Tales of Fear: Impact of Non-Diegetic Sound in a Horror

Pedro Amaral Soares

Thesis to obtain the Master of Science Degree in

Information Systems and Computer Engineering

Supervisors: Prof. Carlos António Roque Martinho
Prof. Sandra Pereira Gama

Examination Committee

Chairperson: Prof. António Manuel Ferreira Rito da Silva
Supervisor: Prof. Carlos António Roque Martinho
Member of the Committee: Prof. Licínio Roque

November 2021
Acknowledgments

Firstly, I want to acknowledge my dissertation supervisors Prof. Carlos Martinho and Prof. Sandra Gama for believing in this project from the start, for all of their insight, support, patience, and sharing of knowledge that has made this Thesis possible. I also want to thank the great feedback and references that Prof. Licínio Roque shared with me, after my presentation of this project’s original proposal.

I want to thank my parents for always being supportive, caring, and present for the good and bad times. You’ve relentlessly encouraged me, since day one, to keep working to achieve my dreams, and I feel truly privileged to have you both. The same words apply to my brother, who has always been my greatest reference of who I aspire to be to others. I couldn’t ask for a better friend.

As this dissertation symbolizes the end of a very special era of my life, I take this opportunity to shed light on everyone that was part of my journey at LabJogos and GameDev Técnico. Each one of you helped me grow as a person and aspiring game developer.

Last but not least, I want to thank all my friends, who not only make life much more enjoyable and special but also did not hesitate in helping me during this project and bringing me extra testers. I feel very grateful and appreciate the kind gesture of every single person who decided to participate in my testing phases.
Abstract

One of the biggest challenges of designing horror games is to produce an experience capable of adequately triggering fear in any player. In particular, the sound is a facet that can have a strong emotional impact in games, especially in this genre. Although there is a considerable amount of works regarding horror techniques in media and the use of sound to enhance its experience, there is yet room to further explore the impact of background audio (including soundtrack music) that changes throughout a horror game, compared to the use of the same audio track during an entire walkthrough. To investigate this problem, we created hypotheses to verify if using multiple audio tracks in a horror game leads to stronger affective reactions (regarding both self-perceived emotions and physiological signals) while listening to different tracks appropriate to the desired state of each game section, and, additionally, if the Heart Rate (HR) is a good physiological metric to reflect the players’ anxiety. To test our hypotheses, we developed a first-person horror video game, able to change its audio generation based on the desired mental state in each section. Users played Version A (default), with different non-diegetic audio tracks according to the desired state in each section, or Version B, always with the same neutral audio track. Results show that Version A triggered more intense positive emotions and promoted a more accentuated change of self-perceived tension and HR between game sections. However, no correlations were found between the evolution of self-perceived tension and HR.

Keywords

Tension, Emotion, Soundtrack, Heart Rate, Horror, Video Games.
Resumo

Um dos maiores desafios no desenvolvimento de jogos de terror é produzir uma experiência capaz de desencadear medo de forma adequada em qualquer jogador. Em particular, o som é uma componente que pode ter um forte impacto emocional, especialmente neste gênero. Embora tenham surgido estudos sobre técnicas de terror e áudio em media, ainda há muito por explorar sobre o impacto do som de fundo (incluindo música) que muda ao longo dum jogo de terror, em comparação com o uso permanente da mesma faixa durante toda a sessão. Assim, criámos hipóteses para verificar se o uso de várias faixas de áudio num jogo de terror conforme o mais apropriado para cada secção leva a reações afetivas mais fortes (tanto em relação às emoções e tensão percecionadas pelos utilizadores, como aos sinais fisiológicos), e, adicionalmente, se a frequência cardíaca (HR) permite definir a ansiedade dos jogadores. Para testarmos as hipóteses de estudo, desenvolvemos um jogo de terror em primeira pessoa, que altera a sua geração de áudio com base no estado mental desejado em cada secção. Os utilizadores jogaram a Versão A, com diferentes faixas de áudio não-diegéticas ao longo da experiência, ou a Versão B, que usa sempre a mesma faixa neutra, de início ao fim. Os resultados mostram que a Versão A desencadeou emoções positivas mais intensas e promoveu uma mudança mais acentuada de tensão relatada e frequência cardíaca entre as secções de jogo. No entanto, não foram encontradas correlações entre a evolução da tensão relatada e HR.

Palavras Chave

Tensão, Emoção, Trilha Sonora, Frequência cardíaca, Terror, Jogos de vídeo.
# Contents

1 Introduction .................................................. 1
  1.1 Motivation and Problem ........................................ 3
  1.2 Research Hypotheses .......................................... 4
  1.3 Objectives / Expected Contributions .......................... 4
  1.4 Document Structure .......................................... 5

2 Background .................................................. 7
  2.1 Concepts of fear, terror, horror, anxiety and suspense ......... 9
  2.2 Effects of Film Soundtrack Music ............................. 10

3 Related Work ................................................. 13
  3.1 Horror Media and Techniques .................................. 15
    3.1.1 Manipulation of emotions in horror media ............... 15
    3.1.2 Impact of Facets Differences in Horror Games ........... 18
  3.2 Biofeedback and Affective Gaming ............................ 19
    3.2.1 Horror Gaming Experiences with Biofeedback ........... 19
  3.3 Discussion .................................................. 22

4 System Overview ............................................... 25
  4.1 Presa ...................................................... 27
    4.1.1 Game Choice ........................................... 27
    4.1.2 Game Concept .......................................... 28
    4.1.3 Game Loop ............................................. 28
    4.1.4 Character Actions ...................................... 28
    4.1.5 Game Sections ......................................... 29
    4.1.6 Non-Diegetic Audio Tracks ............................... 31
    4.1.7 Creatures .............................................. 31
    4.1.8 Tracker ................................................. 32
  4.2 Bio Framework .............................................. 33
    4.2.1 Heart Rate ............................................. 33
List of Figures

4.1 Screenshot representative of the game - Gameplay Snippet of Presa. ............... 27

4.2 Scene displaying the controls layout for the character actions. ..................... 29

4.3 Tutorial scene, where the players explore the essential character actions. ......... 29

4.4 Screenshots representative of different parts of the game - The first line (a, b, c) contains key moments during Sections 1 and 2, which end at the first appearance of the exit gate; the second line (d, e, f) represent Section 3, 5 and 6; The third line (g, h, i) depicts the last two sections, filled with various lost headstones and the appearance of a dangerous creature. Section 4 begins with a quick appearance of a creature and continues with the regular maze paths presented during the first two sections. ............... 30

4.5 Illustrative diagram of the anxiety evolution design for the game, with the non-diegetic audio track selected for each section, for the default version of the game. The diamond shapes represent the collection of each orb. The square shapes represent the moments where the player finds one or multiple creatures (A - quick jump-scare event, where a harmless creature suddenly appears and disappears; B - harmless creature sitting next to an orb; C - static harmless creatures; D - dangerous creature that roams in an area that has to be traversed by the player). ....................... 32

4.6 Tracker detecting the presence of an enemy, battery and orb; the tracker also reveals the direction to reach the gate, indicated by the green arc. ....................... 32
4.7 System architecture, with a module responsible for physiological data and another for
the game. From left to right, it is composed by user raw data (analog Blood Volume
Pulse (BVP)), Bitalino with a PulseSensor attached (a), BioSPPy (b) and data treatment
(c). First, the user produces analog signals which are measured with the Bitalino. These
signals are then digitized by OpenSignals (r)evolution and fed to the BioSPPy which filters
them, performing R-peak detection for the computation of the instantaneous HR. After
that, via UDP sockets (d), the script is able to treat this information and send to the game
and receive gameplay data logs, to synchronize relevant timestamps with the obtained
physiological signals.

5.1 Values of self-reported tension for each audio track per player.

5.2 Main board of the (r)evolution Plugged Kit from Bitalino with one sensor cable connected
to it, for Pulse Sensor.

5.3 Pulse Sensor.

5.4 Boxplots of positive affect for dynamic (A) and constant background audio (B) versions.

5.5 Boxplots of negative negative affect for dynamic (A) and constant background audio (B)
versions.

5.6 Boxplots of accumulated absolute tension change during the game, for Version A and B.

5.7 Tension per Section for Versions A and B.

5.8 Average change of tension between game sections, for Versions A and B.

5.9 Boxplots of accumulated absolute change of average Heart Rate (HR) along the game
sections, for Version A and B.

5.10 Boxplots of normalized average HR per section, for Version A and B.

5.11 Average HR and Normalized average HR, side by side, per section, for Version A and B.

5.12 Scatter plots for self-reported tension vs average HR per section change.

D.1 Normality, Mann-Whitney and Independent Samples tests for versions A and B, for av-
erage self-reported tension difference between each pair of successive game sections
(from $s_{1,2}$, which is the first two sections, to $s_{7,8}$).

D.2 Normality tests for differences of normalized self-reported tension and average HR be-
tween sections ($s_{i,i+1,\text{norm}}$ and $hr_{i,i+1,\text{norm}}$).

D.3 Pearson and Spearman’s Rank Correlations among differences of normalized self-reported
tension and average HR between sections.
List of Tables

3.1 Overview of biofeedback horror-related studies & games . . . . . . . . . . . . . . . . . . . . 22
# Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>BVP</td>
<td>Blood Volume Pulse</td>
</tr>
<tr>
<td>CAM</td>
<td>Congruence-Associationist Model</td>
</tr>
<tr>
<td>ECG</td>
<td>Electrocardiography</td>
</tr>
<tr>
<td>EMG</td>
<td>Facial Electromyography</td>
</tr>
<tr>
<td>EEG</td>
<td>Electroencephalography</td>
</tr>
<tr>
<td>GSR</td>
<td>Galvanic Skin Response</td>
</tr>
<tr>
<td>HR</td>
<td>Heart Rate</td>
</tr>
<tr>
<td>HRV</td>
<td>Heart Rate Variability</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquartile Range</td>
</tr>
<tr>
<td>PANAS</td>
<td>Positive and Negative Affect Schedule</td>
</tr>
<tr>
<td>PCG</td>
<td>Procedural Content Generation</td>
</tr>
<tr>
<td>PPG</td>
<td>Photoplethysmogram</td>
</tr>
<tr>
<td>PRV</td>
<td>Pulse Rate Variability</td>
</tr>
<tr>
<td>SC</td>
<td>Skin Conductance</td>
</tr>
<tr>
<td>VR</td>
<td>Virtual Reality</td>
</tr>
</tbody>
</table>
1.1 Motivation and Problem

Video games have been continuously evolving in the last decades and, currently, are able to produce immersive and believable experiences, in a way that no other media form can reproduce, due to its complex interactive nature, by requiring the audience to take an active role in the action. As in the film industry, we can divide video games into categories by genre, including horror, where a main challenge for developers consists in triggering fear and unease. Similarly to how filmmakers “take advantage of the ways spectators draw upon everyday thinking while viewing a film (...) [and hypothesize] what is likely going to happen next”, game designers, especially in survival horror, are required to “know how to elicit the sort of activities and emotional responses that will create the experience they want the gamer to have”, by considering the several ways one can react to each situation [1].

As we observe the graphical and soundscape aspects of video games achieving a peak level, and with the goal of improving horror gaming experiences, a more recent feature is slowly gathering more attention and being studied, to step up, once again, the gaming experience: biofeedback mechanics. A game with biofeedback is enhanced by physiological signals from the player, to change or add game mechanics, or multiple artistic facets.

Liapis et al. considered six different artistic facets in games’ context to guarantee an emergent interactive experience: visuals, audio, level design, narrative, game rules and gameplay [2]. Considering the fact that “(...) sonic dimensions, like music, are highly suitable to take over other senses during the initial direction of selective attention towards congruent perceptions” [3], and that background music promotes strong affective reactions and memorization of key events in media [4, 5], the audio facet presents itself as a key element yet to be fully studied in games and to be further explored with biofeedback, with the goal of influencing and triggering specific reactions or elicit fear, in the context of horror gaming experiences [6–10]. As such, we identify the importance of investigating how, for instance, background audio can change the experience of horror in a game. Klimmt et al. [11] have studied the effects of soundtrack music on the video game experience, including in the genre of horror, by comparing players’ enjoyment of the game with and without background music. However, there is more to sound in horror to assess than testing solely the differences of having the soundtrack muted or not. Furthermore, if we desire to apply biofeedback for audio choices in future work, we need to examine if the Heart Rate (HR), which is the most commonly used physiological metric in this field, can reflect the tension felt by the players in a horror game.

Sound can be divided in two broad categories: diegetic and non-diegetic. In diegetic sound, the source is visible or can be implied in the action of the experience, unlike non-diegetic sound which comes from a source out of context, including background music. As we start observing and identifying how many different approaches horror media can take regarding non-diegetic audio, from the selection of music to the appearance of unknown and unexpected sounds, we are motivated to raise the following
question: How does the use of distinct non-diegetic audio tracks across a horror experience in sync with the level of tension designed by the game creators can have a different impact on the players' affective response and anxiety, compared to the use of a constant audio track? Additionally, we take this opportunity to also question: Is the HR a good physiological metric to reflect the anxiety state of the player?

