Behind the Door: Exploring Horror VR Game Interaction and its Influence on Anxiety

Marta Ferreira
NOVA School of Science and Technology, NOVA University Lisbon
Caparica, Portugal
mrd.ferreira@campus.fct.unl.pt

Ana Pinha
Lusófona University/HEI-Lab: Digital Human-Environment Interaction Lab
Lisbon, Portugal
ajppessoa@gmail.com

Micaela Fonseca
Lusófona University/HEI-Lab: Digital Human-Environment Interaction Lab
Lisbon, Portugal
LIBPhys – Laboratory of Instrumentation, Biomedical Engineering and Radiation Physics
Caparica, Portugal
micaela.fonseca@ulusofona.pt

Phil Lopes
Lusófona University/HEI-Lab: Digital Human-Environment Interaction Lab
Lisbon, Portugal
phil.lopes@ulusofona.pt

ABSTRACT

Video games are increasingly present and are currently a dominant vehicle not only for entertainment, but also for information. This position has generated a rapid growth and consequent increase in competitiveness in the game development industry, which has required the exploration of new techniques to engage users. Considering the crucial role of emotions in this domain, an increasingly appreciated approach is the integration of psychophysiology in this context. This project is thus included in the field of psychophysiological analysis, framed within the branch of video games, and consists of the study of different actions and how they can influence emotional intensity.

To understand the emotional reactions of players, two versions of a horror game in virtual reality were created, which differ in the way users interact with doors along a virtual scenario. The objective of this work is to understand if more intricate interactions can increase the intensity of human bodily reactions. Thus, to observe this potential phenomenon physiological data was acquired: electrodermal activity, electrocardiogram and respiration; in addition to self-reported questionnaires.

Results showed that the version that had a more positive global evaluation by the participants was also the one that presented a higher variation in some physiological indicators, such as heart rate and respiratory signal amplitude. Despite this, no significant differences were observed between both versions, indicating that further study is necessary for any in-depth conclusions.

FDG 2023, April 12–14, 2023, Lisbon, Portugal

ACM ISBN 978-1-4503-9855-8/23/04.
https://doi.org/10.1145/3582437.3582482

KEYWORDS

Affective Gaming, electrodermal activity, electrocardiogram, respiration, virtual reality, human–computer interaction

ACM Reference Format:
Marta Ferreira, Ana Pinha, Micaela Fonseca, and Phil Lopes. 2023. Behind the Door: Exploring Horror VR Game Interaction and its Influence on Anxiety. In *Foundations of Digital Games 2023 (FDG 2023), April 12–14, 2023, Lisbon, Portugal.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3582437.3582482

1 INTRODUCTION

Video games are interactive tools that elicit complex emotional responses in players [13]. In the process of video game production, it is sought that these games are able to create, in the audiences, experiences that affect their emotional structures in different ways, covering the entire emotional spectrum [35]. Thus, it is unsurprising how Virtual Reality (VR) has been used to study, multiple facets of human emotion and how this emotion can be leveraged for the design of better digital games.

In this sense, psychophysiological analysis is an approach that has gained popularity in game research and development, being quite promising as it provides continuous, quantitative and reliable measures of a player's affective experience, emotions and cognition [9, 32]. The level of arousal of a user results in changes in autonomic nervous system (ANS) activity, through the sympathetic nervous system (SNS) and parasympathetic nervous system (PNS) pathways. This degree can be inferred through measurements of electrodermal activity (EDA), heart rate variability (HRV), obtained from the electrocardiogram (ECG), and respiration (RSP). The combination of these signals is important to correctly determine the response of the two systems, since the heartbeat analysis provides information about the balance between the two and the EDA essentially reflects the activity of the sympathetic nervous system [19]; whereas the respiratory system is affected by the HRV [18], being associated with the activity of the vagus nerve, the main parasympathetic nerve.

While in the past, most research in the field of video games focused on understanding the effects of the act of playing in general, recently studies have shifted their focus of attention and also seek to understand how specific game characteristics can affect the experience of playing [24]. In the video game universe, virtual reality (VR) stands out for offering users a more immersive perspective and modes of interaction, which can lead to a wider range of emotional experiences, as such interactions can often simulate "real world" movements with a fairly high fidelity [25].

Thus, this project seeks to explore psychophysiological analysis in the context of games and consists in studying different actions and how they can influence the emotional intensity of the VR experience.

This work is organized as follows: Section 2 describes the motivation that led to this study. Section 3 briefly presents some related work. Section 4 indicates the materials and explains the methodologies followed. Section 5 presents and discusses the results. Finally, section 6 summarizes our conclusions.

2 MOTIVATION

Physiology and emotion have been extensively investigated, and their relationship is undeniable. Yet, the exact mapping of this relationship is still far from being known [33]. The proposed work arises in this context and aims to study the impact that specific interactions and movements have on emotional magnitude, through a virtual reality video game.

In games, interaction with elements, such as doors or windows, is frequent, so this type of interactions can be used to trigger actions that lead players to a certain experience or emotion [13]. These interactions can be implemented in a variety of ways and with different granularities. More granular interactions promote movements that are close to human actions, while interactions with lower granularity are those that are executed in a more artificial way, such as the case of pressing a button to open a door. The degree of interactive granularity can correspond to direct human perception, such as events in the game that are more or less close to the detail of the human perceptual experience, which, as a consequence, may or may not provide different reactions in users. Thus, this work poses the question whether different methods of interactive complexity of a specific action, such as opening a door, might influence the player's affective state, more precisely, if there are any noticeable changes in the player's physiological reaction.

