed. Using this system, every participant were rated after table 6.2. This score system only takes account for whether the participants understood their *'Mission'* (the correct region and enemy they were suppose to find), and then rates their performance in both the questionnaire listener test and in-game decisions.

In-game	**1st Try**	**2nd Try**	**3rd Try**	**4+ Try**
Tries before destroying the correct core	3	2	1	0
Tries before picking the correct enchantment	3	2	1	0
Questionnaire:	**Yes**	**No**		
Correct Choice on the Mission Region Theme	2	0		
Correct Choice on the Mission Enemy Theme	2	0		
Remembered Mission Region Theme	1	0		
Remembered Mission Enemy Theme	1	0		
Max Points	11			

Table 6.2: My caption

6.1.6 Quantized Observations and Responses

The observations (conductor notes and graph observations) was analyzed of each participant and compared to their responses in the questionnaire. This would achieve a self evaluation from the participants, along with the conductor notes and graphs confirming or refusing their feedback.

6.1.7 Trigger Simulation

One of the less reliable data captures was the trigger data. Because the triggers were disabled in *Hint mode*, so that the player would not trigger another music state while entering a zone but rather stay in the same all the time, all the Foreshadow XML recordings were not captured. However, we already had the player position itself, so by defining multiple trigger masks (see figure 6.6) in the player position data, it was possible to reproduce the trigger zones and even to a more elaborate point than using the previous triggers from the game.

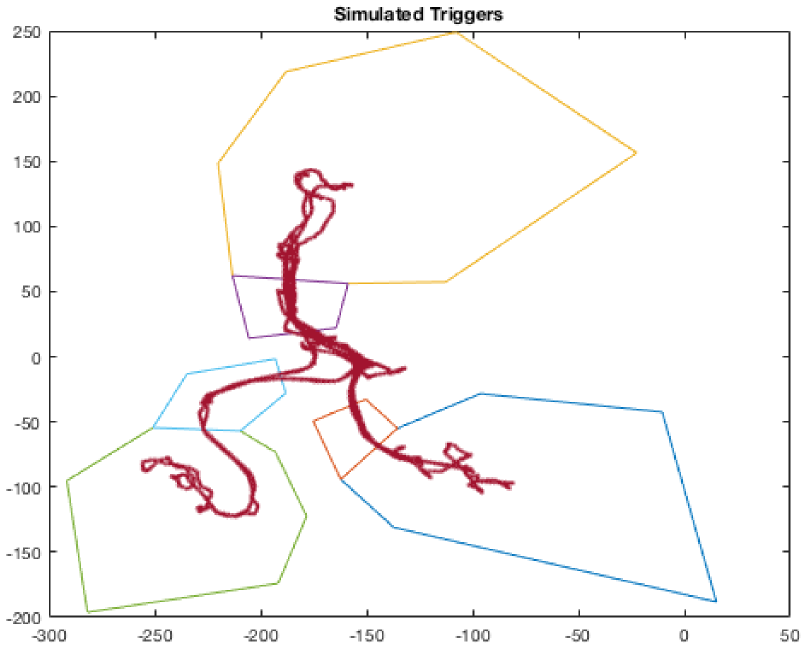


Figure 6.6: Participant Triggers and player position in game space (x,z).

6.1.8 Timelines

Finding the essence hidden in one of the cores, is one of the primary goals of the game. The *Timeline* presented in figure 6.9, was generated for each participant and used to evaluate on the participants performance or reference some of their statistics to what decisions they had during. The blue line represents the *Find the Essence* quest, all from getting past the dream state to beginning the *Find the Coins* quest. In the left side of the figure the *Mission* is displayed.

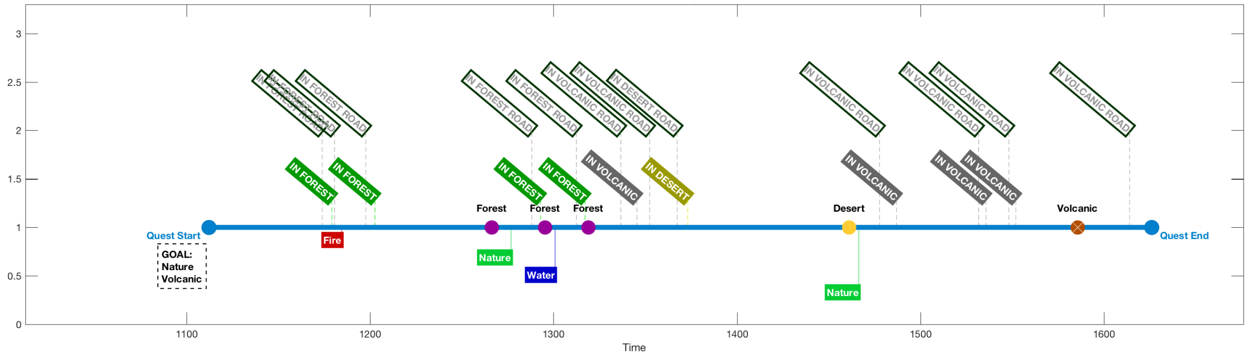


Figure 6.7: Timeline of participant 5.

6.1.9 Crate Polarities

In the final quest (*Find the Coins*) the player objective was to find containers (crates or barrels), destroy them and get coins. As described in 5, these containers would within a certain distance hint whether the container was dangerous (-1) or provided a coin (+1) by the use of a layer in the music. To analyze this interaction container figures was produced like the one seen in figure 6.8.

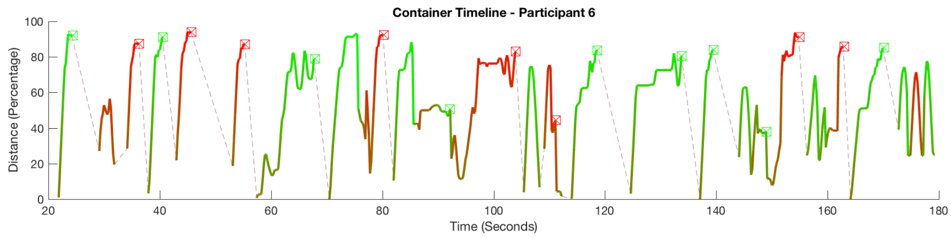


Figure 6.8: Polarity represented as (negative) red and (positive) green.

The horizontal axis is time, the vertical axis is distance to container and the color itself identifies the polarity. A polarity of (+1) is represented by green and (-1) as red. The colors can also blend, meaning if the player would stand with equal distance right next to a positive and a negative polarity container, the line would become yellow.

6.1.10 Gameplay Overviews

The gameplay overview was used to get a sense of the entire play test. It collects all the positions separated in quests and the music states both *Enemy Themes* and *Region Themes*. The destroyed containers are also marked in *Find the Coins* position graph (Green boxes for containers with coins, red X for containers with explosions), as this was mainly what you did rather than explore. The position graphs also received a layer of intensity segments (heat-map), to give a sense of how long time the player was occupied in the area and this can be challenging on the position lines alone.

The two bottom graphs illustrate the music states during gameplay. The graph displays all states including the ones not necessarily used like *None* or *Silent*, evne thought the states are basically the same. This would allow to see when the player was not listening to either enemy or region music, but was also used to monitor whether any changes were not changing as they should.

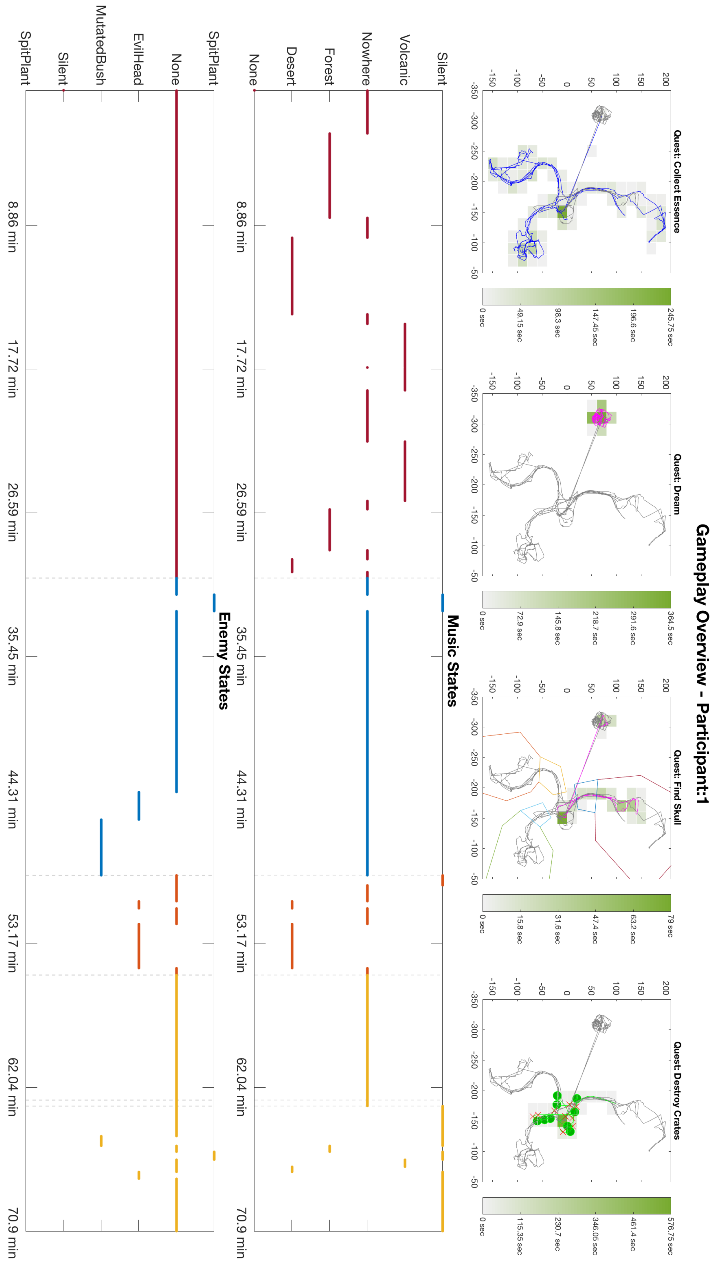


Figure 6.9: Gameplay overview - Top graphs refer to positioning in each quest, bottom two figures refer to music states during the entire play test. (From test 7.2.4)

6.1.11 Statistical Methods

The data gathering will in the conclusively analyzed using statistical methods like the Mann-Whitney U test. To estimate the statistical method the One-sample Kolmogorov-Smirnov test will be used to check if the data is normally distributed. To investigate the general player behaviour, correlation methods will also be used to compare the various different data sets. Correlations will be processed for each test iteration but also separated upon whether its control group or hint mode group.