1.2 Research Hypotheses

To assess this problem, we designed and developed a horror game and a framework to gather in real time the players' physiological signals. Our game is a psychological first-person horror that takes place in a semi-realistic dark environment populated by creatures with a deformed humanoid presentation. The game is able to adapt the background audio according to the game section where the player is located.

We hypothesize that: players have a better horror experience (with higher affect and changes of tension and HR) playing a game that changes its non-diegetic audio according to each game section, compared to a game that constantly keeps the same non-diegetic audio track; the evolution of normalized average HR between consecutive sections correlates to the evolution of normalized self-reported tension between the correspondent consecutive sections. In Chapter 5.3, we explore and subdivide the hypotheses, and share the evaluation process and results.

We crafted and ran two versions of the game, for further comparison in user results, where, in one (Version A), we use the default evolution of desired anxiety, which leads to the dynamic change of background audio among the game sections, and, in the other one (Version B), we permanently use the same background audio track during the entire experience. To prove our hypotheses, players are expected to show a higher self-perceived affect response, accumulated change of tension and HR across the game sections in Version A, as well as a correlation of the changes reported in the tension with the HR between each pair of consecutive game sections.

1.3 Objectives / Expected Contributions

This research aims to understand the impact of non-diegetic audio in the horror experience and analyse if HR is a good physiological metric to define the anxiety state of the player. To fulfill such goal, some steps were set to guide us through the process of this study. Our project aims to contribute with:

- A state of the art survey and analysis about (1) concepts in horror media and techniques to elicit fear, (2) strategies to improve a horror gaming experience and manage the players’ anxiety, by
using sound and biofeedback mechanisms.

- A horror first-person game that adapts its sounds generation depending on the desired mental state.
- A set of physiological data recorded from participants while performing our tests.

Additionally, this work aims to validate the impact of the use of common horror media techniques, by studying the players’ self-perceived level of tension and emotions, as well as their HR, and how these variables change depending on the game events and evolution of the background audio.

1.4 Document Structure

This document is organised as follows: Chapter 2 presents the background of this work, with fundamentals on Horror and Soundtrack Music. Chapter 3 is a description and analysis of other studies that lead to relevant findings in the topics of horror and audio in games and biofeedback. Chapter 4 describes the core components of the implemented system. Chapter 5 is the evaluation of the results where we present the results obtained, accept or reject the proposed hypotheses, and discuss our findings. Chapter 6 presents the conclusions, limitations and future work.
2 Background

<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Concepts of fear, terror, horror, anxiety and suspense</td>
<td>9</td>
</tr>
<tr>
<td>2.2 Effects of Film Soundtrack Music</td>
<td>10</td>
</tr>
</tbody>
</table>
Horror is a complex genre that makes use of multiple concepts relevant to our study and testbed game. In this chapter, we give a brief explanation of fundamentals on horror and soundtrack music.

2.1 Concepts of fear, terror, horror, anxiety and suspense

Human emotions has been a recurring topic in psychology, with several definitions according to the many theoretical models presented. Nowadays, emotions can be conceived of as mental states that are consciously experienced when our well-being is conditioned: for instance, we feel fear when threatened [12], being fear broadly considered as an emotional state. Ekman et al.’s research [13] on facial expressions categorized six basic emotions - anger, disgust, happiness, sadness, surprise and fear - which can respectively lead to physical reactions, including, not only in facial expressions, but also changes in HR, for instance [14,15].

Lazarus’ theory considers emotion as a disturbance that occurs through: cognitive appraisal (assessment of an event, which cues the emotion), physiological response (biological changes such as increased heart rate, triggered by the cognitive reaction) and action, when the individual feels the emotion and chooses how to react (translation of action tendencies to coping activity) [16]. Such manifestations can be easily associated with how we cope with fear, which is often considered, in common sense, as an uncomfortable emotional state. Moreover, as we try to classify emotions within a spectrum, many psychologists and a large body of work divide emotion in two dimensions: valence and arousal. Valence relates to the hedonic value of a mood state, or, in other words, how negative or positive (unpleasant or pleasant) a mood state is, while arousal can be referred to the intensity of an event (relaxing/deactivation or exciting/activation) [17]. Thereby, fear can be instinctively associated to the state of negative valence and high arousal.

Fear, being an emotional state that leads to aversion and agitated behaviours, is often associated with the concepts of terror and horror. Although both involve discomfort, while fear is commonly attributed as a response to a perceived danger, horror is only negative in affect, providing no means of pleasure. “In horror, there is no inherent urge to flee; in fact, the ‘gawking’ impulse would suggest the very opposite” - according to the philosopher C. Solomon, “[...] one can be frozen (or ‘paralyzed’) by fear, but that is when fear becomes horror. Horror involves a helplessness which fear evades” [18]. Horror often encourages fixation and is object-focused, while terror, although connected, refers to the anticipation, being evasive and more situational [7, 18, 19]. In our study, both concepts will be indirectly present, and, to simplify, the term horror will be mainly connected to the genre itself.

A large body of work assesses anxiety, which resembles fear and terror, but refers to a more general, diffuse, objectless [20] and long-term state of distress [21], characterized by worry and tension, not requiring a specific source of threat. According to Freud, anxiety “relates to the condition and ignores the
object, whereas in the word fear attention is focused on the object”. As Bourke [22] mentions, “Freud’s theory does not suggest that the framework of anxiety is devoid of an object, but that the connection between object [or threat] and individual is indirect and distant.”

In the end, fear and anxiety are emotional states hard to define or classify externally, since we assume emotions are feelings, which are states of consciousness, and these require the individual to be aware of their own mental activity [12]. Nonetheless, according to LeDoux, feeling afraid is an additional factor that can help promote survival and occurs in organisms that can be conscious that they are in danger, such as humans. We can refer to fear as the conscious feeling of being afraid, emerging when we know the existence of specific ingredients put together to “compel a certain interpretation of the state we are in”, and anxiety as the “sense of worry or apprehension one has when dwelling on the past and/or anticipating the future, is a variation on this theme” [12].

In this work, as our focus is to create and study a scary experience, the concepts of fear and anxiety will be intrinsically involved recurrently. For the sake of simplicity and due to the ambiguity of soundscapes that can often generate distress without a defined source of threat, we will use ‘anxiety’ as the central definition for the emotional state we desire to elicit throughout the playthroughs, although, naturally, the condition of anxiety might evolve into fear during the experience, due to the visual appearance and confirmation of a real entity that symbolizes threat, or, in other words, the gain of what is painful, like the death that the act of a vicious attack by such entity might provoke [7]. Alongside with ‘anxiety’, the term “tension” will be frequently used as well, to define the same emotional state or degree of distress we pursuit to induce with horror media.

Another term often relevant for this study will be suspens[e, which is commonly referred in fiction media, being a desirable emotional sensation to trigger in any horror experience [7, 23]. Zillman considers it as “an experience of uncertainty regarding the outcome of a potentially hostile confrontation”, “whose hedonic properties can vary from noxious to pleasant” [24, 25], and only possible with the presence of fear, “arguing that concern for the well-being of the narrative characters and anticipation of negative plot resolutions generates the suspense that hooks the viewer” [7].

2.2 Effects of Film Soundtrack Music

Over the years, it has been confirmed by several studies that, if music is present, the film audience will experience stronger affective reactions, often with different interpretations, and improve memory of certain events connected to the soundtrack music, than if music is completely absent [4, 5]. As indicated by one of the main references of our study (Klimmt et al. [11], mentioned in Chapter 3), Cohen [26, 27] proposed the Congruence-Associationist Model (CAM), considered as “the most prominent theoretical account of such influences of soundtrack music on audience responses”. It affirms that music, not only
triggers affective responses in the audience (i.e. emotions), it also has an important cognitive effect, activating knowledge structures independently of what is displayed visually, which will then contribute to the general understanding, interpretation and experience of the film’s story [4, 28]. Since the study of CAM has been successfully applied in film music [4, 29], it can be used similarly as a theoretical foundation for the context of video games.

Indeed, based on the CAM, we can label music as congruent to what is happening in a game situation, considering that, nowadays, audio engines are capable of shifting soundtrack characteristics dynamically so that changing game situations are accompanied by altered instrumentation, rhythm, harmonies, and/or melodicics, to fit the overall narrative setting, mood, and visual aesthetics of the project [11]. As Klimmt et al. and other studies present in Chapter 3 suggest, soundtrack music can feature elements that trigger or amplify specific emotions, to promote a better video game enjoyment: for instance, music with “sudden onset, high pitch, wide range, strong energy in the high frequency range” [30] during a shock event may help the activation/intensification of the emotion of fear, or more generically, anxiety [11].

To fully take advantage of this model, in the next chapter, we further analyse horror media techniques presented in several works and proved to be effective, for different contexts and stages of horrific events, with the recommended diegetic and non-diegetic sounds (including soundtrack music), essential to achieve the expected congruence.
3 Related Work

Contents

3.1 Horror Media and Techniques ........................................ 15
3.2 Biofeedback and Affective Gaming ..................................... 19
3.3 Discussion ................................................................. 22
This section compiles a series of studies with relevant insights into the research of horror-based media, concepts for affective adaptive gaming experiences, and horror gaming experiences. The purpose of this section is to create an understanding of what research has already been done in the domain of audio in the context of horror, what methods were used and what important findings have been uncovered. Besides that, we also want to know what research has been done on biofeedback specifically for horror, to understand the value our research can bring to the current state of the art and what possible future iterations can be developed, based on our work.

3.1 Horror Media and Techniques

In this subsection, we explore the concepts in the lexical field of horror, and strategies to trigger emotional response (e.g. anxiety), giving a particular focus towards the use of sound.

3.1.1 Manipulation of emotions in horror media

Any horror experience makes use of shock, achievable through abruptness, which is crucial to elicit startle stimulus and an instantaneous response to it [31], which is composed by, firstly, the evaluation of the stimulus and, then, the reaction interval, giving, ideally, no opportunity to the user to assess the situation rationally. According to Bradley et al. [31], “regardless of their intrinsic affective meaning, cues signaling shock threat prompt somatic and autonomic reactions consistent with defense, and that intrinsically unpleasant cues continue to prompt defensive activation even when the context of their presentation is specifically nonthreatening.” Furthermore, audio or visual cues are able to launch the ‘horror-pathway’ [7], which makes them core elements for the creation of horrific experiences.

As Garner et al. [7] suggest, in a survival horror context, we can divide the fear behaviour in three stages:

- **Pre-Encounter Defense**, which involves the moment when we are present in a seemingly dangerous environment where predators might emerge at any time;

- **Post-Encounter Defense**, which resembles the definition of terror, involving the moment of heightened fear in response of a predator nearby, whose presence is presumed by its footsteps, for example;

- **Circa-strike Defense**, which resembles the definition of horror, involving the moment with the fight or flight response, after, for instance, the visual revelation and, therefore, confirmation of the physical threat [32].
In horror media, such as films, it is known that prior knowledge of upcoming events would generate increased sensations of fright and upset [33].

Furthermore, to evoke a terrifying reaction, diegetic audio cues that reveal the least possible information about the threat, while being enough to trigger the sense of danger, have the most potential, due to the protection of the uncertainty about the predator, such as location or any characteristics regarding its behaviour - for instance, the occlusion of an audio source’s position from a threat, may augment the intensity of the fear sensation [6,7].

Results by Garner et al. [6] also suggest that “rapid onset/offset (attack and release) of an audio signal relates to a perception of urgency, slower attack relative to faster release increases perceived intensity by way of connoting an approaching source, and all loudness, sharpness and frequency equalisation have the capacity to attenuate and amplify negative emotional activation”. However, no conclusive evidence was found to support any specific assumptions regarding the effects of sound properties on players’ fear, such as pitch, positioning or loudness. This work also relied solely on subjective data from the players, limited to instances when sounds were played, and without any sort of physiological measurements to increase objectivity.

Toprac et al. [34] also studied three audio properties – volume, timing and source – in diegetic sound to promote fear, anxiety or suspense. For these authors, volume is “the relative loudness at which a sound is heard from a loudspeaker” and timing is “the relative synchronization of the sound with its source”. Their results, from self-report surveys, interviews and suggestions, reveal that:

- High volume and sound effects synchronized with a visual element promote fear (however, with no confirmation in these cases that the sound effect by itself may have substantially elicited fear);
- Medium volume is the most adequate to trigger anxiety;
- Low volume is not as effective to elicit such sensations (“due to potential masking by other sounds”)
- Acousmatic\(^1\) and untimed sound effects evoke, mostly, “a feeling of fear and anxiety about the outcome of certain actions”.

Regarding timing, it is noteworthy that the use of white noise with a lagging technique, such as the use of a “radio static sound loop, emanating from the avatar’s pocket radio, to warn players of nearby enemies”, are a valid strategy to elicit suspense, by forewarning the player [34].