Hence, the proposed hypothesis is as follows: "Can we observe a greater physiological impact, through HRV, EDA and RSP, when interactions have a closer proximity to that of their real-world counterpart relative to interactive versions that are more artificial (i.e. opening a door though a button prompt)?"

3 RELATED WORK

3.1 Interactions in Virtual Reality

Interfaces such as VR systems can display immersive virtual environments, where users feel fully present in that environment. The use of this technology is strongly distinguished from "traditional" video games by the level of immersion it provides and by the use of the player's own body as an interface for interaction with the virtual world, leading to unparalleled opportunities that combine player engagement with the amplification of their emotions. In this context, using the game Smash Hit/Smash Hit VR (Mediocre, 2014), a study conducted by Pallavicini et al. [24] sought to explore the differences between an immersive mode, VR, and a non-immersive mode, desktop, in terms of usability, emotional response and sense of presence. Through self-report and by observing the increase in both Skin Conductance Response (SCR) and HR, they concluded that the VR video game was associated with stronger emotional responses and players, in this modality, experienced an increase in happiness and surprise, which did not happen in the non-immersive version.

Additionally, through these systems, it is possible to induce the sense of embodiment in relation to virtual bodies. Embodiment is a complex term that has several meanings according to the field where it is applied. Its explanation is beyond the purpose of this paper, so the focus will be only on the definition according to the field of video games. In video games, embodiment describes the phenomenon experienced by players when they not only observe a virtual body, but also develop the illusion that the parts of that virtual body are parts of their own body. Studies have already proven that this illusion provides means to modulate emotional responses. For example, Gall et al. [10] conducted an experiment in which, by manipulating the temporal synchrony of certain actions, they simulated two conditions: high and low embodiment. With this study, they found that participants experienced greater arousal and more intense valence in the higher embodiment condition, concluding that the embodiment illusion intensifies emotional processing of the virtual environment.

Thus, building interactions that enhance this type of sensation can be useful when the purpose is to elicit or intensify certain emotions. In games, interactions with elements such as doors happen very frequently, so it is common for designers to take advantage of them to trigger actions that lead players to face certain events or experiences. Furthermore, in everyday life, we are used to open and cross doors to move between spaces, so it is a common interaction for users. This facilitates the process of learning the action in the virtual reality environment, which can be an advantage to achieve the desired immersion.

3.2 Horror Media

Horror entertainment is a thriving industry that can be segmented into various fields, such as movies, books, video games, theater; and paradoxical, since people willingly expose themselves to content that triggers negative emotions [7]. According to Zillmann's theory, pleasure from terror comes from the conversion of negative affect (fear, shock, etc.) into positive affect (euphoria, satisfaction) when the suspense ends and the threat is resolved [23]. For these reasons, this is a genre that attracts academic research mainly on topics related to the mechanisms and implications of horror entertainment from psychological perspectives.

Horror games are a very specific genre designed to provoke fear. These kinds of games provide an intense experience, balanced between the sensations of stress and satisfaction [13].

The main goals of terror are to frighten, shock, horrify, and even repulse. For this, several strategies, visual and auditory, are used, such as reference to the supernatural, darkness, mutilation, the

unknown, among others [23]. To build a cohesive game several parameters can be manipulated, namely the lighting, the audio, the existence or not of non-playable characters (NPCs), the game events and the generation of different levels [20].
A study conducted by Garner et al. [12], showed that the manipulation of audio parameters, in particular changes in pitch, volume and source position (3D audio), has the potential to influence the intensity of a player's fear response while playing a survival horror game. In the same context, the results of the study by Toprac et al. [28] suggested that the best sound design for causing fear are high volume sound effects synchronized with the visual element, whereas for triggering anxiety, medium volume sound effects, not timed and without a spatially defined source are preferable.
Besides sound, to intensify certain emotions during horror games, other components that can be manipulated are lights and colors. In video game development, color psychology is a common tool that allows the enhancement of certain sensations. For example, research carried out by Andersson [2], revealed that players considered the version of a horror game with white lights the scariest, indicating that the more realistic colors of that version increased the feeling of immersion, while continuous red lights were perceived as exaggerated or unnatural. In the same study, they also concluded that flickering lights are associated with greater potential for danger and jump scares.
Other elements, in horror games, that allow to drive players' emotions range from introducing enemies and limiting resources in the game world, to restricting certain actions of the character [5]. A classic scare technique is also the incorporation of sudden actions such as surprise attacks, apparitions or other disturbing events that happen without warning.
Furthermore, in the various horror entertainment media, doors also have a special connotation. They are often used to hide something secret or scary, and their appearance and context can give a sense of horror and mystery [1, 15]. If, for example, the movie or game is about a haunted house, it is typical to find an old house with wooden doors that creak and sometimes open and close by themselves. In haunted houses, doors that refuse to open or close register an anxiety that portends a kidnapping or return of something feared [13]. Since the door is seen, by the players, as something dooming, its passage is assumed as a major decision.
One of the advantages of applying this genre of video games for research lies in the fact that, due to their nature, they provide players with emotional states of high arousal, which can be inferred through psychophysiological measurements, such as heart rate (HR), electrodermal activity and respiration [9, 11]. Still, it should be mentioned that, although there are already plenty of reports on the characterization of emotions based on physiological data within the video games field, no studies have been documented that explore interactions as specific as the one we intend to evaluate (i. e. interactions with doors).

