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

Pedro Amaral Soares
pedro.a.soares@tecnico.ulisboa.pt
Instituto Superior Técnico, Lisboa, Portugal

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: Videogames, Audio, Soundtrack, Horror, Psychophysiology, Heart Rate.

1. Introduction

A main challenge for horror 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 [26].

Considering that "(...) sonic dimensions, like music, are highly suitable to take over other senses during the initial direction of selective attention towards congruent perceptions"[15], and that background music promotes strong affective reactions and memorization of key events in media [3, 6], the audio facet presents itself as a key element yet to be fully studied in games and to be further explored with biofeedback[1], with the goal of influencing and triggering specific reactions or elicit fear, in the context of horror gaming experiences [14, 13, 11, 34, 35]. Klimmt et al. [21] 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, and if we desire to apply biofeedback for audio choices in future work, considering its potential to improve and adapt horror gaming experiences to all users, 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 diegetic, where the source is visible or can be implied in the action of the experience, or non-diegetic, which comes from a source out of context, including background music. As we start identifying that horror can take many different approaches regarding non-diegetic audio, from the selection of music to the appearance of unknown and unexpected sounds, we 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?**

We hypothesize that: players have a better horror experience (with higher affect and changes of tension and

[1] A game with biofeedback is enhanced by physiological signals from the player, to change or add game mechanics, or multiple artistic facets.

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 Section 5, we explore and subdivide the hypotheses, and share the evaluation process and results.

We crafted and ran two versions of a short horror game, *Presa*, for further comparison in user results, where, in one (Version A), we use the default evolution of desired tension, 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.

2. Background

In this section, background necessary for developing this work such as fundamentals of Horror and Sound will be introduced, along with relevant work on sound and biofeedback in horror games.

2.1. Concepts and Techniques in Horror Media

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 [24], being fear broadly considered as an emotional state. Ekman et al.'s research [12] 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 [27, 17].

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) [16]. 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 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"[31]. Horror often encourages fixation and is object-focused [13, 29, 31]. In this work, depending on the context, the term *horror* may also refer to the genre itself, instead of the emotional state.

Anxiety, which resembles fear, refers to a more general, diffuse, objectless [1] and long-term state of distress [5], 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"[4]. For the sake of simplicity and due to the ambiguity of soundscapes that can often generate distress without a defined source of threat, in our work we use 'anxiety' as the central definition for the emotional state we desire to elicit throughout the playthroughs. 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 relevant term is *suspense*, commonly referred in fiction media, being a desirable emotional sensation to trigger in any horror experience [13, 22] and considered as "an experience of uncertainty regarding the outcome of a potentially hostile confrontation", "whose hedonic properties can vary from noxious to pleasant" [39, 38], only possible with the presence of fear [13].

2.2. Audio Techniques in Horror

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 [3, 6]. As Klimmt et al. [21] mentions, Cohen [8, 9] proposed the Congruence-Associationist Model (CAM), which 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 [2, 3]. Since the study of CAM has been successfully applied in film music [3, 10], it can be used similarly as a theoretical foundation for the context of video games. 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 melodics, to fit the overall narrative setting, mood, and visual aesthetics of the project [21]. 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[21].

To fully take advantage of this model 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).

In horror media, such as films, it is known that prior knowledge of upcoming events would generate increased sensations of fright and upset [7]. 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 [14, 13]. Also, 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", helps eliciting suspense, by forewarning the player [33].

Sounds derived from what we hear everyday, many generated by people, might be effective at eliciting stress and anxiety [23], especially if they are defamiliarized to lead to perceptions of uncanniness [18]. Therefore, the game created for our study relies on sounds that both generate some familiarity, but also unease, such as heavy and distorted breathing and gurgling.

Foreshadowing is attributed as a technique, commonly in narrative, of incorporating cues of lower significance to anticipate a relevant event [32]. 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 [32]. The technique of foreshadowing is 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. Indeed, sound stimuli are some of the ones that trigger the most terror and anticipation [11, 34, 35], and silence is often a relevant strategy to trigger shock, since it is the most effective way may be to preceed a relevant horrific event by the absence of sound [37]. 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, it is possible to create long lasting tension.