Additionally, silence is another facet in audio to take into consideration, as, to trigger shock, the most effective way may be to precede a relevant horrific event by the absence of sound [35].

According to Kumar et al., the sounds derived from what we hear everyday, many generated by people, might be effective at eliciting stress and anxiety, due to misophonia, an affective sound-processing

\(^1\)Sound that is heard without an originating cause being seen.
disorder characterized by the experience of strong negative emotions (anger and anxiety) [36]. Therefore, the game created for our study strongly relies on sounds that both generate some familiarity, but also unease, such as heavy and distorted breathing and gurgling. An empirical study on the “audio Uncanny Valley” [37] revealed general factors that can be used to trigger uncanniness, which can often lead to fear, including: “Familiar or iconic sounds can be defamiliarized and this can lead to perceptions of uncanniness”; uncertainty regarding the audio source’s location; “An aural resolution that is lower than a high quality, human-like visual resolution might lead to the uncanny”.

Nevertheless, in this study, we focus on the choice of non-diegetic audio tracks and their effect on a horror experience. Foreshadowing is attributed as a technique, commonly in narrative, of incorporating cues of lower significance to anticipate a relevant event [38]. According to Genette [39], we can distinguish two types of advance mention: genuine, which “give[s] the reader correct cues about the development of the story”, and false, which provides misleading knowledge. Scirea et al. [38] applied in their research these concepts into true foreshadowing and false foreshadowing in the game’s background music, by forebonding music before a negative event (true foreshadowing) and happy music that did not lead to a positive event (false foreshadowing). As presented in the article, mixing both types of foreshadowing can promote the audience to study which ones are true and which ones are false. According to their results, based on empirical data, it is revealed that “the use of musical cues for narrative foreshadowing induces a better perceived consistency between music and game narrative. In contrast, false foreshadowing was found to enhance the player’s enjoyment”. Also, the false foreshadowing proved to be more memorable, likely due to its element of surprise.

Ultimately, the technique of foreshadowing is directly tied to the creation of suspense in any horror media and to the selection and timing of the audio produced in a survival horror experience, as it is concluded in selected works examined in the next subsection [8–10], which highlight that sound stimuli are some of the ones that trigger the most terror and anticipation. Moreover, the music, as a non-diegetic sound, is as essential as other audio effects to convey foreshadowing and, therefore, suspense: when an audience is informed about the existence of danger, instead of being triggered through shock surprise events (which, nonetheless, can be used as a tool to quickly change a mental state of the spectators/players), it is possible to create long lasting tension. Hitchcock, known for creating suspense in his films, embraced this technique, by, for example, inserting a stressful background music during the shower scene in Psycho², to foreshadow the upcoming murder event [38].

---
3.1.2 Impact of Facets Differences in Horror Games

Graja et al. [15], used an adaptation of *P.T.*\(^3\) to analyse the players emotional reactions towards context alteration, mainly with changes to light effects, sounds and in-game events. To learn which game effects induce the strongest physiological reactions, an experiment was conducted and correlation between the physiological data, collected through the measure of the Galvanic Skin Response (GSR), and the perceived emotion provided by participants, was investigated. Results show that the order in which effects are arranged can produce extensive emotional responses. They also suggest that psychological impact can be increased not only by the visual horror itself, but also through the process that slowly builds up to it, in particular the usage of sounds, which were the most effective effects in stimulating stress/anxiety/tension, according to all of the collected data - game logs, players’ physiological data and real-time self-reported emotion annotations (emotional intensity).

Other works also reveal the relevance of audio in changing an affective state. For instance, Prager et al. [40] who studied if “the video facet is more important than the audio facet concerning user reactions”, applied Procedural Content Generation (PCG) in a maze game with exchangeable visuals and audio styles (happy and horror facets), and, by comparing different combinations (both homogeneous, i.e. happy with happy and horror with horror, and heterogeneous), concluded that “the audio facet has a bigger impact on the player’s perception of the game content than the visual facet”, according to their predictive model results, which indicated that “the happy audio setting generally performed better than the horror audio setting, regardless of the visual configuration’, which highlights that applying different soundtracks may have a stronger effect than changing the visual landscape of the game. Furthermore, the experiments by [8, 10], mentioned in the next subsection, also highlight sound stimuli as one of the most impactful strategies to trigger terror and anticipation.

Similarly to the direction of our work, Klimmt et al. [11] have studied the effects of soundtrack music on the video game experience, particularly in two different genres - adventure and horror. For the latter, users played an episode of *Alien: Isolation*\(^4\) with soundtrack music present or absent and reported their horror experience. The correlations from questionnaire responses suggested that music increases the players’ horror experience. As concluded by the authors, based on the reported enjoyment and emotions (eight-item scale on video-game enjoyment by Klimmt et al. [41], e.g. “I liked to play the game”; “It was great fun to take over the control in the game”, with a high reliability - Cronbach’s \(\alpha = 0.89\)), the players might have appreciated that experience as a horror game and react to it with a positive emotional response (the positive affect was highly correlated to enjoyment, with \(r = 0.58\)).

Thus, regardless of the effects caused by either foreshadowing technique, both non-diegetic and diegetic sounds are relevant to induce fear on players in this genre, considering the studies analysed.

\(^3\)Konami, 2014 (https://en.wikipedia.org/wiki/P.T._(video_game)).
However, it is required to further examine how the use of distinct non-diegetic tracks across a horror experience in sync with a tension curve designed by the game creators can have a different impact compared to the use of the same track during the entire walkthrough. Moreover, the full absence of background music during the entire experience might trigger an odd feeling to the players and make them feel less connected to the game, considering how the large majority of games always include any sort of background audio.

3.2 Biofeedback and Affective Gaming

A game with biofeedback mechanics is enhanced by physiological signals from the player. Still, biofeedback is a broad concept and can be applied in a gaming experience in different ways, including how it affects the game’s mechanics, how the player’s data is interpreted by the game and if the player acknowledges its existence.

According to Nacke et al. [42], we can subdivide biofeedback in two categories: direct and indirect physiological input, which refers to the contribution of the physiological data to the gameplay mechanics, and is, therefore, also connected to whether the player is oblivious or aware of its existence. For example, in a direct biofeedback game, a game mechanic is manipulated directly by a physiological signal – Facial Electromyography (EMG), respiration, temperature and gaze –, whereas an indirect biofeedback game makes use of the player’s physiological data by indirectly changing the game, such as changing subtly the environment of the game through Electrocardiography (ECG) and GSR. Similarly, Kuikkaniemi et al [43] introduced the definition of explicit biofeedback mechanisms, where the player is informed that their current physiological state directly influences the game, and implicit biofeedback mechanisms, where the player doesn’t know that the game is influenced by their physiological state.

As proposed by Gilleade et al. [44], furnishing a videogame with biological information during gameplay does not make a game affective. In other words, a game is only affective if it propagates affective feedback. As mentioned in their study, “just replacing conventional input (i.e. conscious decisions executed through interactions with a game pad), is what we regard as a straight-forward biofeedback game”: for example, a game that uses the player’s chest volume to solely modulate a game variable is not considered an affective game, since “the players’ emotions are not a factor in their interaction with the game (despite the emotional effect the biofeedback mechanic may have on the player)” [45].

3.2.1 Horror Gaming Experiences with Biofeedback

There is a large body of works that study the aforementioned themes, assessing fear and the biofeedback modulation of affective player experiences in horror. As such, in this subsection we present a
brief overview of relevant experiments that will serve as a blueprint for future work in next iterations of our project, by including the use of physiological data for emotional state identification and consequent manipulation of game parameters in a horror context.

Although the use of biofeedback in video games hasn’t yet reached mainstream popularity, we have, for instance, *Nevermind⁵*, one of the first publicly released games that can be enhanced by the players’ physiological data, focused on fear management based on players’ HR, more specifically by recording the Heart Rate Variability (HRV) [46]. In this game, similarly to the studies we are presenting below, the player’s HR is continuously read into the game which in turn adapts to the player’s momentary levels of negative affective arousal. Greater negative arousal causes the game and its horror-themed settings to become more disturbing. Thus, this game challenges players to improve their emotion regulation skills by encouraging them to down-regulate their negative affective states when facing stressful situations [46]. An example of this game's biofeedback system is a segment where, as the player’s anxiety rises, the higher water/milk rises. If it rises too much, the player’s character can drown and loses the game. If the player relaxes, the water/milk will subside. Although the application of biofeedback in each specific segment is not explained in advance, the player is notified early on about the impact of their physiological signals on the game's behaviour, crucial for the player to succeed in several levels.

In the context of fearful experiences, especially academic studies involving survival horror prototypes described below, we can verify that other several affective biofeedback games use physiological data directly in the game mechanics, with no intermediate interpretation to explicitly associate the signals with emotional states.

In order to study if a biofeedback loop in a Virtual Reality (VR) experience can enhance user engagement, Houzangbe et al. [47], produced a small survival horror experience, where the player is confined in a limited indoor space with a music box playing music that attracts a predatory creature, who can appear from multiple locations. To win the game, the player has to limit the amount of noise coming from the music box and close doors. In the physiologically enhanced version, the game collects the player’s HR throughout the playthrough, which directly impacts the character’s heartbeat sounds, presented as another source of noise to attract the predator – after calibration and measurement of the average HR, for each increase of 15 BPM, the game gets harder, with a faster and louder character’s heartbeat. When the player’s heart rate reaches 40 BPM over the calibrated value, the screen’s field of view starts reducing. The results from a questionnaire based on Presence (Witmer and Singer [48]) and User Engagement Scale (Wiebe et al. [49]), were inconclusive, as no significant differences in involvement were observed between the two groups (control group with no biofeedback and another one with active biofeedback). Additionally, the participants were presented questions about the use of biofeedback and what it brought to the experience. In this experiment it was partially validated that “The users [were]⁵Flying Mollusk, 2015 (https://nevermindgame.com).
capable of influencing their own heart rate in game\textsuperscript{6}. Finally, considering how the “addition of a new game mechanic can increase the difficulty in mastering a game”, it was also hypothesized and verified that “The addition of the biofeedback loop [modulated] the game difficulty making it more challenging to master”, since the majority of the players preferred the non-biofeedback version of the game, confirming “a steeper learning curve and higher difficulty to master the game with the biofeedback loop”.

Outside of the academic circuit, one of the industry giants, Valve Corporation, has also been researching on how to implement biofeedback in gameplay and measure physiology to enhance the experience, with modifications of games such as Left 4 Dead\textsuperscript{7} [50]. With the integration of physiological data, the new game’s version could make use of the actual measured arousal of the player. The director’s algorithm of the game would “represent Survivor intensity as [a] single value, [and] increase it in response to in-game trauma”, decay its intensity over time and create peaks and valleys. According to their results, based on Skin Conductance (SC)\textsuperscript{8} data and questionnaire responses, the actual measured arousal “produced greater enjoyment than [the designed] estimated arousal”, which confirmed that physiological signals are viable inputs and can enhance a gaming experience.

Another example of a study involving an affective biofeedback experience is by Dekker et al. [8], who applied a modified version of Half-Life 2\textsuperscript{9}, by Valve, to incorporate direct biofeedback mechanics, which impacted the player’s abilities and movement. For this experiment, average values for HR and SC would be assessed in the beginning and, then, the environment and player’s character would change their characteristics according to how higher or lower these two variables (HR and GSR) were in the present moment and specific ratios with the baseline values, designed by the authors. Although the results, based on interviews and biometric data, indicated that ECG can effectively help understand the player’s emotional state and “determine to some extent a user’s reaction to an event”, they also showcased how the incorporation of an additional ‘learning’ variable (direct biofeedback for the gameplay) may lead to a lower engagement for players not familiarised with the horror genre, similarly to what Houzangbe et al. [47] concluded as well, considering that the majority of players may prefer to play a game with a gradual learning curve.

Moreover, the contributions of Vachiratamporn et al. [9] about player effective states prior and after a scary event are also to be counted, as their results suggested that players “were more likely to experience fear from a scary event when they were in a suspense state compared to when they were in a neutral state”. However, their next work [10] with the implementation of the affective biofeedback version of a “Slender” inspired game failed to verify possible significant differences in the evaluation of players when implicitly changing the timing of events. One of the key reasons for such conclusion might be

\textsuperscript{6}In this experiment, the users were informed in advance about the biofeedback feature of the game.

\textsuperscript{7}Valve, 2008 (https://en.wikipedia.org/wiki/Left_4_Dead).

\textsuperscript{8}The SC data analysis included: categorization of game events, record survey responses, quantification of waveforms’ spike frequency and other characteristics, and data mining, such as correlation and regression.