7 Testing

As described in section **4. Methods**, the format of this research is highly iterative. The test section have therefore been split into each iteration with its own sub-chapter. Each iteration will describe the general information, results and discuss the content within each sub-chapter. After each iteration a more general analysis and discussion will be presented.

7.1 Preliminary Testing

The highly iterative production process pertained several rounds of preliminary testing in order to fine-tune and improve the experience to be as intuitive and problem-free as possible. This means that a total of 4 preliminary tests were conducted iteratively between production phases.

This section will present highlights from all preliminary tests, including major changes between each iteration, results that amount to new requirements and lead to the next iteration, and so on. Some data correlations will not be available in some tests, especially the early ones (1, 2, 3, 4 & 5), due to simply sparse collection and some of then having too small groups to even perform any correlation statistics.

7.1.1 Test 1

As a first test after several internal functionality tests, this was the first test run of the complete design scheme. At this point the aim of the test was to discover if there were any problems hindering progress or misleading the players, while gathering ideas for improvements to the test design, as well as getting hints at possible future results of a larger scale test.

It should be noted that some game elements were different from what is described in chapter **5. General Design and Implementations**, since this was the first feature-complete iteration of the project. Firstly, the enemies were named differently in this iteration. They had kept their names from the original WPAG, namely Mutated Bush (Crawler), Evil Head (identical), and Evil Spit Plant (Spitter). The themes for each enemy also did not include sound effects from the respective enemies at this point. Furthermore, all information given to the player from the Wwizard was merely in text form, while there were balance issues with the hint sound of the containers.

7.1.1.1 General Information

This test was conducted at Graphics Lab (Location described in 6.0.2.1) 03/04-2017. In-game metrics, observations and interview data was captured from a total of 5 participants with no control group. The median participant play duration was around 17 minutes, and the 5 phases of the test can be seen in table 7.1.

Table 7.1: Overview of phases in tests 1 and 2.

Label	*Objective*	*Goal*
Training Phase 1	Gather items in the three areas.	Exposing the subject to the music themes of each area.
Training Phase 2	Watch a cutscene of the Wwizard presenting enemies.	Show each enemy along with their weakness and musical theme
Test Phase 1	Find the skull hidden in one of the three areas.	Giving away the solution via the music, see if subject goes to the right place.
Test Phase 2	Collect coins in crates and barrels.	Foreshadowing the contents of the container (good/bad) via musical sound effect.
Interview Phase 1	Answer questions about playthrough.	Examine subjects thought process for finding the skull and destroying containers with coins.
Interview Phase 2	Listen to music, pick association from images	See if participants can recognize or identify the music post-playthrough.

Since the solutions to Test Phase 1 were generated randomly (a random combination of area/enemy motifs), the distribution of which area and enemy was featured more prominently as the correct choice is quite uneven. Desert was the correct area 4 out of 5 times, while Volcanic was only the answer once, and forest none. In regards to enemies the distribution was 1, 3, and 1 for Crawlers, Evil Heads and Spitters, respectively.

7.1.1.2 Results and Discussion

Of the 5 participants 2 found the skull in the first attempt, while 1 and 2 participants found it in the second and third attempt respectively. While 40% of participants got the correct area in the first attempt, it should be noted that the tiny sample size is too small to consider evaluating quantitatively and consider meaningful findings (which was out of scope for the initial pretesting anyway). Furthermore, 40% success rate is not an impressive amount over 33%, which can be assumed to be the distribution of first-try successes given random chance or luck (because of the 3 areas).

It was observed that some participants thought there was a connection between the areas and enemies, for instance, the *Mutated Bush* (later Crawler) would be assumed to be found in the forest because it is reminiscent of a bush.

When it comes to destroying containers looking for coins, participants generally destroyed more good than bad containers (table 7.2), with an average of just below 60% in favor of the good containers, despite the issues with the hint sound erroneously being quite low in the mix. A 10% improvement over an assumed 50% success rate by random chance for just 5 participants is not found meaningful.

Participant	*Good Containers Destroyed*	*Bad Containers Destroyed*
1	62.50%	37.50%
2	71.43%	28.57%
3	47.06%	52.94%
4	62.50%	37.50%
5	55.56%	44.44%
Average	59.81%	40.19%
Standard Deviation		9.09%
Average Proximity Time		64 seconds
Average Proximity Periods		8 seconds

Table 7.2: Success rates for container destruction by participants of Test 1.

Looking at the following interview it was found that all participants reported using some variation of trial and error when heading to find the skull, which suggests that they did not actively use the musical hints given to them in their decision making process.

The participants generally performed well in the post-playthrough listening test. P1 was unable to correctly identify a single one of the 6 themes, even though he reported remembering all of them from the recent playthrough.

However, it should be noted that the test conductor had forgotten to have P1 pick from the printed images, so his answers were all based on pure memory. P2, P3 and P4 correctly identified all 3 area themes while P5 could not identify the forest theme. Interestingly, P3 added that he did not directly remember the forest theme from the game, however, he had no doubt that it was the theme for the forest, as it sounded very "forest-ish". None of the participants correctly identified all enemy themes, however, P3 and P4 correctly identified the Spitter and Evil Head themes respectively. These results suggest that the enemy themes might need further work, since so few of the themes were identifiable by the participants, while the area themes were largely okay. However, the small sample size of Test 1 is deemed insufficient for making any conclusions.

7.1.2 Test 2: Preliminary

After fixing bugs and improving the flow of the experience, a second pilot test was conducted at the same location the day following Test 1 (7.1.1).

As an experiment, and in contrast to the first test, participants were told that music is an essential part of the project - an attempt to try and draw the attention of the participants towards the music, with the goal of seeing if there would occur dramatic improvements or differences from Test 1. Additionally, the image of the Village (seen in figure 6.5) was removed from the pool of options in the listener test, as it was observed to act as a sort of fallback answers when the participant was unsure of the musical theme.

7.1.2.1 General Information

The test was conducted at Graphics Lab (Location described in 6.0.2.1) the 04/04-2017. In-game metrics, observations and interview data was captured from a total of 5 participants with no control group. The 5 phases of the test are identical to those of test 1, and can be seen in table 7.1.

The correct area in Test Phase 1 were more evenly distributed with counts of 2, 2, and 1 for Forest, Volcanic and Desert respectively. However, the only enemy that was ever the correct choice in this test was the Spitter, featured all 5 times.

7.1.2.2 Results and Discussion

Of the 5 participants, only 1 found the skull in the first attempt, while the remaining 4 found it in second attempt. Similar to the results of Test 1

(7.1.1), these results can not lead to meaningful conclusions, although it is indeed peculiar that informing the participants of the importance of music in this project certainly didn't positively affect the amount of people who completed Test Phase 1 in the first try. However, 4 participants finding the skull in the second try is an increase from Test 1 (1 second try and 2 third).

Again, the problem with participants assuming connections between areas and enemies was observed several times, leading to the thought that the in-game explanations were insufficient to counter the apparent intuitiveness of the incorrect assumptions.

A slight increase in success was seen in participants of Test 2, with an average of around 62% good crates out of all destroyed crates (table 7.3). Again, the small participant count discourages conclusive remarks on this basis.

Participant	*Good Containers Destroyed*	*Bad Containers Destroyed*
1	83.33%	16.67%
2	70.00%	30.00%
3	38.46%	61.54%
4	45.45%	54.55%
5	73.33%	26.67%
Average	61.54%	38.46%
Standard Deviation		19.20%
Average Proximity Time		157 seconds
Average Proximity Periods		11 seconds

Table 7.3: Success rates for container destruction by participants of Test 2.

It was also observed that some containers were placed too close to each other, causing the system that is in charge of foreshadowing the contents of the containers (described in section 5.5.3) to almost mute the foreshadowing sound.

Similar to the first test, no participant reported using the music in their decision-making process when heading to find the skull, which was more or less supported by the observation notes, as no participant appeared to listen actively in a physically noticeable way, although such a phenomenon is hard to observe.

In the listening test it was observed for the second time that almost all participants correctly identified all the area music, except for P1 who did not recognize the forest theme. Again, it was found that participants found the area music quite obvious, for instance, P2 and P3 having choosing the forest because ”the harp is very characteristic of a forest” (P3). An improvement was seen compared to the first test when looking at the participants' responses to the enemy themes; all participants had at least 1 correct identification, while P4 had 2. Interestingly, only P2 correctly identified the Spitter theme, which was the enemy that was part of Test Phase 1.

In spite of the increased correct identifications, it is still relevant to improve the enemy themes before running further tests.

7.1.3 Test 3: Online Tech Demo

This test acted mostly as pilot test of the technology capturing data and uploading it to a server. The aim was to identify what kind of problems online testing would have and validate whether it would be good enough to launch with *Audiokinetic*.

From the results, observations and feedback from test 2 (7.1.2), it became apparent that the cutscene strategy for introducing participants to the enemy themes (Training Phase 2) was ineffective in establishing connections between the enemies and their respective themes. As such a new approach was implemented since the previous test. Instead of a period of non-interactive cutscene, a new, hidden location was implemented in the game world: The Dream World. At the same point in the playthrough as the previous cutscene, players would instead be teleported to the dream world where the follow scenario would occur:

- The Wwizard presents an enemy (order is random), explains it along with its corresponding weakness (enchantment).
- Player picks the appropriate enchantment form an enchantment stand.
- Player fights 3 instances of the enemy.

This process repeats for as long as there is more enemies to show. For each round, the enemy's theme is played for the entire duration. The aim of this new approach is to better connect the enemies with their corresponding themes, as well as combat the apparent intuition that there is a connection between the areas and enemies. See section 5.3.3 for more information about the design. Furthermore, sound effects from the enemies was added to each enemy theme. For instance, the Crawler theme has been updated to include the growl that the creature makes, as well as the sound of rattled leaves and sticks, all aligned rhythmically to the track. This was with the purpose of easing the association process.