Graja et al. [17], used an adaptation of *P.T.*[2] to analyse the players emotional reactions towards context alteration, mainly with changes to light effects, sounds and in-game events. 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. Other studies, such as Prager et al. [28] also concluded that "the audio facet has a bigger impact on the player's perception of the game content than the visual facet".

[2] Konami, 2014

Similarly to the direction of our work, Klimmt et al. [21] have studied the effects of soundtrack music on the video game enjoyment, including in horror, by exposing players to a version with and another without music. The correlations from questionnaire responses suggested that music increases the players' horror experience. Based on the reported enjoyment and emotions (eight-item scale on video-game enjoyment by Klimmt et al. [20] - 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 $r = 0.58$).

2.3. Biofeedback in Horror

In order to improve and adapt horror gaming experiences to all users, many researchers have approached the use of biofeedback in game mechanics, or multiple artistic facets.

The majority of the current research regarding the use of a biofeedback mechanism in horror 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, such as difficulty [35], 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" [25]. Most common tested changes involve variables directly connected to the gameplay mechanisms or enemy AI [25, 19].

Taking into account how [19] and [11] detected and transitioned between the different affective states, and how [25] considered a positive correlation for HR/arousal and a negative correlation for HR/valence for the horror genre, we 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 is implicit in our affective state of anxiety, depending on the actual danger perceived by each player, which varies from situation to situation. For a low HR, compared to the corresponding average value of the player, we consider they are bored or calm (i.e. not feeling tension); for a high HR, we 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).

3. System Overview

This section describes the two core components of the system and their key development steps.

3.1. Presa

Presa is a psychological first-person horror made from scratch, taking place in a dark cave system, composed of several interconnected sections that need to be traversed[3], populated by creatures with a deformed humanoid presentation. 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, 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.

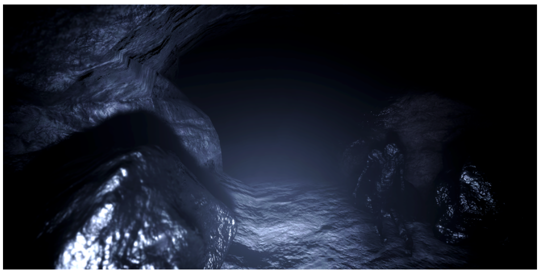

Figure 1: Screenshot of Presa - **Gameplay Snippet**.

In total, the game presents eight sections, which end up describing the progress in the game, 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). These sections' timestamps are then considered in the evaluation for self-reported tension and HR, for further comparison and analysis.

- **Section 1**: From the start of the game, where the player is confronted with the environment, up until the first orb collection;
- **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);
- **Section 3**: While going back, the player observes a different path, with different visual elements;
- **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;
- **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;
- **Section 7**: Region that covers the second exit gate (still locked due to the lack of three orbs), with various static creatures and headstones;
- **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.

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 the previous section, and the results from our preliminary study assessing four original tracks.[4]

- **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 illustrative diagram at Figure 2 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.

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).

Besides walking and interacting with objects, the player can use a sonar tracker, to detect the presence of nearby items (orbs or batteries) and enemies. This

[3] 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.

[4] The suspense track is used in the default version to create contrast between critical moments of the game - although this track was not evaluated previously, its relevance and effect to anticipate a horrific event is theoretically expected based on our related work [37].

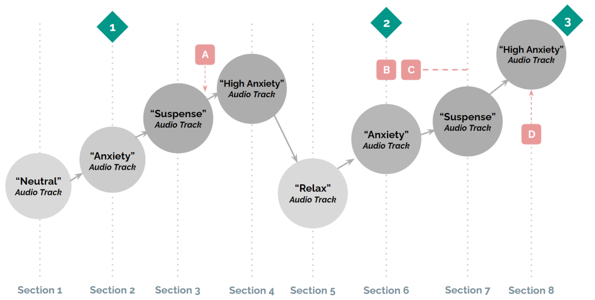

Figure 2: Illustrative diagram of the anxiety evolution design for the game, with the non-diegetic audio track selected for each section, for the default version of the game. The diamond shapes represent the collection of each orb. The square shapes represent the moments where the player finds one or 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).

accessory is not required to complete the game, but is highly recommended during the last section. Inspired by Toprac et al. [33], 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 or enemies' behaviour.

3.2. Bio Framework and Architecture

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.