References | Physiological metrics | Biofeedback Enhancements
---|---|---
Houzangbe et al. [47] | HR | Difficulty, Audio-Visual
Dekker et al. [8] | HR, GSR | Difficulty, Audio-Visual, Player, AI
Vachiratamporn et al. [10] | HR, EEG | Difficulty, AI
Nogueira et al. [45] | HR, GSR, EMG | Difficulty, Visual, Player, AI, Level
Reynolds et al. [46] | HR | Difficulty, Visual

Table 3.1: Overview of biofeedback horror-related studies & games

possible poor game design choices in the affective version, leading to unbalanced game difficulty, due to the forced changes in the predator’s behavior.

Nogueira et al. [45] developed a procedural horror game (Vanish) capable of run-time level, asset, and event generation, to perform a comparative analysis on how different types of biofeedback gameplay mechanics affect players’ gaming experience, studied through Game Experience Questionnaire (GEQ) reports\(^\text{10}\) and physiological data, “in order to assess whether the design of the aforementioned biofeedback mechanics have significant effects”. Two biofeedback versions (explicit and implicit) were tested and results indicated a more concentrated distribution of users on the central region of the arousal-valence space in the explicit version (‘aimed to deepen the empathetic bond between the player and the game character by binding the emotional states of the player and the game character’), which promoted a positive biofeedback loop, unlike the implicit version, where “the player’s arousal level was inversely correlated with the probability of generating a creature encounter”.

As our Table 3.1 reveals, it is noticeable how all of the studies, of which we have knowledge, that approach biofeedback in horror gaming experiences, manipulate directly, not only the game mechanics (by modification or addition), but also, often as a consequence, the difficulty of the game. Secondly, we highlight again the issue of adapting too many variables simultaneously, which causes the inability to discriminate the impact of each game parameter when designing an adaptive affective experience. Finally, regarding the metrics for the physiological input, we can identify HR as the most relevant in the context of horror, as it is present in all of the selected studies and showcased prominent results.

3.3 Discussion

As presented above, there is a large body of works that study the aforementioned themes, many assessing fear and horror experiences, the biofeedback modulation of affective player experiences and the level of fear through soundscape. Nevertheless, we can, firstly, conclude that all the research found regarding the use of a biofeedback mechanism applied it in too many variables of the correspondent game, leading to game design flaws in the enhanced variations and undesired changes in core game aspects,

\(^{10}\)Experience ratings on seven dimensions: immersion, tension, competence, challenge, flow, and positive and negative affect.
such as difficulty [10], or leading to a broad general analysis to answer only if it is possible to “modulate the “scary” experience of playing a horror game using physiological sensors and real-time processing” or if “biofeedback-enabled adaptive mechanics have a significant impact on the players’ gameplay experience” [45]. Most common tested changes involve variables directly connected to the gameplay mechanisms or enemy AI [45,47]. Subtle and non-intrusive manipulations in a game’s environment have not yet been explored to its full potential, especially in the audio field. In fact, even discarding the impact of biofeedback for audio manipulation, we still need to deepen our understanding of how the audio can change the emotions and tension felt in a horror experience. As such, in this study, we have the opportunity to address questions, including: to what extent can relaxing music change the perception of a game area, compared to a neutral audio track? And concerning intense and stressful music - Does it make a section scarier if it already makes use of many visual and gameplay elements that evoke terror? Moreover, although studies (such as Klimmt et al. [11]) already reveal pertinent conclusions, collecting players’ physiological data, such as the HR, may allow more detailed insights and gauge the interest of investigating the application of biofeedback for sound choice and timing manipulation in a horror game.

Taking into account how [47] and [8] detected and transitioned between the different affective states, and how [45] considered a positive correlation for HR/arousal and a negative correlation for HR/valence for the horror genre, we will focus on collecting the HR of the players and using the following assumption: for our study, the higher the HR, the greater the arousal and the less the valence of the player. In essence, a high HR (high arousal and negative valence) translates to, generally, anxiety. Thus, the concept of fear will be implicit in our affective state of anxiety, depending on the actual danger perceived by each player, which will vary from situation to situation. For a low HR, compared to the corresponding average value of the player, we will consider they are bored or calm (i.e. not feeling tension); for a high HR, we will consider they are scared (i.e. feeling tension), unless they are located in an area of the game designed for relaxation and showcased an increase of HR already after the transition from an area prepared for anxiety (i.e. positive valence, which may still lead to a high arousal, but for excitement, rather than fear, due to the exposure of a different context to explore, with new elements to discover).

Based on the collected theory regarding horror techniques with audio and the importance of understanding how non-diegetic audio, including music, can impact an horror experience, we focus on creating a short game demonstration, where it is possible to test two similar scenarios, that only share one difference: the selected non-diegetic audio track. More precisely, in the next chapters we will describe the study of the influence of two approaches: (1) using a dynamic-based audio background version A (default) of the game, which changes the background audio layers depending on the goal anxiety level of the current game section; (2) using the same audio background (neutral state, already used during the first game section) during the entire horror experience.
4

System Overview

Contents

4.1 Presa ................................................................. 27
4.2 Bio Framework .................................................... 33
4.3 Summary ........................................................... 35
With the intention of preparing a system which unites the necessary conditions to answer our current research questions, we created a game, Presa, with an Audio Manager capable of changing its parameters according to the game conditions and, optionally, players’ physiological data. This chapter describes the two core components and their key development steps.

4.1 Presa

This section covers why we chose the game theme and mechanics. It also presents the base game design and the two game versions we have created to attempt to trigger different mental states in the player.

4.1.1 Game Choice

Our short horror gaming experience used in the user tests is a psychological first-person horror made from scratch, taking place in a semi-realistic dark environment populated by creatures with a deformed humanoid presentation. The decision of creating a new game for the study was presented from the beginning, to ensure the access to all in-game information, the communication via socket with the biometric data framework, and the full control over creative and artistic choices. The game was made in the Unity Game Engine, an engine that the author already had experience with.

![Gameplay Snippet of Presa](image)
4.1.2 Game Concept

Presas (see Figure 4.1) is a game inspired by psychological horror games like Outlast\(^1\), Alien Isolation\(^2\), P.T.\(^3\) and Slender: The Eight Pages\(^4\), taking place in a dark cave system, with surreal elements and diverse pathways, which will, nonetheless, always lead the players to the same destinations. The player, confined in this threatening environment, has the goal to escape it, without getting caught by malicious enemies. In order to escape and succeed, the player needs to collect three items, called “orbs”. The player actions are explained in further detail in 4.1.4.

4.1.3 Game Loop

The game consists in several interconnected sections that need to be traversed, in order for the player to progress and conclude their journey. Although multiple paths are presented during many sections (to give the players the illusion of choice and the feeling of getting lost), the state of the game is always updated in an equal manner, triggering the same predetermined events and key regions of the game, to guarantee a linear experience for everyone. To complete the game, the player is only fully required to walk and collect the three orbs, although it is promoted the use of a sonar tracker and a slower movement (by crouching) in the last section, where there is the additional challenge of avoiding a deadly enemy, who wanders around a small region. In this part of the game, the player character’s health can reach zero, which leads to the reset of the last section of the game, maintaining the progress made regarding the items collection.

4.1.4 Character Actions

The actions the player character can do are the following:

- **Walk and look around**: traverse the various zones of the cave system, viewed in a first-person perspective;

- **Crouch and Run**: change the movement speed, by enabling crouch (slow down) and enabling sprint (speed up)

- **Check sonar tracker accessory**: when enabled, this accessory shows items (orbs and batteries) and enemies nearby;

- **Interact with objects**: collect an orb (three in total) or batteries for the sonar tracker and read tombstones;

---


Before starting the horror level, the players have the opportunity to check the controls layout and try the character actions (Figures 4.2 and 4.3).

![Figure 4.2: Scene displaying the controls layout for the character actions.](image1)

![Figure 4.3: Tutorial scene, where the players explore the essential character actions.](image2)

### 4.1.5 Game Sections

For the game, we aimed to follow the usual horror game design with a slow and gradual tension increase over time, which has been proven to be adequate to elicit fear [15]. More precisely, we designed a set of sections that expect to provoke small phases of tension climax, through the use of a minor jump scare and appearance of visually threatening regions (e.g. with static creatures).

In Figure 4.4, we present a set of screenshots covering relevant zones of the game. In total, the game presents eight sections, separated by key events, such as collecting an orb, necessary to reach the end of the experience. These sections, which end up describing the progress in the game, are monitored by an Anxiety Manager, responsible for keeping track of the emotional state that we want to trigger in our players (i.e. if we want to increase or decrease anxiety). Additionally, these sections’ timestamps are then considered in the evaluation for self-reported tension and HR, for further comparison and analysis. It is relevant to note the length of the tracks vary and their completion time depend on the player’s performance.

- **Section 1**: From the start of the game, where the player is confronted with the environment, up until the first orb collection - Figure 4.4 (a, b);

- **Section 2**: Traversing more segments in the cave with a similar layout, until the player reaches the exit gate for the first time (locked due to the lack of orbs) - Figure 4.4 (c);

- **Section 3**: While going back, the player observes a different path, with different visual elements - Figure 4.4 (d);
• **Section 4**: After being triggered by a short and sudden appearance of a creature, the player navigates through more regions of the cave, with a similar layout as the ones presented in Sections 1 and 2;

• **Section 5**: Region with presence light of warm color and a battery - Figure 4.4 (e);

• **Section 6**: Traversing more segments in the cave with a different layout and darker rocks, up until the second orb collection, which triggers a small animation to a laid down creature holding it - Figure 4.4 (f);

• **Section 7**: Region that covers the second exit gate (still locked due to the lack of three orbs), with various static creatures and headstones - Figure 4.4 (g, h);

• **Section 8**: Stealth area of the game, where the player needs to avoid a moving creature and collect the third orb, in front of the final exit gate that opens and leads the player to the end of the game - Figure 4.4 (i).

**Figure 4.4:** Screenshots representative of different parts of the game - The first line (a, b, c) contains key moments during Sections 1 and 2, which end at the first appearance of the exit gate; the second line (d, e, f) represent Section 3, 5 and 6; The third line (g, h, i) depicts the last two sections, filled with various lost headstones and the appearance of a dangerous creature. Section 4 begins with a quick appearance of a creature and continues with the regular maze paths presented during the first two sections.
4.1.6 Non-Diegetic Audio Tracks

Regarding the main focus of this study, the use of non-diegetic audio tracks, we created five distinct tracks in the game, based on the conclusions from the studies introduced in Chapters 2 and 3, and the results from our preliminary study assessing four original tracks.

- **Neutral**: Base track for both versions of the game, composed by sounds of monotonic bass;
- **Anxiety**: Track with the elements of Neutral and additional subtle strings;
- **High Anxiety**: Track with the elements of Anxiety and additional stronger strings;
- **Suspense**: Absence of instruments, with occasional non-diegetic sound effects around the player’s character location (e.g. dust falling, footsteps and quiet growls);
- **Relax**: Track with a calm instrumental, composed by sounds of synths.

The tracks above, with the exception of Suspense, were evaluated isolated during the preliminary tests (for more details, see Chapter 5) and, as their code names suggest, were graded in the following order for tension: relax, neutral, anxiety and high anxiety. The suspense track is used in the default version of the game to create contrast between critical moments of the game - although this track was not evaluated isolated, it triggered the desired reaction in players during the pilot tests of the game, and its relevance and effect to anticipate a horrific event is theoretically expected based on our related work ([35] at Chapter 3.1.1). Thus, the Game Section 3 makes use of Suspense to evoke the anticipation of an intense event, with the goal of eliciting the phenomenon of true foreshadowing.

The illustrative diagram at Figure 4.5 describes the anxiety desired evolution for our horror experience, with the background non-diegetic audio tracks that are expected to be used for each game section, in Version A (default). The other version of the game, Version B, disables the change of background audio tracks, which leads to the permanent use of the same “Neutral” track from the first game section during the entire game. For more details about the game versions and evaluation, see Chapter 5.3.

4.1.7 Creatures

The expectation for the appearance of creatures and their behavior are key to elicit anxiety. Across all of the game’s sections with the exception of the last one, the showcased creatures are not actual threats to the player character’s health. Only near the end the player will be challenged to traverse a region with a creature that moves around in a cyclic path, able to detect and attack the player if they get too nearby. If the player is hit by the enemy two times, the game restarts in the same section, with saved progress (two orbs already collected).
4.1.8 Tracker

Besides moving and interacting with objects, the player is given the opportunity to use a sonar tracker, able to detect the presence of nearby items (orbs or batteries) and enemies. This accessory is not required to complete the game, however its use is highly recommended during the last section of the game. Inspired by the study by Toprac et al. [34] mentioned in Chapter 3.1.1, we decided to incorporate this device into the game to bring, not only additional guidance for progression, but also uncertainty about the danger, especially with the use of radio static sound loops, which aids to elicit suspense. Still, regardless of the audio emitted by the tracker, this source of sound does not affect in any manner the gameplay (i.e. the enemies do not hear the tracker).