7.1.3.1 General Information

This test was conducted online the 21/04-2017. It was sent out to 3 personal connections who had no prior experience with the product or the study. In-game metrics, questionnaire responses and interview data was gathered from a total of 3 participants, whereof two captures were corrupted, rendering them useless for analysis. As a result of the group only consisting of 1 participant, no further analysis is performed on this basis. However, the updated test procedure is shown in table 7.4.

Table 7.4: Overview of phases in test 3.

Label	*Objective*	*Goal*
Training Phase 1	Gather items in the three areas.	Exposing the subject to the music themes of each area.
Training Phase 2	Visit 'Dream Level' and fight each enemy type.	Presented 3 of each enemy along with their weakness and their musical theme
Test Phase 1	Find the essence hidden in one of the three meteors.	Giving away the solution via the music, see if subject goes to the right place.
Test Phase 2	Collect coins in crates and barrels.	Foreshadowing the contents of the container (good/bad) via musical sound effect.
Interview Phase 1	Answer questions about playthrough .	Examine subjects thought process for finding the skull and destroying containers with coins.
Interview Phase 2	Listen to music, pick association from images	See if participants can recognise or identify the music post-playthrough.

7.1.3.2 Results and Discussion

The interviews with the participants (conducted online via the *Discord* service revealed that participants did not notice any hint sounds when near the containers. Even as the system was explained to the participants, most still reported not noticing any hint sounds, pertaining a slight volume increase for next iteration. Others expressed becoming confused because of the two layers ability to overlap and create a mix of both, seen in figure 7.1. As such, the technical implementation of the container hints was changed to no longer overlap when the player is equally near to a good and a bad container at the same time (7.2). This also means that if the participant stands between a good and bad container with perfect distance there would be no hint sound, meaning that the player would have to move closer to either container in order to get more info..

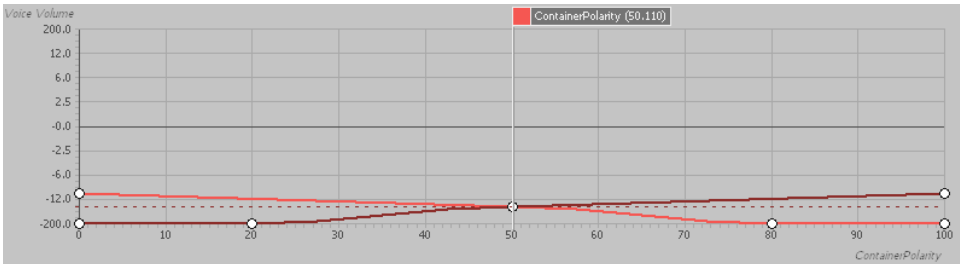


Figure 7.1: Container music layer volumes dependent on average polarity measure.

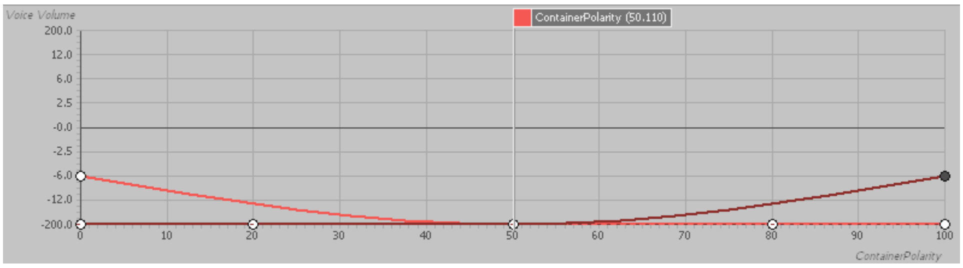


Figure 7.2: Reworked container music layer volumes, no longer overlapping at 50% good and 50% bad container polarities.

The online interview also revealed that participants questioned why they suddenly had to seek out coffins that were not in the world previously, and why a skull had to be retrieved. This pertains a change in the mission design.

Because of the corrupted data, and some data that was incorrectly captured, extra backup captures were henceforth implemented in the case that some of the measures would fail.

Participant	*Good Containers Destroyed*	*Bad Containers Destroyed*
1	55.56%	44.44%
Standard Deviation		0
Proximity Time		343 seconds
Proximity Periods		16 seconds

Table 7.5: Success rates for container destruction by participants of Test 3.

Being left with data from just 1 participant, it is impossible to make any meaningful data presentations, however, it can be mentioned that the participant had a success of 55% in terms of destroying containers for coins, which is around what is expected by random chance.

7.1.4 Test 4: Weapon & Enchantment Test

Because of the recurring problem with participants assuming connections between enemies/enchantments and regions, a special test was set up in which the entire enchantment system was replaced by an alternative. This means that enemies were no longer based on a specific element (water, nature, fire), but instead more non-descript and made of stone/dirt. Furthermore, enchantments were changed into different types of weapons that could be equipped. Enemies would then have weaknesses to certain weapons instead.

The goal of this test was then to investigate whether the change from elements to weapons would eliminate the previously described problem.

7.1.4.1 General Information

This test was conducted at 6.0.2.1 the 24/04-2017. In-game metrics, observations and questionnaire data was captured from a total of 3 participants. The median play duration was 26 minutes. Except for the aforementioned changes, the test procedure was identical to that of test 3 (figure 7.4).

7.1.4.2 Results and Discussion

The weapon-based set up did not seem to result in any changed behavior from the participants. This might be because of the fact that while the look of all enemies were changed to be more non-descript (as to not have an obvious elemental basis), they still had some form of inspiration to real world objects. The Spitter still looked like a flower (in spite of the changed color palette), the Crawler still was reminiscent of a bush, and the Evil Head still was a flying rock-based head. Furthermore, in spite of tries to establish why each weapon was the weakness of each enemy (i.e. "The *Katana* is the only weapon fast enough to hit the Evil Head"), the weapon weaknesses were more arbitrary and not as intuitive as the more generic element-based weaknesses. As such, with the interest of further testing in mind, the weapon-based setup was again discarded in favor of the original implementation.

Participant	*Good Containers Destroyed*	*Bad Containers Destroyed*
1	69.23%	30.77%
2	56.25%	43.75%
3	47.62%	52.38%
Average	56.00%	44.00%
Standard Deviation		10.88%
Average Proximity Time		109 seconds
Average Proximity Periods		8.2 seconds

Table 7.6: Success rates for container destruction by participants of Test 4.

While this test had a different focus, it is still interesting to look at the crate destruction stats (figure 7.6). Quite notable is P1, who had an impressive 70% good crates destroyed, signalling that he might have understood the musical hint.

7.1.5 Test 5: Final Setup Pilot

After test 4, there were several findings pointing towards needed changes in the test setup. After another intense production phase, the time had come to pilot test the final setup. While this test was aiming for improved performances, it also was focused on discovering any last bugs and problems, since the imminent online test would mean that people from around the world could play the game in their own spaces, with no test conductor to help them if any bugs occured.

The quest objective of Test Phase 1, finding a skull, was changed to the meteor setup described in chapter **5. Design**. There were several reasons for this. Firstly, participants of previous tests expressed confusion over the quest with the skull in coffins. Furthermore, the there were different ornamental skulls placed around the world (eg. in the volcanic region, to signify that someone had died in this dangerous area), so when the players were asked to retrieve a skull, some players would be confused (although the Wwizard explicitly stated that the skull was to be found in a coffin). Lastly, there were no narrative support to explain why a skull in a coffin had anything to do with the enemies. As such

The recurring problem with players thinking that there is a connection between enemies/enchantments and areas also resulted in updated quest-descriptions for the Wwizard, along with actual voice over recordings, giving the Wwizard a voice to listen for. Since it was found that many did not read the dialogue, and not noticing the explicit explanations about the quests, it was hypothesised that voice lines could perhaps help making the players listen to the information given. Furthermore, since players hopefully already are accustomed to receiving information by listening, perhaps there would be a rise in success.

As a measure to try and make sure that future online participants would hear the audio of the game, the master volume of the game was locked to a minimum of 10%, discouraging participants from putting on their own playlists in the background while playing the game (which is a common phenomenon). However, in theory, participants could still turn down the system volume, however in that case there is nothing left but to trust their responses in the questionnaire when they are asked about their audio configurations during the testing session.

This test was also the first to include a control group, in order to also test that everything was working in that mode. Furthermore, the score system described in section 6.1.5 was developed and employed for analysis.

7.1.5.1 General Information

This test was conducted at 6.0.2.1 the 25/04-2017. In-game metrics, observations and questionnaire data was captured from a total of 5 participants. The median play duration was 32 minutes. The test procedure was identical to that of tests 3 and 4 (figure 7.4).

The correct area in Test Phase 1 was quite evenly distributed with counts of 2, 1, and 2 for Forest, Desert and Volcanic respectively. Evil Heads were the correct choice 4 times while the Crawler was just 1, and the Spitter 0.

7.1.5.2 Results and Discussion

Test Phase 1 In spite of the small sample size, the score method (described in section 6.1.5) was employed, yielding an average score of 5 points for the musical hint group, and 2.5 points for the control group. However, looking at the standard deviations of 0.7 and 3 for the hint and control group respectively, it can be seen that the group with musical hints performed quite differently within the group (8, 2, and 5 for P1, P3 and P5 respectively).

Not a single participant got it in the first attempt, while 3 and 2 got it in the second and third attempt respectively.

Surprisingly, participants performed rather poorly in the post-playthrough listening test. Of all choices, the the musical themes were in the end only recalled with a common success of 13.3%, which is surprisingly low, considering the success of tests 1 and 2. In contrast, participants reported remembering the theme 70% of the time. This potentially points towards the associations being unsuccessful to a degree where participants thinks that eg. the forest theme is the desert theme - a phenomenon that is very contrasting to the successes of tests 1 and 2. In the same vein, 4 out of 5 participants reported remembering the volcanic theme, however, none correctly identified it when it was played in the questionnaire.

Container destruction information from the hint group can be seen in table 7.7. As evident, there is no immediate bias towards destroying good crates, as both P3 and P5 lie close to 50% success rate, with P1 having a mere 60%.