3.3 Affective Gaming and Use of Physiological Signals

With regard to the measurement and interpretation of physiological responses, considerable progress has already been made, especially in areas such as psychology, neuroscience and cognitive science, ergonomics and human computer interaction. The rise of the field of affective gaming is noticeable by the increasing number of papers published in the last 30 years [26].
In the literature, two possible types of feedback are proposed to be used for adapting games to the player's emotional state, regarding the way emotions are collected and inferred, the direct and the indirect feedback [17]. Direct feedback is obtained through biometric sensors placed on the user, such as EDA and ECG sensors; indirect feedback takes advantage of indirect characteristics, such as the way a player presses the buttons on a controller, or the way he moves his character in the game. In the big picture, because they have the advantage of providing quantitative measures, there are more contributions of studies that use direct feedback approaches, which highlights that, in recent years, there has been a targeted effort in this direction, so that the state of the art, regarding the deduction of emotions in the area of affective gaming, is through the use of physiological activity sensors.
The direct relationship between EDA and sympathetic arousal is well studied and evidenced. As a result, EDA is the most popular method for investigating human psychophysiological phenomena, and skin conductance is currently among the most common ways to measure emotion responses associated with emotions such as stress, frustration, and anxiety [33]. In a recent study Lutin et al. [21] showed that, using a support vector machine (SVM) classifier and by extracting features derived either from the decomposition of the EDA signal into tonic and phasic components or from features in the time-frequency domain, it is possible to successfully use EDA as the only predictor of stress. In this experiment they used a previously collected database, which included twenty participants who had performed three stress-enhancing tasks. Using features extracted from the EDA signal alone, they obtained a classifier with 72.50% sensitivity and 93.65% specificity, similar to the performance results of a previous study with the same dataset, which resulted in a classification model with 72.0% sensitivity and 93.4% specificity, but which additionally included the use of HR.
HRV and HR are obtained by recording heartbeats. These can be detected mainly by two methods, ECG and photoplethysmography. Although the ECG is not as practical, since it requires the application of electrodes on the player's chest, there is consensus in the literature that it provides a much more clear signal, and is therefore considered to be better than the pulse signal [33]. Investigations in the field of psychophysiology that use HRV as a parameter appear, in most cases, associated with the simultaneous acquisition of other signals, such as EEG [29] or EDA [9]. The experimental results of these studies indicated that the acquired physiological measures correlated with gaming experiences, even in relatively simple measurement scenarios, such as the one present in a study by Drachen et al. [9]. In this, participants were asked to play three games, Prey (Arkane Studios, 2017), Doom 3 (id Software, 2004), and Bioshock (Irrational Games, 2007), and, throughout that period, to answer game experience questionnaires (iGEQ), with questions about immersion, flow, competence, tension, challenge, and positive and negative affect; the results indicated that there was correlation between HR and the seven dimensions of the questionnaire, whereas EDA only revealed correlation with the negative affect dimension.
The research presented by Vachiratamporn et al. [29] also showed the promising potential of using only ECG signals in the automatic

prediction of states like anxiety and suspense, through a classifier based on decision trees, in the context of a horror game. Here, the obtained accuracy was superior, compared to the use of only EEG for classification, or even both at the same time. Additionally, this study found that players were more likely to feel fear of a scary event when in a state of suspense, compared to when in a neutral state, which is a result that may be useful in the development of horror games.
Respiratory rate has also been used, essentially in combination with other physiological indicators, to build predictive models of emotions [14, 32]. In addition, several studies show its contribution as a biofeedback mechanism in games, for example, for stress therapy. Zafar et. al [34], developed a version of the game Pac-Man (Namco, 1980), which adapts its gameplay in order to encourage players to control respiratory rate, promoting arousal regulation in stressful environments. One advantage of using this type of physiological signal, in the context of games with biofeedback, is that, unlike other parameters such as HRV, breathing is an indicator that the player can control.

Thus, the study of specific interactions in games, like doors, through physiological measures, ECG, EDA and RSP, by combining the various concepts mentioned, such as embodiment and immersion, and making use of techniques to trigger emotional states provided by terror environments, may be an approach with important contributions, particularly at the level of the development of games in virtual reality, since its results may be useful in understanding which game modes provide better experiences to players.

4 MATERIALS AND METHODS

4.1 Game Design

The game is inspired by psychological horror and survival horror games [13]. The action takes place within a closed environment, consisting of a large room and an office, connected by two hallways, Figure 1. The purpose of the game is for the player, confined in a threatening environment, to try to escape from it. To do so, the player will have to navigate through the scenario over eight turns/sections. Throughout each turn a series of events occur, and to complete the experience, the player only needs to go through all the predetermined cycles in which these events are included.
The game states, across conditions and participants, are always updated equally, triggering the same predetermined events, to ensure a linear experience for everyone. Players can walk and observe the virtual environment, navigating it through a first person perspective in VR.