![Tracker Diagram]

**Figure 4.5:** Illustrative diagram of the anxiety evolution design for the game, with the non-diegetic audio track selected for each section, for the default version of the game. The diamond shapes represent the collection of each orb. The square shapes represent the moments where the player finds one or multiple creatures (A - quick jump-scare event, where a harmless creature suddenly appears and disappears; B - harmless creature sitting next to an orb; C - static harmless creatures; D - dangerous creature that roams in an area that has to be traversed by the player).

**Figure 4.6:** Tracker detecting the presence of an enemy, battery and orb; the tracker also reveals the direction to reach the gate, indicated by the green arc.
As seen in Figure 4.6, the tracker’s screen reveals the items and enemies in polar coordinates and, in the top left corner, the percentage of battery left until it turns off. When the battery is lower than 20%, the screen’s image gets distorted. In this circumstance, we apply a slower energy spending speed, to avoid situations where the player reaches 0% and can’t rely on the tracker anymore, until they find a battery again.

4.2 Bio Framework

In order to save and keep track of the player’s HR in sync with the game and allow the use of an adaptive version of the audio manager for future work, we need a framework capable of trading information between the biometric interface and the game.

This section addresses the physiological measures used for this study and the architecture of the overall system.

4.2.1 Heart Rate

As presented in Chapter 3, among several physiological measures, HR reveals itself as one of the most prominent ways to define user states and apply them in biofeedback solutions for horror games. Additionally, a key criteria for our experiment was to choose physiological measures that require the least intrusive accessories and procedures, to avoid triggering distractions or discomfort to the players. Therefore, methods such as GSR, Electroencephalography (EEG) and ECG which need electrodes were discarded.

For our solution, we decided to use a Photoplethysmogram (PPG) sensor\(^5\) to record the Blood Volume Pulse (BVP), which leads us to calculate the instantaneous HR. Although this signal mostly gives us only an overview on the players’ current arousal, we assume a state of negative valence during most of the game, and, thereby, we considered the evolution of HR during the experience appropriate to define the players’ anxiety and relevant enough to compare it with the players’ self-reported tension. For more details regarding the experiment and its apparatus, see Chapter 5.

Considering that, unlike ECG signals, PPG signals are very sensitive to motion artefact which may lead to poor HRV estimation if false peaks are detected. Although there has been progress in estimating HRV in PPG signals using Bayesian learning approach [51] or through ECG reconstruction [52], we decided to just focus on using the instantaneous HR based on the recorded BVP and measuring the average HR per Game Section, to later compare these values with the self-reported tension per Game Section, obtained in a questionnaire provided after the experience (see Chapter 5.3). On a side note, it is worth mentioning that Pulse Rate Variability (PRV) has also been proven to be a solid alternative to HRV

\(^5\) PulseSensor, connected to a Bitalino, from the (r)evolution Plugged Kit BT.
for tests with healthy subjects at rest, although “stimuli conditions such as exercise [seem] to decrease the correlation [with HRV]” [53], which may remove its relevance in the context of our experiment, with horror events. Still, this strategy could eventually be considered in future iterations of this work.

4.2.2 Architecture

Figure 4.7 depicts the architecture of the framework that allows the collection of physiological signals (in this case, the BVP) and exchange of data with the game. The illustrated process occurs for time intervals of one second. During the experiment, a Python script runs the BioSPPy toolkit and processes raw data from OpenSignals (r)evolution, through Lab Streaming Layer (a). BioSPPy outputs the instantaneous HR (b). Next, the script receives and sends information to the game via UDP sockets, regarding gameplay data, key timestamps for the game sections, and the player’s physiological data (d). The latter and step (c) are not relevant for the purpose of this study, but may be crucial for future work on biofeedback functionalities.

Figure 4.7: System architecture, with a module responsible for physiological data and another for the game. From left to right, it is composed by user raw data (analog BVP), Bitalino with a PulseSensor attached (a), BioSPPy (b) and data treatment (c). First, the user produces analog signals which are measured with the Bitalino. These signals are then digitized by OpenSignals (r)evolution and fed to the BioSPPy which filters them, performing R-peak detection for the computation of the instantaneous HR. After that, via UDP sockets (d), the script is able to treat this information and send to the game and receive gameplay data logs, to synchronize relevant timestamps with the obtained physiological signals.

In order to allow the dynamic change of background audio across different game conditions, we have a selection of managers, each with different responsibilities, already considering a future iteration of the project that could expect the addition of biofeedback:

• **State Manager**: Communicates with both Bio Framework, through UDP sockets, and the Anxiety Manager. In future work, it is expected to have a set of criteria in order to define an emotional state based on the player’s physiological data;
• **Anxiety Manager**: Receives the user state and other key information, mainly at what Game Section they are currently playing, to define what type of mood we want to promote via Audio Manager (i.e. *Relax, Neutral, Anxiety, High Anxiety* and *Suspense*)

• **Audio Manager**: Receives the desired mood from the Anxiety Manager and controls its background audio generator (Music Generator) correspondingly with the correct parameters values. This generator is tightly connected to an FMOD Project, which aggregates multiple audio tracks and continuous parameters that change based on the level of intensity transmitted firstly by the Anxiety Manager.

With the above described managers, the game system can be easily adapted to incorporate biofeedback for background audio choice and timing, based on the game conditions and goal mental states sent by the python script, from the side of the Bio Framework.

### 4.3 Summary

In this section, we presented how the game and framework related to the physiological signals were designed and composed. In short, we started by creating a game that could fit the study requirements and objectives, and apply the current knowledge of horror techniques in media. Then, we explored the devices and libraries needed to obtain the players’ instantaneous HR, possible with the creation of a script that also exchanges data with the game.

Essentially, for the current challenge presented in our study, the main aspects to guarantee in the architecture solution are: (1) the capture of the players’ BVP and computation of HR in real time; (2) the automatic frequent data exchange between the Python script (from the side of “Bio Framework”) and the game’s State Manager to synchronize timestamps and, in future work, allow the use of biofeedback mechanisms; (3) the communication of the game managers with the FMOD Project, to modify the background audio according to the desired state. With all the components implemented, we are able to proceed to the evaluation phase, where we can properly answer our research questions and accept/refute our hypotheses.
5

Evaluation

Contents

5.1 Audio Pilot Tests ................................................................. 39
5.2 Usability and Game Pilot Tests ............................................. 39
5.3 Final Experiment ............................................................... 40
5.4 Discussion ................................................................. 53
In this chapter, the two evaluation phases of this work are described: (1) the audio pilot tests, where the main objective was to gather and compare the self reported tension induced by audio tracks from users; (2) the final evaluation where the main objective was to study the impact of the non-diegetic audio tracks across the game sections in players’ anxiety and self-reported tension.

5.1 Audio Pilot Tests

As a set of original audio tracks was created specifically for Presa, evaluation regarding their perceived level of anxiety was required to guarantee their reliability for future evaluations.

The objective of these tests was to identify the level of self-reported anxiety induced by the audio tracks and verify if all users had a common comparative grading between each track.

5.1.1 Procedure

In order to classify, sort and verify the degrees of anxiety induced by the background non-diegetic/music tracks, we conducted a questionnaire, included in Appendix A, where we presented four tracks and the users were asked to classify each on a 7-Point Likert scale (“1 - very calm” to “7 - very anxious”).

Tracks’ titles make use of symbols which do not provide an alphabetic order bias (such as & and #). Each user was presented a different order of the tracks, with the possibility to change any score until submission.

5.1.2 Results and Changes

In the first iteration of tests, we picked four users to assess if the comparative order of the tracks was consensual. After registering an outlier between two tracks, two more users evaluated the tracks and we verified the presence of one more outlier. Consequently, one of the concerning audio tracks was updated and the procedure was repeated with six different participants in a second iteration of this preliminary experiment. In the end, everyone sorted the tracks in the expected order, with only one draw between two audio tracks in two responses, as seen in Figure 5.1. The results confirmed the overall level of self-reported anxiety in a defined order for the tracks, which allowed the use of the original soundtrack for the study’s main experiments with the testbed game.

5.2 Usability and Game Pilot Tests

Considering the testbed game, Presa, was made from scratch for this study and the experiment pipeline required a list of steps, due to the collection of HR, we conducted pilot tests with four players, to assess
the usability of the game and identify possible issues with the procedure and game features.

In the pilot tests, players were subjected to the final experiment’s procedure and questionnaire further explained in Chapter 5.3, with the additional tasks of filling in an additional set of questions of an adapted core-GEQ, assessing “Flow” and “Sensory and Imaginative Immersion” (see Appendix C), and providing feedback to the assistant in a free interview, regarding any problems felt while playing the game and suggestions for improvement. The assistant was also present in the experiments and had the opportunity to identify any issues and details requiring adjustments.

Based on the feedback and observations during these first tests, minor adjustments were made to the game, while not changing the core experience. The difficulty of the last section was slightly reduced, by providing more powerful batteries and giving more hints for a recommended path to avoid the creature. Additionally, to minimize the total time without battery in the tracker, while maintaining the pressure on the player to search for batteries, we apply a slower energy spending speed when the tracker battery is below 20%, as already described in the previous chapter.

Regardless of the minor adjustments, the players managed to enjoy and finish the game and gave positive feedback regarding the tension felt while playing the game, according to their form responses and comments given in a free interview.

5.3 Final Experiment

With the intent of investigating the impact of non-diegetic audio in the horror experience and analysing if HR is a good physiological metric to reflect the anxiety state of the player, we formulated the following research questions:

- **RQ1**: Is the use of dynamic non-diegetic audio able to trigger more intense emotions and affect
the perception of a horror gaming experience and the players' HR?

• **RQ2**: Is the HR a good indicator of the player's anxiety state?

To answer these questions, we verified the following hypotheses:

• **H1**: Users have a higher positive affect playing the dynamic-based version compared to the uniform-based version

• **H2**: Users have a higher negative affect playing the dynamic-based version compared to the uniform-based version

• **H3**: The use of dynamic non-diegetic audio tracks in a horror experience promotes a higher change of self-reported tension between sections in players.

• **H4**: The use of dynamic non-diegetic audio tracks in a horror experience promotes a higher change of HR between sections in players.

• **H5**: The evolution of normalized average HR between consecutive sections correlates to the evolution of normalized self-reported tension between the correspondent consecutive sections

To be able to accept or refute the hypotheses, for this experiment we elaborated two versions of the game “Presa”, already described in Chapter 4.1, which only differ in the change of non-diegetic background audio tracks:

• **Version A** - “Dynamic non-diegetic audio background”, which refers to the use of different audio tracks across the game sections;

• **Version B** - “Constant non-diegetic audio background”, which refers to the permanent use of the neutral audio track (described in Chapter 4.1.6) during the entirety of the game.

### 5.3.1 Participants

We recruited people interested in participating with at least 18 years old through convenience sampling, such as direct contact or word of mouth.

Our sample was composed of 30 users (24 male, 5 female, 1 other), aged between 18 and 38 \((M = 24.33; SD = 3.575)\). All tests occurred between 10am and 7pm. Only 2 participants reported that they don’t play video games. The remaining participants play when the opportunity presents itself (43.3%) or make time in their schedule to play video games (50.0%). The majority of the users (76.7%) considered easy to navigate in first-person games (by responding 4 or 5, in a Likert scale ranging from “1 - Totally disagree” to “5 - Totally agree” with the statement “I find it easy to navigate in first-person perspective...
games”). Regarding the participants who weren’t as familiar with the genre, we offered additional time before the experiment to try out the gamepad and practice basic movement controls (walk and look around). Everyone managed to complete the game, and no correlation was found between experience with first-person perspective controls and total play time (Spearman’s Roh: $Rs = -0.286, p = 0.125$). The mean play time for all users was around 16 minutes, and no players took less than 9 minutes or more than 26 minutes (play time in MM:SS: $M = 16:17; SD = 03:48; SE = 00:42$).

5.3.2 User Evaluation

In this subsection we present the questionnaires used in the end of the experiment to measure the self-perceived emotions and tension felt while playing the game (for the full questionnaires, see Appendix B).

All users were provided a form to answer the questionnaires, completely written in English. In case of doubts as to the meaning of the questions or other content, users were allowed to ask the assistant for clarification or translation, although in most users such help was not required, with the exception of a few who were not fully comfortable with English.

Tension per Game Section

With the goal of understanding and comparing the impact of the two experiment scenarios along the eight game sections, we rely on subjective and objective data. The latter is assessed through the measurement of HR. The former is guaranteed by providing a questionnaire where we ask the user to rate the amount of tension felt during each game section in a Likert scale ranging from “1 - Not at all tense” to “9 - Extremely tense”.