Participant	*Good Containers Destroyed*	*Bad Containers Destroyed*
1	60.00%	40.00%
3	52.94%	47.06%
5	52.94%	47.06%
Average	55.29%	44.71%
Standard Deviation		4.08%
Average Proximity Time		154 seconds
Average Proximity Periods		8.4 seconds

Table 7.7: Success rates for container destruction by participants of Test 2.

The success rates for the control group can be seen in table 7.8. Since the hint systems were disabled for the control group, there are no continous proximity data captured for this group. As evident, P4 had a surprising 70% success rate, which is higher than every participant in the hint group. However, the small sample size discourages conclusive remarks on this basis, since it is hard to tell whether all the numbers are up to chance.

Participant	*Good Containers Destroyed*	*Bad Containers Destroyed*
2	47.37%	52.63%
4	69.23%	30.77%
Average	58.30%	41.70%
Standard Deviation		15.46%

Table 7.8: Overview of the amount of good containers destroyed compared to bad (%).

As a general note, the addition of the voiced dialogue lines from the Wwizard was received very positively. A lot of compliments were received from the participants of this test, and even some previous participants who happened to drop by (not recorded) also noted that it was a great improvement. However, there were a few issues with the exports of the audio files which caused some of the clips to sound slightly phased, however this was immediately fixed for the following tests.

7.2 Final Testing

While it is tempting to continue to perfect the test setup, resource limitations pertained and end to the preliminary testing phase. Since all systems were now fully functional and bug-free, the final testing phase could be initiated.

7.2.1 Final Iteration Additions

There were no major additions to the contents of the game for the final test phase. However, since test 6 (local) and 8 (online) took place approximately over the same period of time, certain bug fixes were implemented locally, after which the online build was also updated. The build[1] was stored on an online file-sharing server, so it was easy to simple swap the build with a new one as problems got fixed. However, it should be noted that there were no major additions or removals from the contents of the game. All changes since the online test was launched (May 2nd) were only minor bug fixes that had been observed during local testing, or extra precautions for edge cases that were highly unlikely (such as falling off the game world).

7.2.1.1 Switching Groups

To equalize the population sizes of both the control group and the group that receives musical hints, a sort of 'global variable' was implemented on the server. When a player starts the game, the system sends a request to the server, asking for permission to download a file called 'ParticipantCount.txt' (containing just a number for how many participants has played so far) which is then incremented. If the returned number is even, the game is dynamically configured with musical hints, while an odd number yields the control group configuration. In the case that the participant would end the playthrough prematurely (thus not submitting the data), the number would automatically get decremented again to reflect the actual participant count. Each of the final tests saved the data to separate folders on the server, so as to not mix them up in the analysis.

[1] Build: Executable file of the game

Test Procedure As the final 3 tests all use the same build, the final test procedure can be seen in figure 7.9.

Table 7.9: Overview of phases in test 6.

Label	*Objective*	*Goal*
Training Phase 1	Gather items in the three areas.	Exposing the subject to the music themes of each area.
Training Phase 2	Visit 'Dream Level' and fight each enemy type.	Presented 3 of each enemy along with their weakness and their musical theme
Test Phase 1	Find the essence hidden in one of the three meteors.	Giving away the solution via the music, see if subject goes to the right place.
Test Phase 2	Collect coins in crates and barrels.	Foreshadowing the contents of the container (good/bad) via musical sound effect.
Interview Phase 1	Answer questions about playthrough .	Examine subjects thought process for finding the skull and destroying containers with coins.
Interview Phase 2	Listen to music, pick association from images	See if participants can recognise or identify the music post-playthrough.

7.2.2 Test 6: Local

While it was expected that the majority of participants would come from the online test due to accessibility, it was found relevant to do additional testing locally, since it would then be possible to collect observation and interview notes. This way, it was possible to delve into topics that lie beyond the in-game questionnaire This test was the first final test build and was intended to not only play the final build but also discuss it afterwards with the participants. The purpose of this interview would be to investigate elements that the player might not have answered in the questionnaire or if the conductor would discover any odd behaviour during gameplay compared to other participants.

7.2.2.1 General Information

This test was conducted at several locations over two days (the 04/05-2017 to 05/05-2017) at Aalborg University Copenhagen. In-game metrics, observations and questionnaire data was captured from a total of 10 participants

(5 in each group). Test participants were sequentially (by the online system, according to time of participation) divided into either the musical hint group, or the control group. The median play duration was 36 minutes. The test procedure can be seen in figure 7.9.

For the hint group, the correct region in Test Phase 1 was distributed as follows: 1 (Forest), 3 (Desert), and 1 (Volcanic). Enemy counts were 3 and 2 for the Evil Head and Spitter respectively.

7.2.2.2 Results and Discussion

Results from Test Phase 1 (figure 7.10) show that the control group actually performed slightly better than the music hint group, with 2 participants getting it in first try and one in second. This is in contrast to the hint group, where only 1 participant got it in the first attempt, 3 in third and 1 in fifth, notably more scattered than the control group, and quite contrary to the initial hypothesis that the musical hint group would perform better.

Attempts	*Control mode*	*Hint mode*
1	2	1
2	1	0
3	2	3
4	0	0
5	0	1

Table 7.10: The participants attempts before finding the essence.

When calculating scores for each group, this observation is further supported. The musical hint group has an average score of 3.6 (std: 1.5), while the control group scored an average of 4.2 (std: 1.3).

When correlating the captured questionnaire responses of the two groups separately, some interesting findings surface. In the hint group, their self-assessed confidence when heading to find the evil essence in Test Phase 1 correlates negatively with the amount of time they have listened to the enemy theme of the correct region (r: -0.86, p: 0.063), and the amount of time they have listened to the theme of the correct region (r: -0.88, p: 0.051) almost to statistical significance. This shows that the higher the confidence, the less amount of time has been spent listening to the correct themes of the solution in Test Phase 1. This indicates either that the musical hint has contributed negatively to their decisions, which is odd, considering it is the only hint they have to perform better than random chance; or that the more confident participants used other elements to guide their decisions than the music, thus reporting higher confidence since they were more sure that their other intuition was right. The control group also showed a significant correlation (r: 0.98, p: 0.003) between how experienced participants considered themselves as gamers, and how long they had listened to the enemy theme of the solution to Test Phase 1, which surprisingly shows that more experienced players had spent more time in the dream world fighting the monsters. This might suggest that experienced gamers were more interested in the combat situations, or perhaps playing with- or exploring the enemy behavior.

Listening Test In terms of the following listening test there is a remarkable 23,4% difference between the groups when it comes to the self-report of remembering the themes. Of the total votes of remembering themes, the control group voted 86.7%, while the hint mode is 63.3%. This seem to accord with the generally better performance in the control group. However, when it comes to actually identifying the themes, there were only 6.7 % correct answers in the control group, contrasting the 13.3% in the hint group. These numbers are remarkably low, and they seem to contradict the high amount of themes remembered. A notable mention is the fact that the Spitter theme was correctly identified by all 5 participants in the control group, while in the hint group, only 2 participants correctly identified this theme.

Participant	*Good Containers Destroyed*	*Bad Containers Destroyed*
2	47.37%	52.63%
5	52.17%	47.83%
7	47.37%	52.63%
9	56.25%	43.75%
10	42.86%	57.14%
Average	49.20%	50.80%
Standard Deviation		5.14%

Table 7.11: Percentage amounts of destroyed crates for the control group.

Container destruction success percentages from Test Phase 2 are shown in figures 7.11 and 7.12, for the control group and musical hint group respectively. Both groups seems to have similar performances, with an average of 21.4 crates destroyed per participant in the hint mode and 19.6 in the control group, very similar variances and both groups reported that there might have been visual differences on the crates even though there were none. A success rate of around 50% for both groups indicates that the musical hints did not have an influence on the participants, and the coin collection was left to random chance, or trial and error for both groups.

Participant	*Good Containers Destroyed*	*Bad Containers Destroyed*
1	50.00%	50.00%
3	47.37%	52.63%
4	53.85%	46.15%
6	52.94%	47.06%
8	47.37%	52.63%
Average	50.47%	49.53%
Standard Deviation		3.03%
Average Proximity Time		224 seconds
Average Proximity Periods		8.3 seconds

Table 7.12: Percentage amounts of destroyed crates for the hint mode group.

Observation highlights While observations were captured for both groups, notes from participants in the control group are not included here, since we're looking for pointers to how well music performed as the sole information giver in the test setup.

There were three (out of five) in this group who were considered 'avid listeners' (with a big interest in sound and music), or who reported themselves that they paid much attention to the music during their playthrough. Interestingly, none of these seemed to catch on to the information that the music was giving them. P1 even said *"It was confusing that there were no hints towards what core was the right one"*, indicating that he did not at all include the music into his thought process for going to find the core with the evil essence. However, P1 did report that he recognized both the fact that each area had its own musical theme, as well as each enemy having its own theme. Instead, P1 picked an enchantment based on the assumption that enemies would belong in specific areas - an issue that has been observed persistently throughout the entire development, even though countless measures has been taken to explain to the participants (directly in the game) that it is **not** known where each enemy is located, and each type could be anywhere. In other words, P1 picked the fire enchantment when heading to the forest, because fire beats nature (which is correct), but fire does not necessarily beat the enemy type that is located in the forest (which could just as well be water-based or fire-based - making the fire-enchantment useless). This behavior is shared by P3 and P4 as well, who are both also considered avid listeners. However, this does not necessarily mean that they did not listen to the music throughout the playthrough, as all reported paying much attention to the music (e.g. P4: *"[About paying attention to the music] I did that a lot. Really cool for the mood, nice combat modes, etc."*).

Most seem to have positive feedback about the music, although for some it seemed to not stick in memory (e.g. P3: *"When I was in the desert i remember thinking 'Wow this is awesome music!', but now I actually don't remember how it sounded."*). Going back to the topic of so-called 'obvious instrumentation' as discussed in section 3.0.3, it seems to have worked at least to some extent, since P6 mentions the following when asked about his choices in the music association phase of the questionnaire: *"I would have guessed the same even if I didn't just play the game"*. It has occurred several times throughout previous preliminary tests that participants reported that, for instance *The forest music sounds like forest music"*, or compared the music to the music of big, well-known games like World of Warcraft (Entertainment, 2004).