Among several physiological measures, the heart rate 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 Galvanic Skin Response (GSR), Electroencephalography (EEG) and Electrocardiography (ECG) which need electrodes were discarded.

For our solution, we decided to use a Photoplethysmogram (PPG) sensor 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.

Figure 3 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.

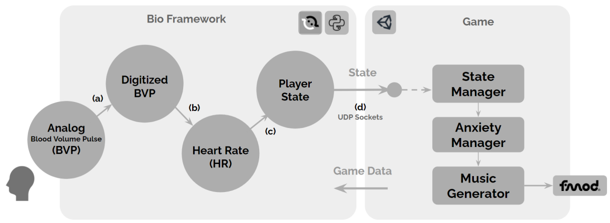

Figure 3: System architecture, with a module responsible for physiological data and another for the game. From left to right, it is composed by user raw data (analog BVP), Bitalino with a PulseSensor attached **(a)**, BioSPPy **(b)** and data treatment **(c)**. First, the user produces analog signals 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.

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, based on the game conditions and goal mental states sent by the python script, from the side of the Bio Framework.

4. Evaluation

This section introduces the second testing phase of this work, presents the results and discusses them.

4.1. 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.

In order to classify, sort and verify the degrees of anxiety induced by the background non-diegetic/music tracks, we conducted a questionnaire 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").

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. 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 Presa.

4.2. 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: "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?"; "Is the HR a good indicator of the player's anxiety state?". To answer these questions, we validate a set of hypotheses described in 4.3.2.

For this experiment we elaborated two versions of the game "Presa", 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 during the entirety of the game.

4.2.1 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: $M = 16 : 17; SD = 03 : 48; SE = 00 : 42$).

4.2.2 User Evaluation

This testing phase had a consent form and an adapted version of the Game Engagement Questionnaire (GEQ) [21]. We used five items from the adapted GEQ. Four address flow and one presence. We use them to see which version led to a higher self-perceived flow

All users were provided a form, completely written in English, which included a questionnaire about their self-perceived tension per section (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"), a PANAS questionnaire to assess one's self-perceived well-being and affectivity, with the original twenty items [36].

4.3. Apparatus

To collect the instantaneous HR, we used a BITalino[5], from the (r)evolution Plugged Kit BT[6], to record the Blood Volume Pulse (BVP) of the player with a PPG sensor - the PulseSensor[7]. 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

[5] https://bitalino.com/
[6] https://bitalino.com/products/plugged-kit-bt
[7] https://plux.info/sensors/42-pulsesensor.html

and which fingers would be least needed.

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)[8].

4.3.1 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 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[9], with high values of valence and low-mid values of arousal, which relates to relaxing stimuli (in a 7-point Likert scale - valence: $M > 5.9; SD <= 1.0$; arousal: $M < 4.0, SD <= 2.0$). 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.

4.3.2 Results

Firstly, we use the data obtained from the PANAS to perform comparisons regarding positive and negative emotions felt while playing the game.

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

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: $D(15) = 0.958, p = 0.663$). 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, we identified that Version A has considerably higher minimum, maximum and 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.

One of the objectives of this work is to research

[8] bitalino.com/downloads/software

[9] 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.

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_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)[10]:

$$s\_change = \sum_{i=1}^{8} |s_i - s_{i-1}|$$

Considering the circles on the lower end of the boxplot for Version A (see Figure 4), 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 for data in both version (Version A: $D(13) = 0.960, p = 0.746$; 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: $M = 14.8462, SD = 2.23033$; Version B: $M = 11.9333, SD = 4.46361$) using a paired-samples t-test, we verified that there is statistically significant difference: $t(26) = 2.130, p = 0.043$.

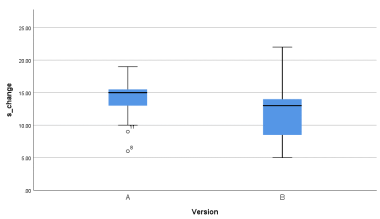

Figure 4: Boxplots of accumulated absolute tension change during the game, for Version A and B.

Moreover, it is important to note that the line graph **(b)** at Figure 5 reveals that our game, especially in Version A, presented a similar tension curve to the one we desired and proposed for Presa. 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.