PANAS

Watson et al. [54] developed the Positive and Negative Affect Schedule (PANAS) scale with the objective to assess one’s self-perceived well-being and affectivity. For this experiment, we used the original scale which consists of twenty items that aim to evaluate positive and negative affect on a Likert scale ranging from “1 - Very slightly or not at all” to “5 - Extremely”. In this version of the questionnaire, ten items belong to the positive affect component, e.g., enthusiasm, inspiration, interest, and the other ten to the negative affect component, e.g., irritation, fear, nervousness.
5.3.3 Apparatus

In Chapter 4.2, the choice of physiological metrics for our study is discussed. As our final decision was to collect the instantaneous HR, we used a BiTalino\textsuperscript{1}, from the (r)evolution Plugged Kit BT\textsuperscript{2}, to record the Blood Volume Pulse (BVP) of the player with a PPG sensor - the PulseSensor\textsuperscript{3}. The BiTalino, with a PulseSensor connected to it, was placed well distanced from the players on the table in each experimental condition and during the baseline measurement. Before starting the experiment with the relaxing segment, tutorial and game, the assistant explained each player how to place the PulseSensor on the tip of a finger and started capturing the signal, to verify if the raw data was being correctly measured and received in the system framework. Users were recommended to use the index finger of their right hand, based on the controls layout defined for the game and which fingers would be least needed.

![Figure 5.2: Main board of the (r)evolution Plugged Kit from Bitalino with one sensor cable connected to it, for Pulse Sensor.](image)

![Figure 5.3: Pulse Sensor.](image)

For the experiment, one computer was required in any test, to capture the physiological data and run the necessary programs, including the game. Players were given the opportunity to choose between interacting with the game through a mouse and keyboard or a gamepad, according to their preference and familiarity with the options. Additionally, all users were requested to put on and use a pair of headphones (Audio-Technica ATH-M40x). The data for the BVP was recorded at 100 Hz with OpenSignals (r)evolution OpenSignals 2.1.1 software (December 2020)\textsuperscript{4}.

5.3.4 Procedure

The tests were conducted in two laboratories in the same temperature interval and ambient lighting, with reduced light sources to create a dark environment, necessary to heighten the horror experience. The participants were left alone in the room with the assistant. The assistant started by explaining the purpose of the study, what they would be doing and that they should avoid moving the finger with the

\textsuperscript{1}\url{https://bitalino.com/}
\textsuperscript{2}\url{https://bitalino.com/products/plugged-kit-bt}
\textsuperscript{3}\url{https://plux.info/sensors/42-pulsesensor.html}
\textsuperscript{4}\url{bitalino.com/downloads/software}
PulseSensor attached, to prevent noise/incorrect readings or detachment of the wire from the Bitalino. Then, users were asked to fill a consent form for the experiment and a small form dedicated to demographic data, while the assistant added and connected the Bitalino to OpenSignals (r)evolution, and, afterwards, initiated the scripts to prepare the collection of biometric data, and started recording the baseline values for five minutes while the player had to look at a selected set of neutral pictures from the Open Affective Standardized Image Set (OASIS) database\(^5\), with high values of valence and low-mid values of arousal, which relates to relaxing stimuli (in a 7-point Likert scale - valence: \( M > 5.9; SD \leq 1.0 \); arousal: \( M < 4.0, SD \leq 2.0 \)). Simultaneously, the user would be asked to put on a pair of noise cancelling headphones and relax.

After these first steps, the users played a short tutorial-demo of the game to try the basic controls and in-game interactions. This segment generally took less than a minute to complete, considering that it skips to the game after the execution of the key tasks. If the player presented difficulty finishing the basic actions or misses important information regarding the controls or what to do, the assistant would provide help during this phase. Additionally, the users who never engaged with games with a first-person perspective were invited to try the basic controls before starting the experiment.

The users took approximately 8-20 minutes to complete the game. If the player’s character died in the final section of the game (which is the only part of the game where this event is possible), they would restart near the same zone. After playing, the users were asked to:

- Review, with the help of the assistant, the sections of the game (in a recorded video of the playthrough) and fill a form to evaluate the tension felt during each of them in a 9 point-Likert scale;
- Fill a PANAS form;
- Optionally, give additional comments about the game and experience.

**5.3.5 Results**

Following the user tests, we performed statistical data analysis in order to answer our research questions. Subsections 1 and 2 address the relevant information from the questionnaires to investigate, correspondingly, the players’ positive and negative affect, and the self-perceived tension per game section. Subsection 3 presents results regarding the collected mean HR per section of the players and a comparison with the self-perceived tension.

For the following reports, we start by performing Shapiro-Wilk normality tests to decide which follow-up to use.

---

\(^5\)Open-access online stimulus set containing 900 color images depicting a broad spectrum of themes, including humans, animals, objects, and scenes, along with normative ratings on two affective dimensions - valence and arousal.
PANAS

For the first and second hypothesis H1 and H2 - Users have a higher positive/negative affect playing the dynamic-based version compared to the uniform version - we analyse the PANAS scores.

H1: Users have a higher positive affect playing the dynamic-based version compared to the uniform version.

This hypothesis states that playing with dynamic non-diegetic tracks across the sections of the game induces a significant difference in the positive affect of the user. Shapiro-Wilk test was not significant for data in both versions of the game (Version A: $D(15) = 0.943, p = 0.420$; Version B: $D(15) = 0.958, p = 0.420$). By comparing means of positive affect between the two versions (Version A: $M = 33.20, SD = 5.583$; Version B: $M = 28.93, SD = 5.637$) using a paired-samples t-test, we verified statistically significant difference ($t(28) = 2.083, p = 0.047$). Thus, we accept Hypothesis 1, as the results suggest that the players enjoyed the game more while playing Version A.

H2: Users have a higher negative affect playing the dynamic-based version compared to the uniform version.

This hypothesis states that playing with dynamic non-diegetic tracks across the sections of the game induces a significant difference in the negative affect of the user, which may suggest that the player felt a stronger sense of danger or discomfort, for instance, compared to the uniform version. Normality testing through Shapiro-Wilk showed no evidence against normality (Version A: $D(15) = 0.985, p = 0.993$; Version B: $D(15) = 0.940, p = 0.383$). By comparing means of negative affect between the two versions (Version A: $M = 26.33, SD = 5.538$; Version B: $M = 23.07, SD = 5.063$) using a paired-samples t-test, we verified no statistical significance: $t(28) = 1.686, p = 0.103$. As such, we refute this hypothesis. However, by observing Figure 5.5, we can identify that Version A has considerably higher minimum, maximum and Interquartile Range (IQR) values than Version B, which indicates, alongside with results of H1, that the emotional response to the version with dynamic non-diegetic audio (Version A) is slightly more intense, compared to the same game that keeps a neutral background audio track (Version B). With further testing including a higher number of participants, we will be able to see whether this difference is statistically significant.

Tension per Section

One of the objectives of this work is to research the impact that a dynamic background audio across different game sections can have in a horror experience, more precisely in the tension felt by the player. For such, we expect to have stronger changes in self-reported tension between sections, both in parts where we expect to scare the player and parts that have the intention to relax them. We obtained data
for the tension from the post-game questionnaire, to create and accept or refuse the following hypothesis:

**H3**: *The use of dynamic non-diegetic audio tracks in a horror experience promotes a higher change of self-reported tension between sections in players.*

This hypothesis states that playing version A (normal) induces a significant difference in the accumulated absolute tension changes across the game sections, which may suggest that the switch of background non-diegetic audio layers trigger stronger degrees of anxiety and/or relaxation during the experience.

The variable we use is $s_{\text{change}}$, which is the sum of the absolute difference of self-reported tension.
between each pair of game sections (from $s_1$ to $s_8$):

$$s_{\text{change}} = \sum_{i=1}^{8} |s_i - s_{i-1}|$$

Shapiro-Wilk test was not significant for data in both versions of the game (Version A: $D(15) = 0.921, p = 0.197$; Version B: $D(15) = 0.959, p = 0.668$). By comparing means of accumulated change of tension across game sections between the two versions (Version A: $M = 13.8667, SD = 3.35659$; Version B: $M = 11.9333, SD = 4.46361$) using a paired-samples t-test, we verified no statistically significant difference: $t(28) = 1.341, p = 0.191$.

However, as we can observe in Figure 5.6, both conditions have a considerably different IQR and Version A presents a higher median, which may still suggest that using the same background audio track in the entirety of the game in Version B leads to weaker changes in tension. Consequently, considering the circles on the lower end of the boxplot for Version A, we identified two outliers, with values below $Q1/1.5IQR$ ($Q1 = 13; IQR = 2.5$). By removing the two outliers, Shapiro-Wilk test was not significant once again for data in Version A ($D(13) = 0.960, p = 0.746$), and we observed that with the updated data (Version A: $M = 14.8462, SD = 2.23033$) there is statistical significance using paired-samples t-test: $t(26) = 2.130, p = 0.043$.

Figure 5.6: Boxplots of accumulated absolute tension change during the game, for Version A and B.

Moreover, it is important to note that both boxplots (a) and line graph (b) at Figure 5.7 reveal that our game, especially in Version A, presented a similar tension curve to the one we desired and proposed at

\[\text{\textsuperscript{6}}\text{For the sake of simplicity and coherency with the naming conventions we used for the sections during the experiment, we consider the horror sections starting from number 1 to number 8. The tutorial phase of the game, which was out of scope for the evaluation in questionnaires and heart rate, is considered as $s_0$, irrelevant for our formulas.}\]
Chapter 4.1.6. Both curves show an evolution of self-reported tension composed by two peaks: the first one due to the buildup followed by a jump scare, and the second one, more intense, related to the active threat of a moving creature, requiring the player to apply stealth techniques to survive. In addition, as expected, we have Section 5 as the area of the game responsible for relaxing the players.

Once again confirming our hypothesis, we can also verify how Version A promoted a more accentuated change of tension during the game, as from Section 1 to Section 4 the mean tension increases and in Section 5 drastically decreases to a value lower than the one present in Version B. Immediately thereafter, the mean tension scores for the three remaining sections are always slightly higher in Version A.

Considering these observations and the steep changes of tension around Section 5 in Version A, we further analysed the mean differences of tension between each pair of sections (i.e.: $s_{i,i+1}$) and observe the bar chart at Figure 5.8, which depicts the average change of mean self-reported tension between game sections.

By performing Shapiro-Wilk tests for all pairs (normality tests table in Appendix D, at Figure D.1), we showed no evidence against normality ($p > 0.05$) for three pairs in both versions ($s_{1,2}$; $s_{3,4}$; $s_{4,5}$), allowing the use of paired-samples t-tests. Regarding the remaining pairs, there was at least one version with evidence against normality ($p < 0.05$), requiring in these cases the use of non-parametric methods to further analyse them.

We verified statistically significant difference in the transitions adjacent to Section 5, designed for relaxation: from Section 4 to 5 ($s_{4,5}$) with $t(28) = -3.419, p = 0.002$; from Section 5 to 6 ($s_{5,6}$) with Wilcoxon test’s $z$-score of $-2.547$, significant at $p = 0.011$. Additionally, there is also a statistically significant difference from Section 2 and 3 ($s_{2,3}$): while Version A leads to a higher mean of tension,
Version B decreases its value (Wilcoxon test’s $z$-score of $-2.124$, significant at $p = 0.034$). Thus, we observe that the contrast between anxiety inducing music and relaxing music amplify the change of self-perceived tension of the player ($s_{4,5}$ and $s_{5,6}$) and that cutting off background music while introducing short audio effects around the player in a suspenseful context in Version A is a strategy which sustains and elevates the tension felt by the player ($s_{2,3}$).

Regarding the remaining transitions, as we already observed by the absolute scores of tension of each section in Version A, we can detect differences in Figure 5.8, although we cannot prove statistically the significance for all of them. Nonetheless, according to our observations and the statistic evidence mentioned above, we are able to confirm a higher change of self-reported tension between sections in Version A. Hence, we accept Hypothesis 3.

Heart Rate

Unlike the answers we get from questionnaires, biometric signals offer us the possibility to objectively evaluate how players react to a game. In our experiment, we obtained data for the BVP to calculate and analyse the average HR of the players per game section, to accept or refuse the following hypothesis:

**H4:** The use of dynamic non-diegetic audio tracks in a horror experience promotes a higher change of HR between sections in players.

This hypothesis states that playing version A (normal) induces a significant difference in the accumulated absolute HR changes across the game sections, which may suggest that the switch of background non-diegetic audio tracks induce moments of higher anxiety and/or relaxation during the experience. The variable we use is $hr_{change}$, which is the sum of the absolute difference of average HR between each pair of game sections (from $hr_1$ to $hr_8$, i.e. from the average HR of Game Section 1 to Game
Section 8):

\[
hr_{\text{change}} = \sum_{i=1}^{8} |hr_i - hr_{i-1}|
\]

Shapiro-Wilk test was significant for data in both versions of the game (Version A: \(D(15) = 0.863, p < 0.05\); Version B: \(D(15) = 0.656, p < 0.05\)), hence we turned to non-parametric methods to further analyse our data. The Wilcoxon test’s \(z\)-score is \(-2.385\) and this value is significant at \(p = 0.017\). As such, the results indicate that Version A leads to a more intense change of HR during the experience (see Figure 5.9), and therefore, we accept Hypothesis 4.