Similarly to what was observed in test 5 (section 7.1.5), one participant replied 'No' to remembering all of the themes in the listening test, which was noted by the conductor as following: "Participant seems confident when choosing correctly for associations with the music, but always says he doesn't remember from the game". Why this happens can only be speculated. A theory is that the obvious instrumentation has helped in creating a foundation for the association of the themes, meaning that even though the participant did not remember the themes from the game, it was still intuitive to him which region each theme belonged to.

In regards to the crates and barrels, all participants in the non-control group mentioned looking for visual clues for figuring out which containers were explosive, and which contained coins. For instance, P1 mentions *"I thought it had something to do with being red on the top, and maybe some [containers] were smaller. At least a difference in size"*, and P6 reported looking for colors as well. This suggests that even though the only way to tell the difference between the containers were through audio, visual methods are clearly more natural to the players of this group.

7.2.3 Test 7: Hint Only Test

Due to the surprising discoveries in Test 6, and while waiting for more participants in the online test, it was decided to try an approach similar to in Test 2, where participants are explicitly told to listen for the music, as it will provide valuable information. This was with the aim of seeing if there were drastic improvements in performance when explicit focus is put on the music. Additionally, if this test would still show weak results in regards to performance, doubt can be cast on the validity of the test setup of this entire study. Weak results could potentially indicate that the way that it was chosen to test the problem statement might be suboptimal, and require additional work. In contrast to Test 2, where only a slight hint was given that the music played an important role, this time it was decided to go to greater lengths. When starting the test, the conductor would explain the following to the test participant:

> "Pay attention to the music, and try to recognize it. You will need to use it later."

After returning from the dream, initiating Test Phase 1, the conductor would then explain the following:

> "The music will provide you with information."

This way, it can reasonably be assumed that at least some attention is drawn to the music at the most critical time.

7.2.3.1 General Information

This test was conducted at several public locations at Aalborg University Copenhagen in the period 08/05-2017 - 09/05-2017. In-game metrics, observations and questionnaire data was captured from a total of 10 participants. The median play duration was approximately 27 minutes.

7.2.3.2 Results and Discussion

Test Phase 1 Participants of this test scored an average of 5.2 (std: 2.658320272), which is the highest average score seen at given point. 50% of the participants got the correct solution to Test Phase 1 in the first try, which is the best result so far, with the largest participant count, which indicates that the explicit guidance has indeed had an effect. Those who did not get it in the first attempt got it in their second attempt, except for P4 and P9 who got it in their third.

A statistically significant negative correlation was found between the self-assessed gaming proficiency and amount of time listened to the correct enemy theme of Test Phase 1 (r: -0.94, p: 0.00005), while a similar negative correlation was found for gaming proficiency and Test Phase 1 region theme listening time (r: -0.576, p: 0.081), although not statistically significant within a 5% significance level. This shows that participants who rated themselves as experienced gamers spent less time listening to the important themes, which might be a sign of gameplay proficiency. This is further backed by a significant anti correlation between self-assessed gaming proficiency and total playtime (r: -0.94 ,p: 0.00006).

While the scores for the participants were generally higher than previously observed, when looking at the 5 participants who scored 6 or higher, all but one reported high knowledge of the process of game development, which most likely stems from testing at a university with a high population of technologically adept individuals.

Listening Test Of the 10 participants 15% correctly identified the correct themes of Test Phase 1, and 65% reported remembering the themes, which is also the best seen so far. This is relatively in accord with the higher scores and generally better performance.

Test Phase 2 With an average of around 70.71 good containers destroyed, participants of Test 7 (seen in figure 7.13) performed quite well, and as expected. Additionally, they only destroyed an average of 15.3 containers, compared to 19.6.

Participant	*Good Containers Destroyed*	*Bad Containers Destroyed*
1	75.00%	25.00%
2	71.43%	28.57%
3	66.67%	33.33%
4	58.82%	41.18%
5	84.62%	15.38%
6	83.33%	16.67%
7	76.92%	23.08%
8	69.23%	30.77%
9	73.68%	26.32%
10	47.37%	52.63%
Average	70.71%	29.29%
Standard Deviation		11.17%
Average Proximity Time		150 seconds
Average Proximity Periods		8 seconds

Table 7.13: Percentage amounts of destroyed crates.

It can safely be said that these findings suggest that explicitly informing participants of the fact that the music provides information does indeed positively affect the performance.

7.2.4 Test 8: Online

This test can be considered 'the main test', as it was expected to provide the largest amount of participants. The test procedure was identical with that of tests 6 and 7 (shown in figure 7.9). While observations and interview notes were captured in previous tests, this was not possible for this test, as participants from all over the world could play the game in the comfort of their own homes. As such, it was also impossible to ensure that participants would even listen to the game while playing (and not having the system volume muted, or speakers shut off), however, in this case we are left to trust the questionnaire responses about audio configurations during testing. Additionally, participants were prompted by the intro screen (described in section 5.16) that it is highly suggested to play with good headphones or speakers.

7.2.4.1 General Information

This test was conducted online in the period 02/05-2017 to 25/05-2017. In-game metrics and questionnaire data was captured from a total of 31 participants with a median play duration of 38 minutes.

Due to unfortunate corruption of data, some data submissions (not counted) were corrupted, and could thus not be used for further analysis, which also slightly messed up the system on the server that took care of putting participants into the two groups. As such, this test has 17 participants in the hint group, and 14 participants in the control group.

The distribution of correct themes in the hint group is shown in table 7.14.

Table 7.14: Occurance of region and enemy themes in Test Phase 1

Correct Region	*Count*	*Occurance (%)*
Forest	8	47.06%
Desert	6	35.29%
Volcanic	3	17.65%
Correct Enemy	*Count*	*Occurance (%)*
Spitter	6	35.29%
Crawler	8	47.06%
Evil Head	3	17.65%

7.2.4.2 Results and Discussion

Test Phase 1 The average scores for the participants in the control group and hint group were calculated as 5.6 (std: 2.84) and 4.8 (std: 2.76) respectively. Only 3 participants in the hint group got it first time, while 5 got it first time in the control group. In fact, in most areas, the control group seems to have performed better than the group that received musical hints, which directly contradicts the initial hypothesis. However, a two-tailed

A notable mention is that multiple participants scored 10, which is the highest score seen so far, while a single participant scored 0. However, those who scored 10 also reported working with audio on a professional, or experienced level, meaning that they might have a natural interest in game audio, and thus an increased focus on the audio of the experience. This is supported by their self-assessed knowledge about the process of game development which comes in at around 0.8 (on a scale from 0-1) for both of them.

There was found no significant correlations between any of the data from the questionnaire responses or other in-game metrics within the hint group, however, in the control group there was found several. A significant correlation (r: 0.53, p: 0.05) was found between play duration and score, which shows that the longer time people played the game, the better score. This is in part counter-intuitive, as one could expect that eg. more failed attempts would lead to generally longer play sessions. However, one should also not forget that people have vastly different ways in which they like to play games, so it is possible that it is by pure coincidence. Additionally, a statistically significant correlation between self-assessed gaming proficiency and confidence when heading out to find the evil essence was found (r: 0.77, p: 0.001). This means that players who consider themselves inexperienced, also were not as confident in where to go in Test Phase 1, which makes sense, given that they did not receive any information about where to go. Meanwhile, experienced gamers seem to trust their intuition based on other things, given no information.

Listening Test The control group had 11.1% correct pick in themes, and the hint group had 11.8%, which is the lowest of all hint groups so far. Surprisingly, and in contrast to the low amount of correct identifications, the hint group reported remembering the themes 78.4% of the time, which is the highest percentage, with the biggest sample size of all the tests. Likewise, the control group reported remembering the themes 76.7% of the times. Once again, this is found extremely odd, as if participants did in fact remember the themes, it is peculiar that they do not correctly identify them as well.

Test Phase 2 In Test Phase 2, when participants had to destroy containers to get coins, small differences could be found between the two groups. As evident from figures 7.3 and 7.4 the percentages of good containers destroyed of the control group generally lies around 50% or lower, with a few exceptions, whereas in the hint mode, the majority lies at 50% or higher. When running the Mann-Whitney Wilcoxon U (MWW) test on the container success percentages of the two groups, no statistically significant differences are found. However, when comparing the amount of containers destroyed for each group, a one-tailed MWW test yields a non-significant (within a 5% significance level), but non-trivial result (p: 0.057). This shows that the group receiving the musical hint generally didn't destroy as many containers as the control group, which indicates that the hint group used the musical hints to their advantage.

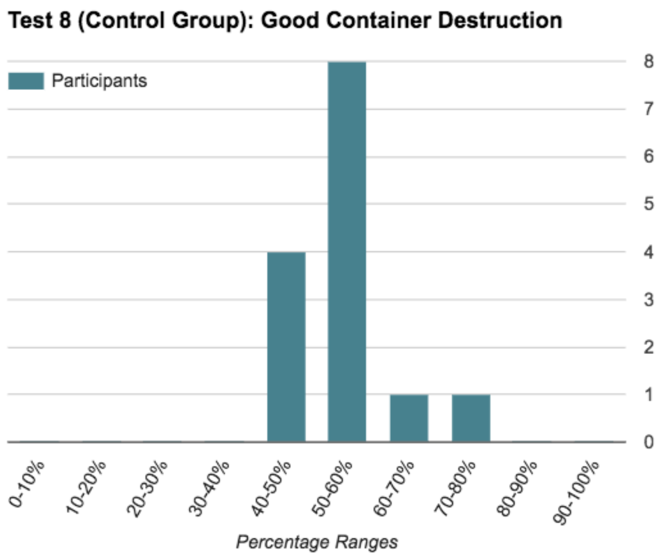


Figure 7.3: Good container destruction histogram (control group)

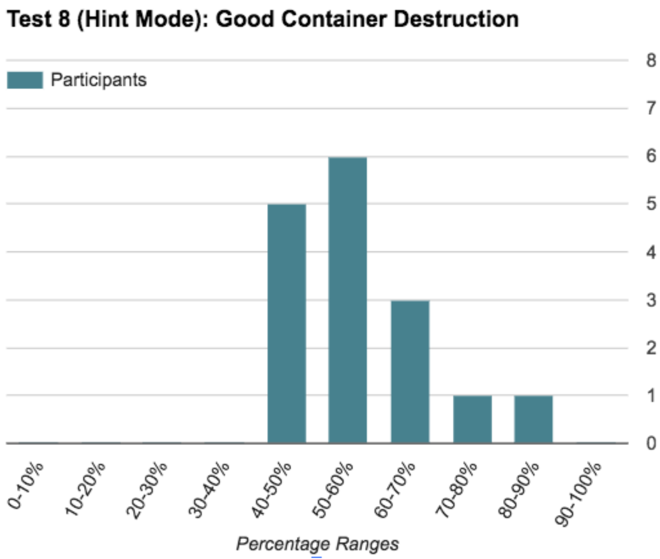


Figure 7.4

7.3 Overall Results

This section presents some general results from comparing the results of each individual test. As there are many results from this test that cannot be easily represented using tables or diagrams, please refer to the appendix for directions to all extra figures and spreadsheets.