4.1.1 Game Sections.

For the game, the typical horror game design was followed, with a slow and gradual increase in tension over time (see Figure 2), which has already been proven to be adequate to induce anxiety/fear [13]. Thus, a set of loops were designed, with the goal of eliciting several stages of tension, essentially through changes in lighting, sounds, and appearance of "threatening" regions or figures. The duration of each section, and consequently of the entire game, varies and depends on the player's performance, and no time limit is imposed for the experience. Some parts of the game are represented in Figure

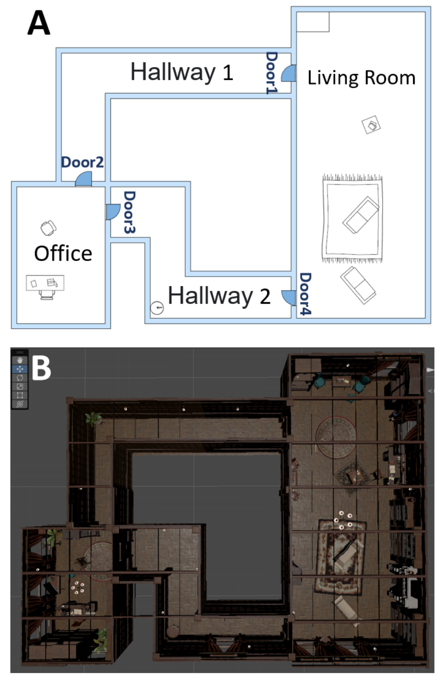

Figure 1: A: Blueprint of the game scenario, with the respective identification of doors and rooms. B: Real perspective, through Unity environment, of the developed scenario.

3. A full demonstration of the game and the two interaction systems can be accessed here[1].

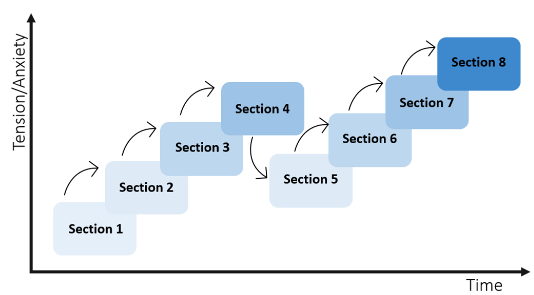

Figure 2: Diagram illustrating the progression of anxiety designed for the game.

In general, the sections can be summarized as follows:

Section 1 First turn of the game, where the participant is confronted with the scenario and the environment. There is no specific event, and the lighting level is high. It is thus intended to provoke a low degree of anxiety.

Section 2 A background sound of light rain is introduced and the brightness of the entire map decreases slightly. Throughout the rooms, small events, such as sounds and flashing

[1] https://tinyurl.com/fdg23-45

lights, occur. The goal is to start creating some tension in the player.

Section 3 In this section, it is intended to create a sense of suspense and hesitation. For this, the narrative is established only by changing lights and sound and the appearance of shadows. It is aimed to generate the perception that the character is being watched by some entity.

Section 4 A significant increase in the tension level is expected. The player is finally confronted with more concrete events, beyond just changes in lighting and sounds, such as a painting falling in the living room and a rocking chair moving in the office. In this section, the player has to interact with a key, to open a door that is locked.

Section 5 Everything returns to the beginning, as in section 1, except for objects that have been moved throughout the previous sections. A sense of return to a more calm state is expected.

Section 6 A large storm begins and the lights start to fail. It is intended to bring the player back into a growing state of anxiety, with the introduction of bloody objects and human sounds.

Section 7 At the beginning of the seventh section, there is a general blackout of the lights and the player is required to walk across the map with the aid of a flashlight. In terms of scenery there are no changes, creating suspense and hesitation in the player.

Section 8 In the last section, a maximum degree of anxiety/fear is intended, compared to the previous sections. There is the appearance of two creatures. Finally, when the player goes through the last door, door 4, the experience ends.

4.1.2 Game Versions.
Considering the goal of the study is to observe different interactions, two alternating versions were built:

- Game A: The more granular interaction, where the participant has to physically place their virtual hand near the door handle and physical turn it;
- Game B: The less granular interaction, where the participant only needs to be adjacent to the door and press the prompted button.

4.2 Data Collection

4.2.1 Physiological Data.
For this work, the physiological data collected were ECG, EDA and RSP. The monitoring and recording of these signals was done using the BioPLUX system from Plux - Wireless Biosignals, S.a. This device is controlled via Bluetooth connection. These data were acquired, in the three channels, at a sampling rate of 1000 Hz, with a resolution of the analog digital converter (ADC) of 16 bits, characteristic of the equipment.

For the integration of the physiological data synchronously with the game events, the Lab Stream Layer (LSL) [16] software was used, which is a protocol with libraries and tools that allow the synchronization and transfer of data and biosignals in real time. Thus, all data of the participant's physiological activity is transmitted to the computer and saved using LabRecorder, LSL's standard recording program, in a single file in XDF format.