Furthermore, with Figure 5.10, we spot a more focused normalized mean HR between players across many game sections. For example, it is noteworthy to highlight the last game section, where the IQR in Version A is much smaller and concentrated in the highest values, when compared to Version B, which spreads its values over almost the entire scale. This means that, in Version A, most players reached their relative higher mean HR during Section 8, as opposed to the players in Version B, who reacted very differently from each other in this part of the game, relatively to the other game sections. In other words, we interpret that most players’ HR in A peaked during Section 8, likely to the high anxiety inducing background music, while players’ HR in B peaked more sparsely throughout the experiment.

Moreover, in Figure 5.11 (b), it is noticeable that in Version A, on average, players had a higher normalized mean HR during Sections 4, 5 and 8. This may lead us to conclude that the use of high anxiety inducing background music during Section 4 and Section 8 led these phases of the game to standout and trigger a stronger reaction comparing to the rest of the game, unlike in Version B, where the players’ most intense sections are not collectively as easy to identify. As picture (a) in the same Figure 5.11 suggests, the average HR of players in Version B stays more consistent along the game sections. In fact, the mean average HR per section ranges between 85.66 bpm and 92.86 bpm in Version A, while in Version B it ranges between 84.45 bpm and 86.07 bpm. However, it is relevant to highlight that the results for A present a higher standard deviation in all sections comparing to B (in A, across the eight sections, the minimum is for Section 7 with \(SD = 15.83\) and the maximum is for Section 5 with \(SD = 19.04\); in B, the minimum is for Section 8 with \(SD = 9.59\), and the maximum is for Section 2 with \(SD = 13.26\)). Although we could try to relate the differences in variance to the application of each version of the game, we verify this phenomenon already during Game Section 1, which uses the same background audio layers in both versions (Section 1 in Version A: \(SD = 16.62\); Section 1 in Version B: \(SD = 10.71\)). Therefore, we conclude that these differences in variance values don’t relate to the game experience, but are likely related to the group of users for each version and/or the quality of measurements.\(^7\)

Moreover, it is noticeable how Section 5, originally designed for relaxation, presents a higher average

\(^7\)Although players were given similar conditions, it is relevant to consider the possible imprecisions of the physiological data collected by the PulseSensor, due to the nature of this gadget, its placement position and other unknown factors.
HR than the previous section in both versions of the game, and, right after, during Section 6, we finally spot a decrease of average HR, especially in Version A. This phenomenon may be due to two key reasons: (1) Players reported that they felt attentive and intrigued while discovering this zone of the game, which may have led them to a higher arousal, which is what the HR mostly captures in a user, instead of their valence. Due to its novelty factor, the presence of a different background music may have also contributed to the continuation of the considerably high average HR in Version A. (2) Since Section 5 is relatively short (duration of less than 1 minute), its relaxing effect might have only truly affected the player’s body reactions afterwards, already during Section 6.

Figure 5.9: Boxplots of accumulated absolute change of average HR along the game sections, for Version A and B.

Figure 5.10: Boxplots of normalized average HR per section, for Version A and B.
Comparison: Self-Reported Tension and Heart Rate

In relation to how the results from the questionnaire can be compared to the collected HR of the players, we elaborated the following hypothesis:

**H5:** The evolution of normalized average HR between consecutive sections correlates to the evolution of normalized self-reported tension between the correspondent consecutive sections

This hypothesis states that the changes of normalized average HR between each pair of consecutive sections correlate to the changes of the correspondent normalized tension values. By studying this hypothesis, we expect to answer if the evolution of the HR obtained from the BVP translates to the tension reported by the players and, consequently, if it is enough to capture how the player felt during the horror experience, regarding their level of anxiety.

We decided to perform correlation tests for all users ($N = 30$), using the difference of normalized self-reported tension and average HR between each pair of sections (from $s_{1,2,norm}$ to $s_{7,8,norm}$; from $hr_{1,2,norm}$ to $hr_{7,8,norm}$). We checked the normality of each pair of variables to be studied. The normality Shapiro-Wilk tests were not significant for data, with the exception of the difference of normalized tension between Section 1 and 2 ($s_{1,2,norm}$) and Section 2 and 3 ($s_{2,3,norm}$), as seen in the normality tests table in Appendix D, at Figure D.1. By studying the correlation tables in Figures D.2 and D.3, we verify no statistically significant correlation in the pairs of self-reported tension and average HR for each game section transition ($p > 0.05$), with the exception of Section 1 to Section 2, presenting a negative correlation of: $p = 0.005$, $r = -0.502$. This specific pair may indicate how the perception of tension for the players can strongly dissociate from their actual body response and the anxiety they might truly feel during the experience. Furthermore, when observing the scatter plots in Figure 5.12, we conclude that

---

*Pearson Correlation was used in all tests, except the first two pairs (Sections 1/2 and Sections 2/3), with Spearman’s Rho.*
the average HR is very unreliable to dictate what is the current level of anxiety of the players. Although we are using the mean HR during entire game sections for this analysis and the player’s emotional state can vary more frequently, the discrepancy of these values with the self-reported tension and the fact that, for correlation, many of these section transitions have $r < 0$ and no statistical significance ($p > 0.05$), reveal that it is not possible to relate directly the HR with the players’ self-perceived anxiety. Thus, we refute the Hypothesis 5.

5.4 Discussion

Overall, our results show some effects of using different background audio tracks across the game sections in the players’ self-perceived tension and emotions, and HR. We were able to prove that it led to more intense positive emotions and triggered a stronger accumulative change of tension and HR. We also verified a stronger change of self-reported tension when transitioning to and from Game Section 6 (designed as a relaxation zone), compared to a version of the game that keeps the same background track in all sections. However, two hypotheses were refuted: negative emotions such as fear were slightly higher in the dynamic-based audio background version (A), but with no statistical significance; we did not obtain significant results to prove a correlation between the players’ evolution of self-reported tension and their collected HR.

There were some limitations to this study that may explain the results and that should be addressed in future studies. One possible limitation in this experiment was the personality and vulnerability differences between each person towards a horror experience. We asked the users about their familiarity with the genre and if they enjoy it, but, unfortunately, such information doesn’t clarify how each player is expected to react to horror content that is designed to trigger anxiety, for instance. Furthermore, these issues are aggravated with the reduced number of users in our sample. As such, in order to obtain better results we should repeat the experiment with a larger population.

Regarding the use of physiological signals in this study, we highlight again the fact that HRV could be potentially a better indicator of the impact left by the game, instead of the instantaneous HR. However, in order to calculate HRV, it would be necessary to use a different and less comfortable setup, where, instead of collecting the BVP through the PulseSensor, we would need to use electrodes to obtain the ECG. And, as suggested before, other methods such as EEG, followed by the construction of classifiers based on many users’ data, may lead us to deeper insights regarding our research questions and the implementation of biofeedback for sound choice and timing manipulation. Indeed, our HR data and results of correlation with the self-reported tension showed how we cannot rely solely on a PulseSensor to gauge the real level of fear felt by the player, mainly because it does not provide a full view on the emotional state of a user and presents severe inconsistencies in its readings, sometimes due to its
Figure 5.12: Scatter plots for self-reported tension vs average HR per section change.
placement position, high sensitivity to light and movement, and other unknown factors. Even so, the collected HR provides many hints regarding the players’ arousal, considering that, for instance, many users had their peaks of HR during the final section, agreed as the most tense part of the game.
Conclusion and Future Work
This study was conducted with the goal of better understanding what is the impact of non-diegetic background audio (including soundtrack music) that changes across different sections in a horror game, compared to the use of the same background audio track during the entirety of the walkthrough.

To investigate this problem, we created hypotheses to verify if using multiple audio tracks in a horror game leads to stronger affective reactions (regarding both self-perceived emotions and physiological signals) while listening to different audio tracks appropriate to the desired state of each game section. To answer our research questions and accept or refute these hypotheses, a psychological survival horror game was created from scratch, Presa, taking place in a dark cave system, where progress is made by trying to escape it, without getting caught by malicious enemies. Users played Version A (default), with different non-diegetic audio tracks according to the desired state in each section, or Version B, always with the same neutral audio track.

Although there were empirical evidences that players would have stronger affective reactions while listening to different audio tracks appropriate to the desired state of each game section (Version A), compared to a version of the same game showcasing always the same track (Version B), we had the opportunity, in this study, to gauge the real impact of these differences by measuring the players’ self-perceived affective response to the game, self-perceived tension per section and their mean HR per section. Results showed statistical evidence for positive emotions, which suggest that players with Version A reacted more intensely to the game. Although we had no statistical evidence for negative emotions, we also noticed slightly stronger negative emotions in Version A. We also found that players showed more intense changes of self-perceived tension across the game sections while playing Version A, which confirms, once again, the importance of using tense and calm tracks in the appropriate moments of a horror game. Furthermore, regarding the collection of the players’ HR, we were also able to confirm that Version A led to a higher average change of mean HR between game sections. However, no correlations were found between the change of normalized average values of HR and the change of normalized means of tension, for each pair of consecutive game sections.

While we have verified the presence of promising results with the instantaneous HR, calculated from the BVP, regarding the players’ arousal during the horror experiment, it is relevant to consider other physiological alternatives, measurements and strategies to better reflect the actual level of anxiety felt by the players. Although we chose a PPG sensor for its convenience and comfort, we suggest the use of ECG sensors in future iterations of the work, since they allow a more reliable measurement of the players’ HR and HRV, which may offer new insights regarding how players objectively felt during the experience. Additionally, signals such as EEG are frequently used in experiments with biofeedback as well and might also provide a deeper understanding of the differences between our two versions of the game. Many researches that already incorporated biofeedback in horror games also developed predictive models, with multiple physiological signals, which may provide more promising results in the
long term, even though such strategy might require larger populations.

In the end, we managed to accomplish our main goal which was to understand the impact that the change of background music during different game sections can have on the players in a short psychological horror game, both on their self-perceived tension and HR, although the latter is confirmed not to be sufficient to define the player’s emotional state. Besides the approach of new evaluation strategies as the abovementioned, we also encourage the use of other effective horror games for similar experiments, since it is not possible to generalize such results to the entire genre, considering that the execution and presentation of horror elements and all artistic facets, including audio, may differ to a great extent between each game.
Bibliography


Audio Assessment Questionnaire

This appendix contains a sample of the questionnaire for the audio assessment preliminary experiment. Each participant received the same form content with a different order of questions.
Audio Assessment Questionnaire
As a part of a MSc Dissertation at Instituto Superior Técnico, we are conducting a questionnaire to evaluate the level of anxiety that a set of audio tracks can induce.

We guarantee that the data collected will be analyzed and treated confidentially. If you have any questions, contact pedro.a.soares@tecnico.ulisboa.pt.

The estimated time to answer this questionnaire is 3-5 minutes. Thank you for your participation!

**NOTE: Headphones are highly recommended for this experience**

*Obrigatório

| Audio Tracks | Please listen to the entirety of the following tracks and evaluate from 1 to 7 the level of anxiety each one triggered. **NOTE: Headphones are highly recommended for this experience** |

1. Audio "&": [https://youtu.be/6B6LYAcf-as](https://youtu.be/6B6LYAcf-as)

   *Marcar apenas uma oval.*

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Calm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


   *Marcar apenas uma oval.*

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Calm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Very Anxious
3. Audio "??": [https://youtu.be/7MSi-lhT5wM](https://youtu.be/7MSi-lhT5wM)

Marcar apenas uma oval.

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Very Calm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Very Anxious</td>
</tr>
</tbody>
</table>


Marcar apenas uma oval.

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Very Calm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Very Anxious</td>
</tr>
</tbody>
</table>

Este conteúdo não foi criado nem aprovado pela Google.
This appendix contains a sample of the questionnaire for the final experiment, composed of a demographic form and a post-game form regarding the level of tension felt during each game section and the emotions felt during the entire experience.
This questionnaire is part of a Master’s Thesis by Pedro Soares, a student of Instituto Superior Técnico, University of Lisbon.

We are conducting an observational study on the impact of the presence and choice of background sound effects and environments on a playable horror experience. The objective of this session is to play a short horror game (survival horror), identify the tension felt at different stages of the experience, and collect your physiological signs of cardiac output during the demonstration.

The experience will take the following steps: firstly, we ask you some questions about yourself and your gaming habits. Secondly, we setup the necessary equipment, watch a set of relaxing pictures and play the game. Lastly, you will answer to questions regarding the experience. The full experiment should take about 35-50 minutes.

All data collected will be kept confidential and will be analyzed exclusively by the researchers of this project. The data may also be used for presentation or exhibition of results, duly pseudonymized, in scientific publications, conferences or similar events.

Before proceeding, verify if you’ve already carefully read and filled in the consent form. By proceeding, you are giving your consent to the conditions mentioned in the consent form.

**Obrigatório**

Demographic Questionnaire

Before you start playing our game and follow our next steps, please fill out this demographic survey.