7.3.1 Time per Playthrough

The participant playthrough durations were fairly varied. Testing a problem statement that required several training phases for establishing connections with associative themes also contributed to what in the end became a 20-40 minutes test duration.

Test No.	Minutes (avg)	Standard Deviation	Participants
8	39.17	17.55	31
7	26.34	7.88	10
6	35.21	9.43	10
5	31.82	5.68	5
4	25.59	1.78	3
3	38.72	0.00	1
2	22.07	1.90	5
1	17.22	6.43	5

Table 7.15: Average playtime durations, standard deviations, and the amount of participants for that test.

In figure 7.15 a list of playthrough durations, their standard deviations and the participant count is shown. Generally the tendency is that the gameplay durations gets longer with the test iterations. A notable mention is P19 of Test 8 (the online test), who had a playthrough duration of around 46 hours. This is obviously an oddly long duration, and it is expected that the participant had minimized the game at some point, only to return to it later. Because of the way that the playthrough duration is saved, this resulted in a 46 hours playthrough which had to be accounted for during analysis.

7.3.2 Image and Music Associations

Through the in-game questionnaire, the participants were asked to select an image based on the music that they are currently hearing. This was a way of checking whether the associations had successfully been established during the playthrough. In table 7.17, an overview of correct response rates is shown.

Test No.	Comparisons		Total Votes	
	Control	*Hint*	*Control*	*Hint*
8	12.12%	8.33%	66	84
7		15.00%		60
6	6.67%	13.33%	30	30
5		13.33%	0	30
4		22.22%	0	18
3		0.00%	0	6

Table 7.16: Percentages of participants replying a correct answer in the image listener test.

Since the listening tests in tests 1 (7.1.1) and 2 (7.1.2) were of a slightly different format (physical edition with variations from the final version) they are not included in the table.

Due to unforeseen complications with recording a state directly from Wwise, the original music state recordings were unfortunately periodically invalid, however, as for safety measures a music state recorder was also built into the game, and using this it was possible to retrieve the music states during these images.

Test No.	Comparisons		Total Votes	
	Control	*Hint*	*Control*	*Hint*
8	71.21%	77.38%	66	84
7		65.00%		60
6	86.67%	63.33%	30	30

Table 7.17: Percentages of participants remembering correctly.

7.3.3 Playtest Scores

From the scoring board (described in further detail in section 6.1.5) the scores of every test was generated in table 7.18.

Test	Groups Control/Hint	Average Score Control	Average Score Musical Hint
Test 8: Online	14/17	5.642857143	4.823529412
Test 7: Local	5/5	4.2	3.6
Test 7: Hint Only	0/10	#N/A	5.2
Test 5	2/3	2.5	5
Test 4	0/3	#N/A	4.666666667
	0/0	#N/A	0

Table 7.18: Average Scores of participants from each test iteration.

7.3.4 Time of Listening

Due to the nature of WPAG being dependent on how and when the player makes choices, the participants had sometimes longer listening durations on some themes than others. Figure 7.5 shows a numbered heat-map of how many seconds the participants listened to certain music sequences.

	CONTROL GROUP						
Test No.	DESERT	FOREST	NOWHERE	VOLCANIC	EVIL HEAD	CRAWLER	SPITTER
8	148.5	89		105	148.5	89	105
7							
6	172	85	243	133	172	85	133
5	183.5	212.5	298.5	136.5	183.5	212.5	136.5
4							
3							
	HINT MODE						
	DESERT	FOREST	NOWHERE	VOLCANIC	EVIL HEAD	CRAWLER	SPITTER
8	120	106	246	107	120	106	107
7	123.5	89.5	195	112.5	123.5	89.5	112.5
6	137	85	249	121	137	85	121
5	141	114	240	108	141	114	108
4	100	79	196	105	100	79	105
3	117	111	285	125	117	111	125

Figure 7.5: Heatmap of seconds listening to music sequences.

7.3.5 Attempts

As every core interaction was recorded, the participants attempts were counted and drawn in figure 7.6. The vertical axis account for percentage of participants using that amount of attempts rather than numbering, since the groups are of different sizes.

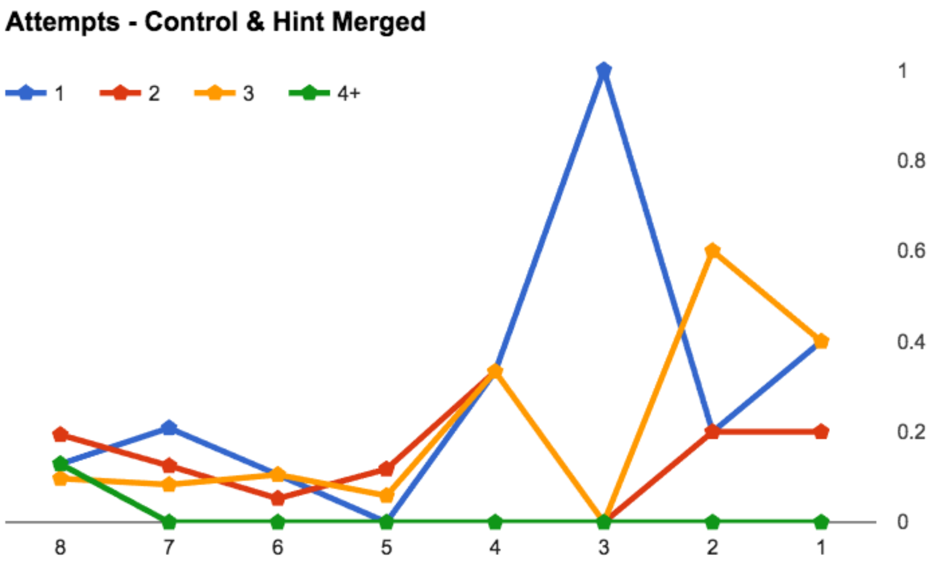


Figure 7.6: Count of attemps in each group.

As there were control groups in many of the tests, the participants were also recorded in comparison to the size of their group and separated in figure 7.7.

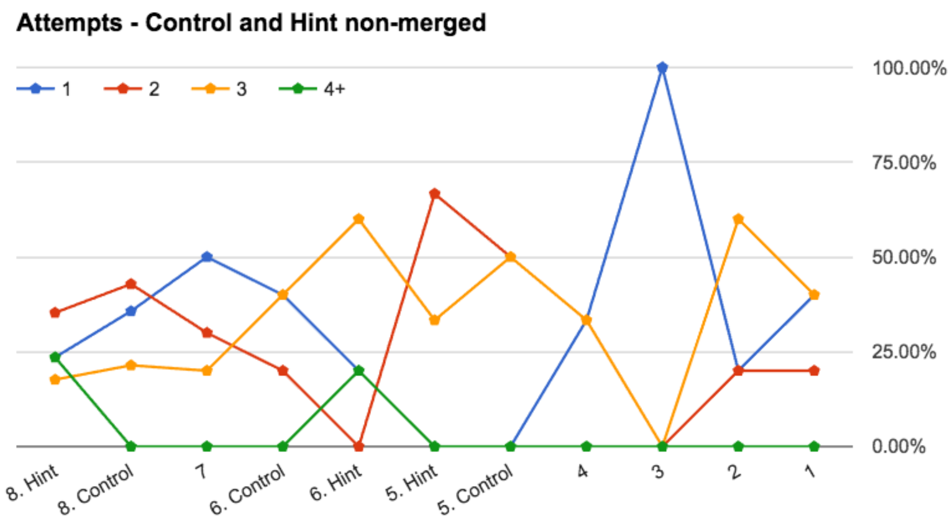


Figure 7.7: Count of attempts in groups, with split control and non-control groups.

7.3.6 Interesting highlights and topics

As test 7.2.2 and test 7.2.4 were using the same build, it was decided to merge the scores of both tests and try to investigate whether there any statistical differences between the groups, seen in figure 7.20.

The *Kolmogorow-Smirnov* test was first used to investigate whether the data is parametric or not. As the data was not parametric, the *Mann-Whitney-Wilcoxon* (MWW) non-Parametric test was used to measure whether the two data sets had any significant difference. Since we're looking for a one-tailed difference (better than control), we're testing against a null hypothesis that there's no difference and left-tailed ranksum was used in Matlab. The p value returns 0.745 which not even close to the 0.05 significance level, and there are therefore no significant difference.

Effect size was measured using the formula below.

$$\frac{Z}{\sqrt{populationSize}}$$

As the effect size returns 0.114 it is hereby categorizes as a small sample size according to listing 7.19.

Effect size	**r**
Small	0.10
Medium	0.30
Large	0.50

Table 7.19: Effect sizes

Score board List with Statistics

Control (c)	*Hint Mode (h)*
8	1
10	8
4	0
10	4
5	5
3	6
7	5
5	1
9	7
3	1
7	5
4	4
2	8
2	5
6	5
2	9
5	8
4	3
4	3
	3
	6
	3
kstest2(c,h)	0
kstest(c)	1
kstest(h)	1
ranksum	0.745
zVal	0.633
effect size	0.114

Table 7.20: Online & Local collective test statistics.

7.3.7 Container Destruction

When comparing the container destruction performance (proportions of crates destroyed per participant) of test 6 (hint mode) and test 7 (hint mode), a two-tailed MWW test reveals a significant improvement with a p-value of 0.007. In other words, the proportions of good containers destroyed were significantly different from test 6 to 7.