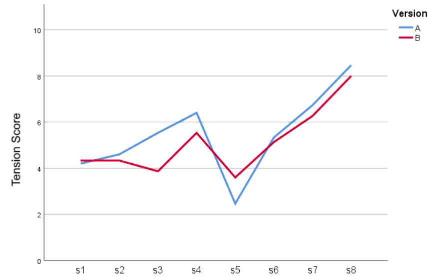

Figure 5: Means of Tension per Section for Versions A and B.

Furthermore, 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):

$$hr\_change = \sum_{i=1}^{8} |hr_i - hr_{i-1}|$$

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

[10]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

to non-parametric methods to further analyse our data. The Wilxocon 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, and therefore, we accept Hypothesis 4.

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 ($s_{2\_3\_norm}$). We verified no statistically significant correlation[11] 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. 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.

4.3.3 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.

5. Conclusions

This study was conducted with the goal of better understanding user differences in their HR and self-perceived tension and emotions, by playing two different versions of the same horror game, only differing in the use of non-diegetic background audio. More precisely, we had players hearing distinct tracks across the game sections, according to a desired mental state, and other players hearing the same background track during the entire game.

To answer our research questions and accept or refute their consequent 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.

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), 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, as expected and hypothesized (**H3**), players suffered 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

[11] Pearson Correlation was used in all tests, except the first two pairs (Sections 1/2 and Sections 2/3), with Spearman's Rho.

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 (**H4**). 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, which led us to refute **H5**. 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.

References

[1] D. H. Barlow. *Anxiety and its disorders: The nature and treatment of anxiety and panic.* Guilford press, 2004.

[2] M. Boltz. Musical soundtracks as a schematic influence on the cognitive processing of filmed events. *Music Perception*, 18:427–454, 2001.

[3] M. Boltz. The cognitive processing of film and musical soundtracks. *Memory & Cognition*, 32:1194–1205, 2004.

[4] J. Bourke. *Fear: A cultural history*. Counterpoint Press, 2006.

[5] J. S. Brown, H. I. Kalish, and I. Farber. Conditioned fear as revealed by magnitude of startle response to an auditory stimulus. *Journal of experimental psychology*, 41(5):317, 1951.

[6] C. Bullerjahn and M. Güldenring. An empirical investigation of effects of film music using qualitative content analysis. *Psychomusicology: A Journal of Research in Music Cognition*, 13:99–118, 1994.

[7] J. Cantor, D. Ziemke, and G. G. Sparks. Effect of forewarning on emotional responses to a horror film. 1984.

[8] A. Cohen. Music as a source of emotion in film. In *P. N. Juslin & J. A. Sloboda (Eds.), Music and emotion: Theory and research*, pages 249–272. Oxford University Press, 2001.

[9] A. Cohen. Congruence-association model of music and multimedia: Origin and evolution. 2013.

[10] A. Cohen. Congruence-association model and experiments in film music: Toward interdisciplinary collaboration. *Music and the Moving Image*, 8:24 – 5, 2015.

[11] A. Dekker and E. Champion. Please biofeed the zombies: enhancing the gameplay and display of a horror game using biofeedback. In *DiGRA'07-Proceedings of the 2007 DiGRA International Conference: Situated Play*, pages 550–558, 2007.

[12] P. Ekman and W. Friesen. Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17 2:124–9, 1971.

[13] T. Garner and M. Grimshaw. A climate of fear: considerations for designing a virtual acoustic ecology of fear. In *Proceedings of the 6th Audio Mostly Conference: A Conference on Interaction with Sound*, pages 31–38, 2011.

[14] T. Garner, M. Grimshaw, and D. A. Nabi. A preliminary experiment to assess the fear value of preselected sound parameters in a survival horror game. In *Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound*, pages 1–9, 2010.

[15] H.-P. Gasselseder. Those who played were listening to the music? immersion and dynamic music in the ludonarrative. In *2014 4th International Workshop on Cognitive Information Processing (CIP)*, pages 1–8. IEEE, 2014.

[16] A.-L. Gilet and C. Jallais. Valence, arousal and word associations. *Cognition and Emotion*, 25(4):740–746, 2011.

[17] S. Graja, P. Lopes, and G. Chanel. Impact of visual and sound orchestration on physiological arousal and tension in a horror game. *IEEE Transactions on Games*, 2020.

[18] M. Grimshaw. The audio uncanny valley: Sound, fear and the horror game. 2009.