1. What is your gender? *

   *Marcar apenas uma oval.*

   - Female
   - Male
   - Other

2. What is your age? *
3. How much time do you spend playing video games? *

    *Markar apenas uma oval.*

    ☐ I don't play video games
    ☐ I play video games when the opportunity presents itself
    ☐ I make time in my schedule to play video games

4. Do you like horror games/films? *

    *Markar apenas uma oval.*

    ☐ Yes
    ☐ No
    ☐ Neutral

5. Do you agree or disagree with the following statement: "I find it easy to navigate in first-person perspective games (e.g. "P.T.", "Firewatch", "Resident Evil 7", "The Witness", "Call of Duty")" *

    *Markar apenas uma oval.*

    1 2 3 4 5

    Totally disagree ☐ ☐ ☐ ☐ ☐   Totally agree

---

Preparation for the demonstration

Now that you filled in our consent and demographic forms, the assistant is going to prepare everything for the experience and clarify any doubts you may have, before you start playing the game.

The estimated duration of the game is 10-20 minutes. Before playing, you'll watch a set of relaxing pictures during 5 minutes.

Have fun!
Thank you for playing the game! Now, please fill this final form (divided in two pages).

In this page, we ask you to evaluate the tension you felt during each section.

To help you remember each part of the game, the assistant will show you the recorded playthrough and describe what each section includes in the game.

6. Please indicate the level of tension you felt during each section of the game. *

*Marcar apenas uma oval por linha.*

<table>
<thead>
<tr>
<th>Section 1</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 - Moderately tense</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9 - Extreme tense</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section 2</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 - Moderately tense</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9 - Extreme tense</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section 3</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 - Moderately tense</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9 - Extreme tense</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section 4</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 - Moderately tense</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9 - Extreme tense</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section 5</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 - Moderately tense</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9 - Extreme tense</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section 6</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 - Moderately tense</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9 - Extreme tense</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section 7</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 - Moderately tense</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9 - Extreme tense</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section 8</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 - Moderately tense</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9 - Extreme tense</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
This final page consists of a number of words that describe different feelings and emotions.

Read each item and then mark the appropriate answer (1 to 5) to that word.
7. Indicate to what extent you felt the following emotions and feelings during the experience.*

*Marcar apenas uma oval por linha.*

<table>
<thead>
<tr>
<th></th>
<th>1 - Very slightly or not at all</th>
<th>2 - A little</th>
<th>3 - Moderately</th>
<th>4 - Quite a bit</th>
<th>5 - Extremely</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interested</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distressed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excited</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upset</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strong</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guilty</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scared</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hostile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enthusiastic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proud</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irritable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alert</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ashamed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inspired</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nervous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Determined</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attentive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jittery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afraid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Pilot Experiment Questionnaire - Additional Question

This appendix contains a sample of the additional set of post-game questions of an adapted core-GEQ, assessing “Flow” and “Sensory and Imaginative Immersion”, for the pilot tests. The remaining of the questionnaire presents the same content as the provided at B.
7. Please indicate how, in general, you felt while playing the game for each of the items. *

Marca apenas uma oval por linha.

<table>
<thead>
<tr>
<th></th>
<th>1 - Not at all</th>
<th>2 - Slightly</th>
<th>3 - Moderately</th>
<th>4 - Fairly</th>
<th>5 - Extremely</th>
</tr>
</thead>
<tbody>
<tr>
<td>I was fully occupied with the game</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I forgot everything around me</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I lost track of time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I was deeply concentrated in the game</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I lost connection with the outside world</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I was interested in the game’s story</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>It was aesthetically pleasing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I felt imaginative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I felt that I could explore things</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I found it impressive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>It felt like a rich experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Post-Game Statistical Tests Results

This appendix contains tables with the results of normality tests, paired samples t-tests, Man-Whitney U tests, and bivariate Pearson correlations, from the final experiment’s data.
### Tests of Normality

<table>
<thead>
<tr>
<th>Version</th>
<th>Statistic</th>
<th>df</th>
<th>Sig.</th>
<th>Kolmogorov-Smirnov*</th>
<th>Statistic</th>
<th>df</th>
<th>Sig.</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td>s1_2 A</td>
<td>.180</td>
<td>15</td>
<td>.200*</td>
<td>.885</td>
<td>15</td>
<td>.056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>.233</td>
<td>15</td>
<td>.027</td>
<td>.901</td>
<td>15</td>
<td>.099</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s2_3 A</td>
<td>.208</td>
<td>15</td>
<td>.081</td>
<td>.932</td>
<td>15</td>
<td>.283</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>.196</td>
<td>15</td>
<td>.124</td>
<td>.852</td>
<td>15</td>
<td>.019</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s3_4 A</td>
<td>.201</td>
<td>15</td>
<td>.105</td>
<td>.919</td>
<td>15</td>
<td>.187</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>.181</td>
<td>15</td>
<td>.200*</td>
<td>.946</td>
<td>15</td>
<td>.468</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s4_5 A</td>
<td>.223</td>
<td>15</td>
<td>.043</td>
<td>.923</td>
<td>15</td>
<td>.210</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>.217</td>
<td>15</td>
<td>.055</td>
<td>.884</td>
<td>15</td>
<td>.055</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s5_6 A</td>
<td>.281</td>
<td>15</td>
<td>.002</td>
<td>.832</td>
<td>15</td>
<td>.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>.181</td>
<td>15</td>
<td>.200*</td>
<td>.938</td>
<td>15</td>
<td>.358</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s6_7 A</td>
<td>.279</td>
<td>15</td>
<td>.003</td>
<td>.844</td>
<td>15</td>
<td>.014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>.198</td>
<td>15</td>
<td>.117</td>
<td>.957</td>
<td>15</td>
<td>.641</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s7_8 A</td>
<td>.251</td>
<td>15</td>
<td>.012</td>
<td>.798</td>
<td>15</td>
<td>.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>.322</td>
<td>15</td>
<td>.000</td>
<td>.770</td>
<td>15</td>
<td>.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* This is a lower bound of the true significance.
a. Lilliefors Significance Correction

### Mann-Whitney Test

<table>
<thead>
<tr>
<th>Ranks</th>
<th>Version</th>
<th>N</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td>s2_3</td>
<td>A</td>
<td>15</td>
<td>18.80</td>
<td>262.00</td>
</tr>
<tr>
<td>B</td>
<td>15</td>
<td>12.20</td>
<td>183.00</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s1_2</td>
<td>A</td>
<td>15</td>
<td>19.47</td>
<td>262.00</td>
</tr>
<tr>
<td>B</td>
<td>15</td>
<td>11.53</td>
<td>173.00</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s6_7</td>
<td>A</td>
<td>15</td>
<td>15.67</td>
<td>235.00</td>
</tr>
<tr>
<td>B</td>
<td>15</td>
<td>15.33</td>
<td>230.00</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s7_8</td>
<td>A</td>
<td>15</td>
<td>14.50</td>
<td>217.50</td>
</tr>
<tr>
<td>B</td>
<td>15</td>
<td>18.00</td>
<td>247.50</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Test Statistics*

<table>
<thead>
<tr>
<th></th>
<th>s2_3</th>
<th>s1_2</th>
<th>s6_7</th>
<th>s7_8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>63.025</td>
<td>53.020</td>
<td>110.000</td>
<td>97.500</td>
</tr>
<tr>
<td>Wilcoxon V</td>
<td>183.000</td>
<td>173.000</td>
<td>230.000</td>
<td>217.500</td>
</tr>
<tr>
<td>Z</td>
<td>-2.124</td>
<td>-2.547</td>
<td>-1.007</td>
<td>-0.947</td>
</tr>
<tr>
<td>Asymp. Sig (2-tailed)</td>
<td>.034</td>
<td>.011</td>
<td>.305</td>
<td>.318</td>
</tr>
<tr>
<td>Exact Sig (2-tailed)</td>
<td>.441*</td>
<td>.201*</td>
<td>.939*</td>
<td>.539*</td>
</tr>
</tbody>
</table>

*a. Grouping Variable: Version
b. Not corrected for ties.

### Independent Samples Test

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>df</th>
<th>Sig.</th>
<th>Leven's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>r</td>
<td>Sig</td>
<td>l</td>
</tr>
<tr>
<td>s1_2</td>
<td></td>
<td></td>
<td></td>
<td>.150</td>
<td>.701</td>
<td>.599</td>
</tr>
<tr>
<td>s1_2</td>
<td></td>
<td></td>
<td></td>
<td>.599</td>
<td>26.483</td>
<td>.554</td>
</tr>
<tr>
<td>s3_4</td>
<td></td>
<td></td>
<td></td>
<td>.016</td>
<td>.901</td>
<td>-1.375</td>
</tr>
<tr>
<td>s3_4</td>
<td></td>
<td></td>
<td></td>
<td>-1.375</td>
<td>27.928</td>
<td>1.80000</td>
</tr>
<tr>
<td>s4_5</td>
<td></td>
<td></td>
<td></td>
<td>.001</td>
<td>.979</td>
<td>-3.419</td>
</tr>
<tr>
<td>s4_5</td>
<td></td>
<td></td>
<td></td>
<td>-3.419</td>
<td>27.806</td>
<td>.002</td>
</tr>
</tbody>
</table>

### Figure D.1: Normality, Mann-Whitney and Independent Samples tests for versions A and B, for average self-reported tension difference between each pair of successive game sections (from s1_2, which is the first two sections, to s7_8, A).
Tests of Normality

<table>
<thead>
<tr>
<th>Kolmogorov-Smirnov*</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
<td>df</td>
</tr>
<tr>
<td>s1_2_norm</td>
<td>215</td>
</tr>
<tr>
<td>s2_3_norm</td>
<td>197</td>
</tr>
<tr>
<td>s3_4_norm</td>
<td>155</td>
</tr>
<tr>
<td>s4_5_norm</td>
<td>137</td>
</tr>
<tr>
<td>s5_6_norm</td>
<td>141</td>
</tr>
<tr>
<td>s6_7_norm</td>
<td>135</td>
</tr>
<tr>
<td>s7_8_norm</td>
<td>132</td>
</tr>
<tr>
<td>h1_2_norm</td>
<td>071</td>
</tr>
<tr>
<td>h2_3_norm</td>
<td>099</td>
</tr>
<tr>
<td>h3_4_norm</td>
<td>102</td>
</tr>
<tr>
<td>h4_5_norm</td>
<td>099</td>
</tr>
<tr>
<td>h5_6_norm</td>
<td>099</td>
</tr>
<tr>
<td>h6_7_norm</td>
<td>104</td>
</tr>
<tr>
<td>h7_8_norm</td>
<td>112</td>
</tr>
</tbody>
</table>

Figure D.2: Normality tests for differences of normalized self-reported tension and average HR between sections ($t_{ij+1, \text{norm}}$ and $h_{ij+1, \text{norm}}$).

<table>
<thead>
<tr>
<th>Spearman's rho</th>
<th>s1_2_norm</th>
<th>hr1_2_norm</th>
<th>hr2_3_norm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation Coefficient</td>
<td>-402</td>
<td>403</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.005</td>
<td>.027</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>30</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spearman's rho</th>
<th>s2_3_norm</th>
<th>hr2_3_norm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation Coefficient</td>
<td>.237</td>
<td>-159</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.207</td>
<td>.400</td>
</tr>
<tr>
<td>N</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Figure D.3: Pearson and Spearman's Rank Correlations among differences of normalized self-reported tension and average HR between sections.
This appendix contains the consent form, which students filled out to confirm that they wished to participate in the final experiment.
Informed Consent

Dear participant,

We are conducting an observational study on the impact of the presence and choice of background sound effects and environments on a playable horror experience. The objective of this session is to play a short horror game (survival horror), identify the tension felt at different stages of the experience, and collect your physiological signs of cardiac output during the demonstration.

During the session, your blood volume pulse data will be recorded. All data collected will be kept confidential and will be analyzed exclusively by the researchers of this project. The data may also be used to present or display results, duly pseudonymized, in scientific publications, conferences or similar events.

Your participation is voluntary and you can always withdraw at any time without any penalty or consequence.

To participate in this experiment, we ask that you read the informed consent and if you agree to participate in accordance with the terms below, we ask that you sign the form at the location indicated.

Thanks for your collaboration!
1 - I read and understood the meaning of this study. I had the opportunity to ask questions, if necessary, and collect the corresponding answers.

2 - I understand that participation in this study is voluntary and that I can withdraw at any time, without giving any explanation. If this happens, I will not be subject to any penalty and the data relating to my experience will be removed and destroyed.

3 - I authorize the recording of my physiological data (cardiac output) during the session.

4 - I authorize the processing of physiological, game and questionnaire data within the scope of this project for the purposes of analysis, investigation and dissemination of results in scientific publications or conferences in the project area, by the researchers of this project.

5 - I understood that the data collected in this study will be used as mentioned above.

6 - As described above, I authorize my participation in this study and accept its conditions.

The participant Date

_________________________________ _____________________________

Responsible Researcher:

Pedro Soares
Instituto Superior Técnico - University of Lisbon

The participant will be given a signed copy of this form.