7.4 General Discussion

This chapter describes potential biases that were thought of in hindsight after conducting the tests. It then discusses some general findings and elements from the tests.

7.5 Biases

In hindsight there are certain things that could have been set up more optimally. These are either things that might have yielded different results if done differently, or that potentially could have had an impact on the presented results.

7.5.1 Change of heart

Observing participants answering the in-game questionnaire revealed that some participants were trying to select multiple images as the associations to a theme during the listener test. However, the system was built to only allow one selection, and since some participants were reconsidering their choice, it might have been interesting to see what other options they might have chosen. Furthermore, it might have been possible to tell if, for instance, all enemy themes blurred together for the participants, leading to more general conclusions about what themes were ineffective at establishing associative connections.

7.5.2 Degrees of Remembering Music

In the questionnaire the participants were asked whether they remembered a certain theme as follow up question. This was done in a binary way, whee the only options were 'Yes' and 'No'. Having only these options is in reality a bit imprecise, as memory is not simply there or not. Some participants showed signs of being unsure whether they fully remembered the theme, or in other words, some participants 'might' have remembered the theme, but could have ended up choosing 'No'. As such, it might have been better to implement this as a scale, similarly to the question about self-assessed gaming proficiency. This would have allowed participants to enter that they *seem* to recall the theme, although not that much, or for those who definitely remember it, that would be reflected in their response as well.

7.5.3 Interaction

The WPAG game accommodates mostly open-world exploration. In WPAG the third mission was to either find a skull or the essence, when this was achieved the player would be notified by a message to return to the *Wwizard*. However, many participants continued to either combat enemies or visit the remaining cores after completing the quest by finding the essence. One explanation of this could be due to some participant expressing that they were simply interested in completing everything. But despite the notifications on screen when completing a quest, many participants expressed confusion rather than intentional behaviour. From a game design perspective the WPAG also accommodates interaction after the mission objective is reached, by still having enemies to fight, cores to destroy, etc. The participant is therefore only given directions by a small user interface message saying "Return to the Wwizard", whereas the rest of the environment presents even more exploration.

Lastly, some participants expressed that they simply were following the rhythm of their choices in the training phase. One example is participant 20 from test 7.2.4 seen in figure 7.8, who describes the decision making as "I just happened to follow how I had explored the first part of the game". Notable mention is that the *Wwizard* tells the participant "One of the meteors that landed protects an evil core", but if this message was either not payed attention to or in any way missed the participant could believe that the mission was not only finding one core but all.

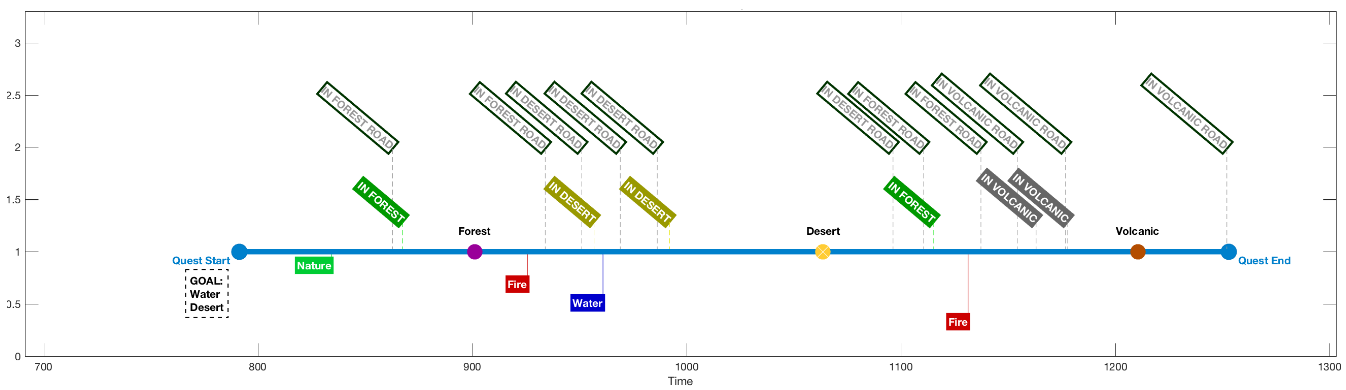


Figure 7.8: Timeline!

7.6 Voice over

Voice over was added to every dialogue to both deliver information through audio as well as visuals, but also as a tool for directing more attention towards listening for information, hopefully directing more attention the music.

It seems that the voice over has contributed positively to the entire gameplay experience. Some of the comments for the *Wwizard* Voice was "I like the voice acting!" (P27 text 7.2.4) or [translated from danish] "Wwizard is really awesome" (P21 in test 7.2.4), and multiple conductor notes described behaviour like "Smiles when hearing the Wwizard voice" (P5 in test 7.2.3). Were the general feedback from participants described the voice over as helpful or commented its aesthetic values in WPAG, none of the captured data seems to be able to support this claim. However, taking into account that 21/45 participants (this data was only captured in test 7, 8 & 9) described themselves as audio professionals or having hobby interests with audio, the participants should be quite qualified to estimate whether it was good in comparison to other professional products. Therefore, by no complaints about the voice over only compliments, the tests suggest that the voice over lived up to professional standards.

7.7 Map distances

The size of the game world was intentionally designed to make sure the player would be exposed to the music themes for a certain amount of time. However, some participants expressed that they thought the distances were too far, like e.g. P4 from Test 8 (7.2.4): "Maybe less walking between the different areas"' or P24 again from Test 8: "the running is extremely slow. It"s killing me!", although it should be noted that P24 did not notice that it was possible to sprint using the shift key. Some participants also expressed distress regarding Training Phase 1 of retrieving objects and that it was at times uninteresting, like P17 (Test 8): "too many fetch quests. The issue might be that they feel too much like fetch quests because they don't have more story dressing".

While this is certainly sub-optimal game design, it was a conscious decision made at development time, in order to force the participants to be in a certain area for a certain amount of time, in order to associate the themes with their element.

7.8 Player Guidance

As with any other game, there is often a fine line between helping too much, and too little. In the test setup of the present study, players were mostly left to themselves, and in the hint group, the musical hints. This is a highly unusual way of guiding the player - something that is rarely (if ever) experienced in other games, and as such, much less expected by the player. This became evident from the improvements in performance when the players are explicitly told that the music will provide information, and suggests that the implementation was not good enough to let the players depend on the music, or that it is such an unusual way of guiding the player, that it is directly counter-intuitive.

7.9 Enemies and Regions

The participants generally also expressed confusion regarding the element type of the enchantments and the region element types. For example some would think the fire enchantment would generally be better in the forest region, however, this would not necessarily the case. Test 4 tries to accommodate this problem with switching to weapons, but the weapons did not seem to be well connected to the enemies which seemed much more obvious with enchantments. One could have changed the regions to being not related to a specific element, but the participants also commented that the fire enchantment might have belonged in volcanic or desert. The problem of participants associating the elements with any region would be there nevertheless, and therefore it was decided to have the approach of having characteristic element types in regions and therefore make the musical association more clear with e.g. desert it's natural association to sitar. However, it should be noted that in order to reduce this issue, extra effort was put on the Wwizards instructions who directly explains that "it could be anywhere!".

As the enemy themes received improvements during test 3, they went from being completely unique and not necessarily having any cultural relations, to actually having the enemy SFX inside the themes. However, doesn't seem to have had a huge influence, and the participants still choose regions in many of the enemy theme questionnaire picks. This suggests that they didn't notice the relation between the original enemy sounds and the enemy sounds implemented in the music. However, by comparison of the different enemy SFX, they do have somewhat similar characteristics, which therefore would only be noticable to practised listeners.

7.10 Scores

Due to the many sources of information, the score system was a very helpful method to get a general performance estimate of the different groups and the individual participant. When looking at how well the participants were at achieving their goals, the scores seemed to represent their performance well while compared to evaluating a participant using all the different sources of data. One example is the well performing test iteration 7, which also had the highest hint mode score at 5.2. The reason the score is not higher in this group is due to their lower performance in the questionnaire.

7.11 Questionnaire Listening Test

By evaluating the participants performance in the listener tests it seems that some participants had a drastic decrease in correct answers compared to the observations. As the questionnaire returns were invalid and a simulation was used to capture the last music states during the questionnaire, further analysis could be attended to whether there are other problems with the questionnaire responses, other whether the questionnaire combination with the simulated music states could be improved.

Even stranger, the observation notes noted the following about P3 in test 5: "Replies 'NO' in each remember question, even though he seems to remember them" (correct identification). This suggests that something is not right, and perhaps the backup system has incorrectly captured the current music states. Although this is undocumented, a test conductor also remembers thinkin that the performance of a participant was surprisingly good, where when later examining the results, he was surprised to see the bad performance.

7.12

Other

Releasing a highly solid product online was not only a priority to the study itself, but also to the collaboration with *Audiokinetic* as an article was released describing both the game and the students, and this would therefore be many of the AAA game companies first impressions of the students and their future work.

Despite having quite a solid build on release, some problems occurred during that either hindered some participants in playing or gave them problems during playthrough. Some *Windows* computers would have installed the user as non-administrator, which would not allow them to upload the gameplay data. The problem did not occur to the students beforehand, because both their computers and University computers were always logged in as Administrator. The issue only occurred to the installer version of the game and was quickly fixed by deleting the installer, however, until then we could easily have given some participants a bad impression of the game. In follow up to this issue, a mail as sent from *Audiokinetic* to all who had signed up, reminding them about bug fixes and the new builds uploaded.

With 31 participants, Test 8 was by far the biggest, however, it was expected to be way larger, especially due to the online announcement. The participants collectively submitted around 21 hours of playtime (not including participant 19 with a play duration of 46 hours), which meant that the amount of data in this test was enormous. With very minor differences in both correct picks, destroyed crates, the support by musical constructs does not seem to have had any effect on the participants. As a matter of fact the control group performed better than the hint group in achieving their goal with 35.71% compared to 23.53%, which suggests that the hint mode only made it more confusing. However, with only 14 participants in the control group and 17 in the hint mode, the sample size is too low to say anything for certain.