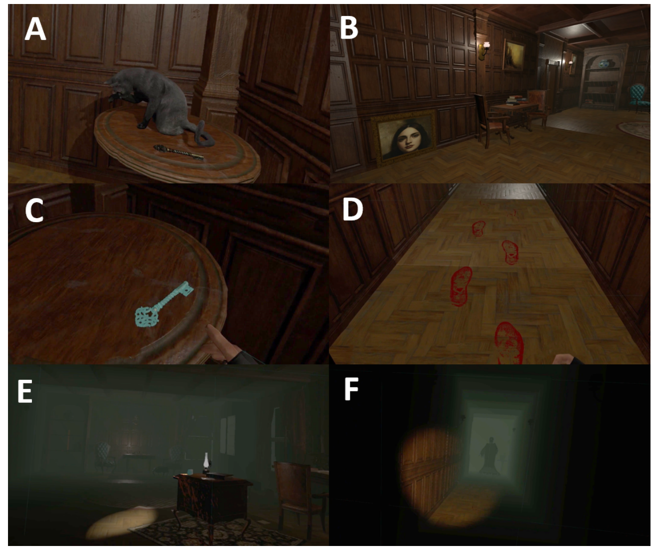

Figure 3: Examples of different moments in the game. A - Appearance of a cat (Section 2), B- Frame drop (Section 3), C - Key needed to open a locked door (Section 4), D- Bloody footprints (Section 6), E - General blackout of lights and flashlight (Section 7), F - Ghost (Section 8).

4.2.2 Game Markers.
In the same XDF file where the data of the signals are stored, the event markers of the game are also registered. This registration was done through a timestamp and an event identifier. The instants associated with the markers selected as of interest for this work were: beginning and end of the baseline, beginning of the tutorial, start of the game and opening and passing each door.

4.2.3 Postexperimental Questionnaire.
At the end of the experiment, a questionnaire was presented to the participants. This questionnaire is divided into ten questions. These resulted from a review and adaptation of several usability and gameplay questionnaires, namely the System Usability Score (SUS) [6], SlaterUsoh-Steed Presence Questionnaire [27] and Presence Questionnaire [31], from which we selected questions that would be adequate to the present project. In all questions, where the answer is given according to scales, except for question six, the seven-point Likert scale was chosen, from "1 - Strongly Disagree" to "7 - Strongly Agree", since it presents a higher degree of reliability [8]. Question six, which aims to evaluate the self-perception of well-being and affectivity, was based on the Positive and Negative Affect Schedule (PANAS) scale [30]. A Likert scale ranging from "1 - Not at all or very Slightly" to "5 - Extremely" was used for this question.

4.2.4 Experimental Protocol.
Each volunteer begins by reading and accepting an informed consent document. Once this step is completed, the participants are asked to fill out the demographic questionnaire. Next, the electrodes are positioned in the appropriate locations and the VR headset is placed and adjusted. The participant is positioned in the room and the tasks begin. These can be segmented into three phases: the baseline phase, which lasts one minute and where the participant

is instructed not to move or interact with the system; the tutorial phase, where a short tutorial is presented to familiarize the participant with the virtual reality environment and the game mechanics; and the game phase, in which the participant performs the activity itself. At the end, the sensors and other equipment are removed and then the game experience questionnaire is filled out. Once the questionnaire is completed, the session is considered over.
Among different participants, the choice of which activity to perform, i.e. game A or game B, was alternated, so that half of the participants performed the experiment with the button application and the other half with the movement of turning the handles and pushing the doors.
The questionnaires and methodology for this study were approved by the Ethics and Deontology Committee for Scientific Research of Lusófona University.

4.2.5 Participants.
The initial sample was composed of thirty individuals, twenty male subjects and ten female subjects, however, and after the preprocessing step, described in section 4.3, only seventeen of the participants' data were evaluated (12 male and 5 female). More than half of the participants reported that they played video games often or almost always (52.94%), the rest being divided into casual gamers (29.41%) and individuals who rarely or never play (17.65%). Regarding VR experience, only 17.65% responded that they played casually or often, and of the surplus, 41.175% stated they had never tried it or had done so a long time ago, and the other 41.175% had their first VR experience not long ago. Finally, the majority stated they never or rarely play horror games (76.47%).

4.3 Data Pre-Processing

The signals were visualized one by one, in order to detect possible anomalies during the acquisitions. At this stage, thirteen of the thirty initial signals were discarded. This elimination was due to the fact that the EDA signal had saturated in ten participants and in the remaining three there was no information corresponding to the complete experiment, due to early interruptions by cybersickness and errors in the game.