[19] S. Houzangbe, O. Christmann, G. Gorisse, and S. Richir. Fear as a biofeedback game mechanic in virtual reality: Effects on engagement and perceived usability. In *Proceedings of the 13th International Conference on the Foundations of Digital Games*, pages 1–6, 2018.

[20] C. Klimmt, T. Hartmann, and A. Frey. Effectance and control as determinants of video game enjoyment. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society*, 10 6:845–847, 2007.

[21] C. Klimmt, D. Possler, N. May, H. Auge, L. Wanjek, and A.-L. Wolf. Effects of soundtrack music on the video game experience. *Media Psychology*, 22(5):689–713, 2019.

[22] C. Klimmt, A. Rizzo, P. Vorderer, J. Koch, and T. Fischer. Experimental evidence for suspense as determinant of video game enjoyment. *CyberPsychology & Behavior*, 12(1):29–31, 2009.

[23] S. Kumar, O. Tansley-Hancock, W. Sedley, J. S. Winston, M. F. Callaghan, M. Allen, T. E. Cope, P. E. Gander, D.-E. Bamiou, and T. D. Griffiths. The brain basis for misophonia. *Current Biology*, 27(4):527–533, 2017.

[24] J. E. LeDoux. Feelings: What are they & how does the brain make them? *Daedalus*, 144(1):96–111, 2015.

[25] P. A. Nogueira, V. Torres, R. Rodrigues, E. Oliveira, and L. E. Nacke. Vanishing scares: biofeedback modulation of affective player experiences in a procedural horror game. *Journal on Multimodal User Interfaces*, 10(1):31–62, 2016.

[26] B. Perron. Sign of a threat: The effects of warning systems in survival horror games. In *COSIGN 2004 Proceedings*, pages 132–141, 2004.

[27] R. W. Picard. Affective computing: challenges. *International Journal of Human-Computer Studies*, 59(1):55–64, 2003. Applications of Affective Computing in Human-Computer Interaction.

[28] R. P. Prager, L. Troost, S. Brüggenjürgen, D. Melhart, G. Yannakakis, and M. Preuss. An experiment on game facet combination. In *2019 IEEE Conference on Games (CoG)*, pages 1–8. IEEE, 2019.

[29] W. H. Rockett. *Devouring whirlwind: Terror and transcendence in the cinema of cruelty*, volume 21. Praeger, 1988.

[30] K. Scherer, E. Coutinho, T. Cochrane, and B. Fantini. How music creates emotion: a multifactorial process approach. 2013.

[31] S. J. Schneider. Toward an aesthetics of cinematic horror. *The horror film*, pages 131–149, 2004.

[32] M. Scirea, Y.-G. Cheong, M. J. Nelson, and B.-C. Bae. Evaluating musical foreshadowing of videogame narrative experiences. In *Proceedings of the 9th Audio Mostly: A Conference on Interaction With Sound*, pages 1–7, 2014.

[33] P. Toprac and A. Abdel-Meguid. Causing fear, suspense, and anxiety using sound design in computer games. In *Game sound technology and player interaction: Concepts and developments*, pages 176–191. IGI Global, 2011.

[34] V. Vachiratamporn, R. Legaspi, K. Moriyama, and M. Numao. Towards the design of affective survival horror games: An investigation on player affect. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, pages 576–581. IEEE, 2013.

[35] V. Vachiratamporn, K. Moriyama, K.-i. Fukui, and M. Numao. An implementation of affective adaptation in survival horror games. In *2014 IEEE conference on computational intelligence and games*, pages 1–8. IEEE, 2014.

[36] D. Watson, L. Clark, and A. Tellegen. Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of personality and social psychology*, 54 6:1063–70, 1988.

[37] M. Xu, L.-T. Chia, and J. Jin. Affective content analysis in comedy and horror videos by audio emotional event detection. In *2005 IEEE International Conference on Multimedia and Expo*, page 4. IEEE, 2005.

[38] D. Zillmann. The logic of suspense and mystery. *Responding to the screen: Reception and reaction processes*, 7:281–303, 1991.

[39] D. Zillmann. The psychology of suspense in dramatic exposition. *Suspense: Conceptualizations, theoretical analyses, and empirical explorations*, pages 199–231, 1996.