The iterative test method, mimicking a game production company's approach, seemed to in general generate incomprehensible amounts of data. Despite the vast amount of analyzed in this study, there are still data to be examined. It can be argued that fewer tests would in the end have been better, however, this would have distanced the project from its initial motivations of helping industry game developers.

8

Future Work

8.1 Confusion

In WPAG some elements seemed to have cause special confusion, which could hopefully in the upcoming version be designed more clearly. Some participants were confused over the fact that when a incorrect core was destroyed, it would just disappear and they would not know what to do like in test 7 where "He smashes it, and then seems a bit confused of why it wasn't there and why there is nothing. " (P4). With this in mind some kind of notification could have supported the player when destroying a wrong cores, since the participant were looking for any hints in this situation and might have missed it initially from the *Wwziard* when not knowing it was important.

8.1.1 Dramatic Memorization

As described in the research section 3, one of the ways to accomplish association to a leitmotif is by making sure the presentation of the leitmotif is when something dramatic or eventful happens. In WPAG this was attempted in the Dream region but wasn't very empathized when first introduced to the regions. One could have tried to introduce something more eventful when first presented the regions, e.g. a cinematic camera event or likewise.

9 Conclusion

The study problem statement was declared as:

> ***"How do players perform in a game scenario when information necessary to successfully progress is provided solely through associative music?"***

with the hypothesis that when split into two groups (one with musical hints, and one without), the group that receives the information through music will perform better.

A total of 8 tests were conducted, and the analysis of the test data did not reveal any significant differences in the player performances between the two groups. In fact, certain test iterations even had the control group perform better than the group that got the hint. It can be concluded that there are many signs that point towards players not being used to consider the music in their decisions, however, even when players are explicitly told to listen for the music, only a slight increase in performance is seen.

However, the study did find some non-trivial results when considering those who were explicitly guided to listen for the music, such as a difference between tests 6 and 7 in crate destruction numbers, where participants in test 7 destroyed markedly fewer crates, indicating successful guidance by the musical hints.

After all, a product of this project is a very successful game experience that received a lot of positive feedback. It was published to the world, and industry-professionals have spoken well about it. In general it seems that the product has been successful, in spite of it being an ambitious product with a 2-man production team in charge of the entire development pipeline. The collaboration with Audiokinetic was strengthened, and they are interested in further supporting the project.

The huge effort to gather a lot of participants with the aim of ending up with a decent sample size was unfortunately unfruitful, leaving the final test with only 31 participants, split in 2 groups. As such, most findings from this study remain statistically inconclusive, and requires additional efforts

Bibliography

AB, E. S. (2015), 'Elias adaptive music'.

Aljanaki, A., Wiering, F. and Veltkamp, R. C. (2016), 'Studying emotion induced by music through a crowdsourcing game', *Information Processing & Management* **52**(1), 115 – 128. Emotion and Sentiment in Social and Expressive Media.
URL: *http://www.sciencedirect.com/science/article/pii/S0306457315000424*

Andersen, M. S. (2016), 'Inside: A Game That Listens', `http://www.gamasutra.com/view/news/277712/Video_Designing_the_audio_for_Inside_a_game_that_listens.php`.

Apple (2017), 'Logic Pro X'.

Audiokinetic (2017), 'A wwise wwizard for wwise adventurers'.

Audiokinetic Wwise (2015), `https://www.audiokinetic.com/products/wwise/`.

Avid (2017), 'Pro Tools - Digital Audio Workstation', `http://www.avid.com/pro-tools`.

Batson, C. D., Shaw, L. L. and Oleson, K. C. (1992), 'Differentiating affect, mood, and emotion: Toward functionally based conceptual distinctions'.

Beedie, C., Terry, P. and Lane, A. (2005), 'Distinctions between emotion and mood', *Cognition and Emotion* **19**(6), 847–878.
URL: *http://dx.doi.org/10.1080/02699930541000057*

Berndt, A. and Hartmann, K. (2008), *The Functions of Music in Interactive Media*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 126–131.
URL: *http://dx.doi.org/10.1007/978-3-540-89454-4*$_1$9

Berndt, A., Hartmann, K., Röber, N. and Masuch, M. (2006), 'Composition and arrangement techniques for music in interactive immersive environments', *Audio Mostly* **2006**, 53–59.

Berndt, A. and Theisel, H. (2008), *Adaptive Musical Expression from Automatic Realtime Orchestration and Performance*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 132–143.
URL: *http://dx.doi.org/10.1007/978-3-540-89454-4*$_2$0

Boltz, M., Schulkind, M. and Kantra, S. (1991), 'Effects of background music on the remembering of filmed events', *Memory & Cognition* **19**(6), 593–

606.
URL: *http://dx.doi.org/10.3758/BF03197154*

Bribitzer-Stull, M. (2015), *Understanding the Leitmotif*, Cambridge University Press.

Cohen, A. J. (1990), 'Understanding musical soundtracks', *Empirical Studies of the Arts* **8**(2), 111–124.

Cook, Nicholas, . and Everist, M. (2001), *Rethinking music*, reprinted with corrections edn, Oxford Oxford University Press. First published: c1999.

Cross, I. (2005), 'Music and meaning, ambiguity and evolution', *Musical communication* pp. 27–43.

Csikszentmihalyi, M. (1991), *Flow: The Psychology of Optimal Experience*, Harper Perennial, New York, NY.
URL: *http://www.amazon.com/gp/product/0060920432/ref=si3*$_r$*dr*$_b$*b*$_p$*roduct*/104− 4616565 − 4570345

DICE, E. (2008), 'Mirror's edge'.

Dog, N. (2011), 'Uncharted 3: Drake's deception'.

Eladhari, M., Nieuwdorp, R. and Fridenfalk, M. (2006), The soundtrack of your mind: mind music-adaptive audio for game characters, *in* 'Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology', ACM, p. 54.

Entertainment, B. (2004), 'World of warcraft'.

Fireflight Technologies (2015), 'FMOD Studio', `http://www.fmod.org`.

Gabrielsson, A. (2002), Emotion perceived and emotion felt: Same or different?, *in* 'Musicae Scientiae [Special issue 2001-2002], 123-147'.

Giacchino, M. (2009), 'Main Theme - Disney Pixar's Up', `https://cdn.instructables.com/ORIG/FSZ/RBOU/IDOVPLP4/FSZRBOUIDOVPLP4.pdf`.

Gomez, P. and Danuser, B. (2007), Relationships between musical structure and psychophysiological measures of emotion, *in* 'Emotion [2007, Vol. 7, No. 2], 377-387'.

Green, J. (2010), 'Understanding the score: Film music communicating to and influencing the audience', *The Journal of Aesthetic Education* **44**(4), 81–94.

Hoover, T. (2009), *Keeping score: Interviews with today's top film, television, and game music composers*, Nelson Education.

Inception Cinematic Trailer (n.d.), https://www.youtube.com/watch?v=66TuSJo4dZM.

Jack Grillo, P. L. (2015), 'Scoring Dynamic Combat with Procedural Music', http://tombraider.tumblr.com/post/131986215580/dev-blog-scoring-dynamic-combat-with-procedural.

Jørgensen, K. (2006), 'On the functional aspects of computer game audio'.

Juslin, P. N. and Västfjäll, D. (2008), 'Emotional responses to music: The need to consider underlying mechanisms', *Behavioral and Brain Sciences* **31**, 559–575.

Kastbauer, D. (2017), *Game Audio: Tales of a Technical Sound Designer*, Vol. 1, LostChocolateLab.

Kramer, L. (2002), *Musical meaning: toward a critical history*, Vol. 1, Univ of California Press.

Kungel, R. (2008), *Filmmusik für Filmemacher*, Dpunkt. Verlag, Reil, Germany.

Livingstone, S. R. and Brown, A. R. (2005), Dynamic response: real-time adaptation for music emotion, *in* 'Proceedings of the second Australasian conference on Interactive entertainment', Creativity & Cognition Studios Press, pp. 105–111.

Lucas, G. (1977), 'Star wars'.

Lynggaard Olsen, J. and Maretty Sønderup, M. (2015), 'Design and evaluation of the affective composing model'.

Lynggaard Olsen, J., Maretty Sønderup, M., Nygaard, M. and Korinek, J. (2015), 'The influence of dynamic music in THINQ'.

Nintendo (1997), 'Pokémon'.

Paulus, I. (2000), 'Williams versus wagner or an attempt at linking musical epics', *International Review of the Aesthetics and Sociology of Music* pp. 153–184.

Phillips, W. (2014), A composer's guide to game music, The MIT Press.

Schmid, J. (2016), 'Wwise Tour 2016 - Playdead INSIDE (2 of 3) - Voice', https://www.youtube.com/watch?v=gRF8Gt5hys4.

Scirea, M., Cheong, Y.-G., Nelson, M. J. and Bae, B.-C. (2014), Evaluating musical foreshadowing of videogame narrative experiences, *in* 'Proceedings of the 9th Audio Mostly: A Conference on Interaction With Sound', AM '14, ACM, New York, NY, USA, pp. 8:1–8:7.
URL: *http://doi.acm.org/10.1145/2636879.2636889*

Tan, S.-L., Spackman, M. P. and Bezdek, M. A. (2007), 'Viewers' interpretations of film characters' emotions: Effects of presenting film music

before or after a character is shown', *Music Perception: An Interdisciplinary Journal* **25**(2), 135–152.
URL: *http://mp.ucpress.edu/content/25/2/135*

Unity3D (2015), `https://unity3d.com`.

Universal, V. (1999), 'Counter strike'.

Wingstedt, J. (2004), Narrative functions of film music in a relational perspective, *in* 'ISME 2004, 26th International Society for Music Education World Conference, 11-16 July 2004, Tenerife, Spain', International Society for Music Education.

10 Appendix

10.1 Links to Data Sets

To simplify the large amount of data, a online folder have been set up where datasets, figures, etc. can be access all in one place.

https://drive.google.com/open?id=0B5bhO5IOfjg8bTFZbURDeV9fX1k

The online test 8's released build can be accessed either through the article at this link: https://blog.audiokinetic.com/a-wwise-wwizard-for-wwise-adventurers/

or directly here: https://drive.google.com/drive/folders/0B-fmACLM8vz-QWtIaDlpbkE5VTQ