4.4 Data Processing

The signal processing was done in Python using Neurokit2. This is an open-source package that provides a comprehensive set of methods for processing various physiological signals, e.g. ECG, EDA, EMG and RSP [22]. The approach followed is outlined in Figure 4.
After the signal acquisition, raw ECG, EDA and RSP signals should be processed. The raw signals are often corrupted by the external noise such as powerline interference, baseline drift and other biological artifacts like breathing and muscle noise. Various approaches have been proposed for signal processing. In this study, we selected the frequency filtering, which is efficient and easy to implement. Before applying the filters, a z-score standardization of the data was performed.
For the ECG, a high-pass butterworth filter of order 5 and cutoff frequency 0.5 Hz was applied to eliminate slow fluctuations and DC offsets, followed by a moving average filter to remove 50 Hz power supply noise, which can be induced by poorly filtered power supplies.
For the EDA, a butterworth low-pass filter of order 4 and cutoff frequency 3 Hz, suitable for preserving the low frequencies characteristic of these signals, was applied.
Finally, for the RSP signal, we applied a butterworth bandpass of order 2 and cutoff frequencies 0.05 Hz and 0.5 Hz, which removes slow fluctuations from the baseline and high-frequency noise and preserves breathing frequencies between 3 and 30 breaths per minute.
After filtering the signals, they were segmented in order to study the interactions with the doors. For this purpose, ten second windows were created, using the door opening marker as reference. These windows comprise two seconds immediately before the marker and eight seconds after it. The duration chosen for the window interval resulted from a consideration of several aspects, including the average time of door opening by the participants, the possible overlap of windows of two consecutive doors, and the minimum acceptable time to obtain certain indicators of physiological signals.
Finally, the features of interest were extracted, namely the average HR, number and amplitude of the SCR peaks, average respiratory rate and average amplitude of the RSP signal.

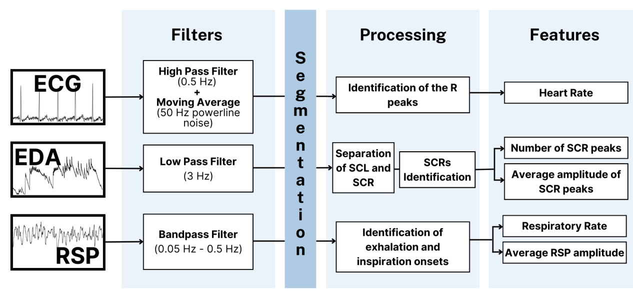

Figure 4: Processing steps for each of the biosignals.

5 RESULTS

5.1 Game Design

To verify the degree of anxiety induced in each section of the game, the physiological signals collected were divided in windows, corresponding to each lap of the game. Then, several physiological indicators were extracted per section, namely the mean HR, Root Mean Square of Successive N-N Interval Difference (RMSSD index), powers in the low and high frequency range of HRV signal, number, and mean amplitude of SCR peaks, mean respiratory rate and mean respiratory amplitude.
In general, the results obtained allowed us to understand that, although in many cases there were no statistically significant changes between consecutive sections, it is possible, according to the indicators studied, to observe certain response patterns to the different moments of the experience. We observed a slight increase in HR throughout the game, reduction of the RMSSD indicator, increased power in the low frequency range of the HRV signal and a decrease in high frequencies and sharp amplitudes of the phasic responses at certain moments. Thus, a degree of tension and anxiety triggered in the participants was confirmed. This could be used as a basis

for future modifications, with the intent of optimizing the game to meet any needs that further studies may require.

5.2 Questionnaire

For the present study, a reduced version of the PANAS questionnaire was used, composed of five items to assess the positive affect (PA) - determined, active, enthusiastic, inspired and interested - and another five for the negative affect (NA) - nervous, afraid, scared, guilty and upset.

Positive Affect: After verification of normality, when comparing the means of the PA between the two versions (Condition A: M = 15.30, SD = 3.61; Condition B: M = 18.57, SD = 2.61), using t-test, there was no statistical significance (p-value = 0.073). However, in Figure 5A, we can see that Condition B has the minimum and median values higher than Condition A, with the points clustered in higher scores, indicating that the PA in B may be higher. Thus, in this condition, participants experienced a greater sense of pleasure and subjective well-being. However, this is a conclusion that may be conditional due to the small sample size, and with a larger number of participants, it could be proven whether this difference would actually be statistically significant.

Negative Affect: The normality test showed no evidence against it. When comparing the mean NA scores (Condition A: M = 8.90, SD = 2.51; Condition B: M = 10.57, SD = 3.66), there was no statistical difference (p-value = 0.310), and by inspecting Figure 5B, we can see that the data distributions between the two conditions are not distinct. In this case, we also observe that in both conditions, the NA scores are not as high as the PA scores analyzed earlier, which reflects that participants did not evaluate their experience as a displeasing activity.

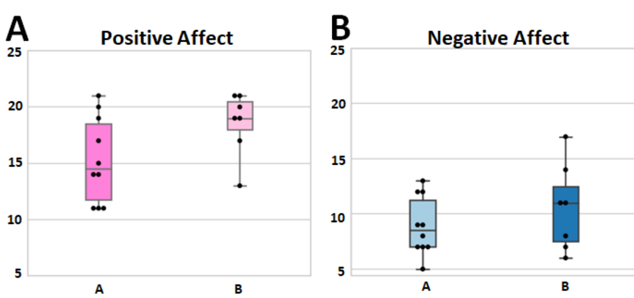

Figure 5: A: Boxplots of the PA for conditions A and B. B: Boxplots of the NA for conditions A and B.

The questionnaire also included questions about the gaming experience, namely:

1. *I think that having sensors placed on my body affected my experience and may have made me more nervous.*
2. *I felt comfortable during the games and did not experience any feelings of nausea.*
3. *I lost awareness of my surroundings while playing the game.*
4. *I performed the game activities with high ease.*
5. *I would like to use the game system frequently.*
6. *Interacting with the game world didn't feel as real to me as it would in the "real world".*
7. *I got so involved in the task that I lost track of time.*
8. *The mechanics and interactions felt quite natural to me.*

In only one of the items, question 8., was there a statistically significant difference between the groups (p-value = 0.025 < 0.05). Thus, the participants who performed the tasks of condition B had the perception that the interactions of this version were more natural, as opposed to what was initially hypothesized. This may be related to limitations in the development of the condition A interaction. Figure 6 shows the distributions of the participants' responses to the questionnaire. We can see that, although the difference was not significant, there was a greater tendency for participants in condition B to feel sick. Also in condition B there was a greater loss of perception of the environment and, therefore, greater immersion in the game. With regard to the ease of execution of the tasks, although in condition A we noticed some participants with more limitations, the region of greater condensation of data is similar in both conditions, and the participants agreed that the activities were performed with ease. Regarding question 5., it is possible to observe a positive deviation in the answers of condition B, which suggests that the individuals who performed the activity in this version were more pleased than the others. On question 6., there is a great dispersion in the answers, with no distinction between the two conditions. Finally, in 7., for both condition A and condition B, there is a greater density of answers in values higher than five, which indicates that most participants had the sensation of losing track of time while playing, which is another indicator of immersion, and this, again, is higher in condition B.

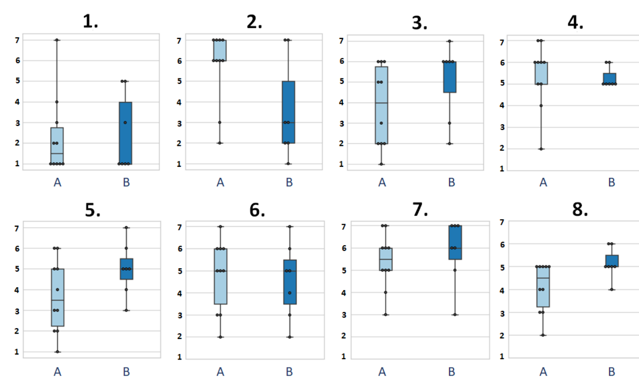

Figure 6: Boxplots of the questionnaire responses (1, 2, 3, 4, 5, 6, 7, and 8) for each of the conditions.

5.3 Signals

5.3.1 ECG.

For the comparison of HR between conditions, we used the variable ΔHR, which refers to the difference of the mean HR in the window and the mean baseline HR. The results are shown in Figure 7.

It is possible to see, through Figure 7A, that, in general, the HR

variations are higher and more positive in Condition B, suggesting that this interaction stimulates more intense changes in the emotional state of the individuals, because there is a greater sympathetic activation. After checking the normality of the data, when comparing the means of ΔHR between the two versions (Condition A: M = 7.21, SD = 9.03; Condition B: M = 12.57, SD = 10.65), using t-test, statistical significance was verified (p-value = 1.00e-09 « 0.05), which corroborates the observed difference.
Regarding the variation across sections, Figure 7B, there are not many differences in condition A, whereas in condition B there is more fluctuation in HR throughout the game. When comparing the interaction between the two conditions (A vs B), there were statistical differences in sections 2, 3, 4, 6 and 8 (p-value < 0.05).

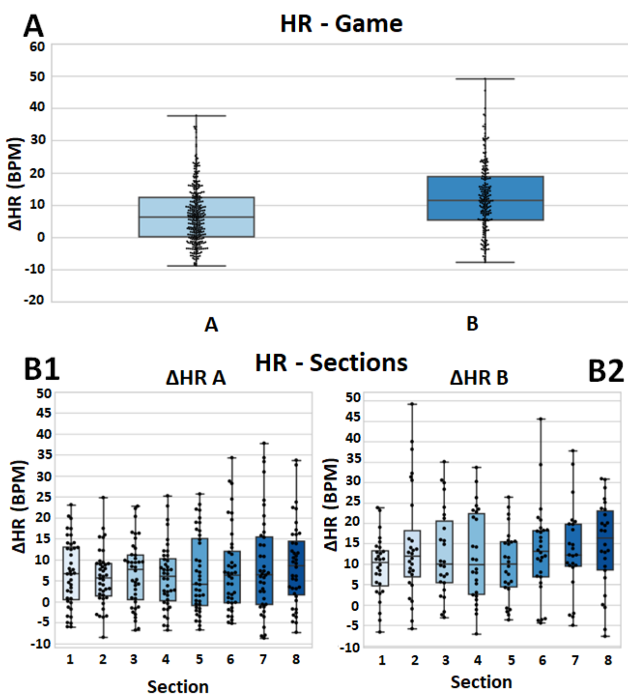

Figure 7: A: Boxplots referring to the ΔHR of the total number of doors of a complete game, for conditions A and B, B: Boxplots referring to the ΔHR during the interaction with the doors, along the sections, B1: Condition A, B2: Condition B

5.3.2 EDA.
The analysis of the EDA signal was based on the identification of the SCR peaks and their amplitudes. With regard to the number of peaks per window, an almost identical overall result is observed between the two conditions (Condition A: M = 0.90, SD = 0.76; Condition B: M = 0.89, SD = 0.68), Figure 8A, which was verified through the t-test, which was not significant, p-value = 0.969 > 0.05. By inspection of the collected data, it was found that in most situations, there is one SCR peak per window, thus per door interaction.
Looking at the results by section, Figure 8B, again there are no evident differences between the two conditions. Still, it is possible to see that in condition A, there are three sections that present, on average, more than one peak, unlike condition B, where this only occurs in the first lap. This may be related to the way each interaction is executed. In section 1, participants in both conditions were still in a learning stage, which may be a stress-enhancing task, resulting in activation of the sympathetic component of the ANS and leading to the emergence of more phasic peaks. After a short period of time, most of the individuals assigned to condition B had already mastered the task, while in the case of condition A, several participants showed a greater delay in this learning process, because it is a more complex interaction. This may justify the existence of more peaks, not only in the first section, but also by the third round. Finally, throughout the acquisitions, it was noticed that the moment of the game in which the lights turned off (sections 7 and 8) was more frustrating and stressful during experiment A, since the position of the flashlight was not optimized for the players' height, so they had some difficulty in visualizing the door handle and, consequently, in performing the interaction with ease, which could be a possible justification for the increase observed in Figure 8B1, section 8. Still, although slight variations were noted, the results of the statistical tests did not exhibit significant differences between the two conditions.

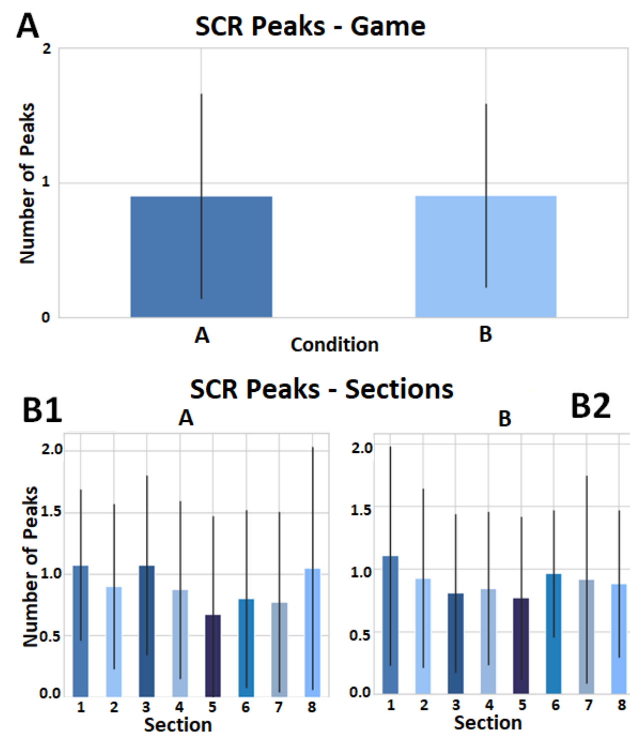

Figure 8: A: Mean value of the number of phasic peaks per window and their standard deviation considering all doors crossed in a full game, for conditions A and B, B: Mean value of the number of phasic peaks per window and respective standard deviation over the sections of the game, B1: Condition A, B2: Condition B.

As for the amplitude of identified SCR responses representative of the interaction with the doors, overall no considerable differences were observed, Figure 9A, which was verified by the statistical test, p-value = 0.271 > 0.05 (Condition A: M = 0.45, SD = 0.56; Condition B: M = 0.49, SD = 0.50). Across the sections of the game, Figure 9B, between equivalent sections of the two conditions, again, no concrete criteria for distinction are evident. In statistical tests, only one of the sections revealed a different distribution between the two conditions (p-value < 0.05), namely section 6, where higher phasic peak amplitudes were observed in condition B.

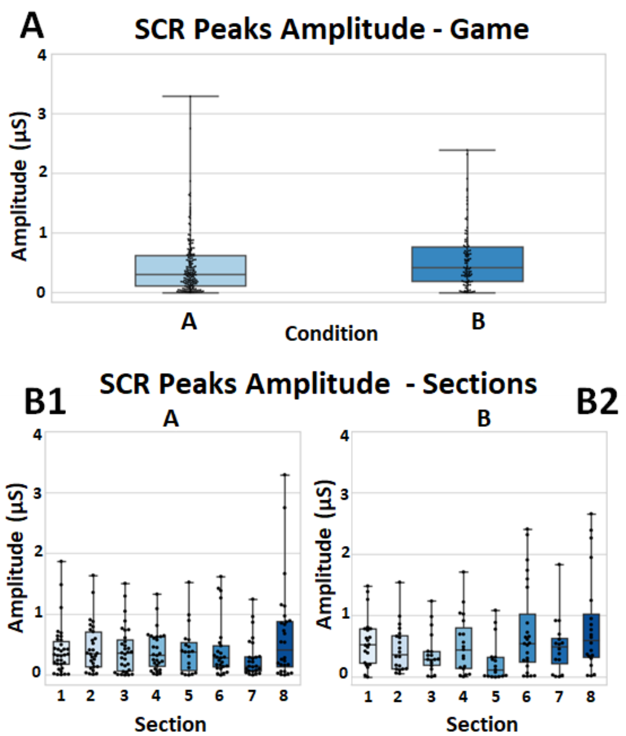

Figure 9: A: Boxplots concerning the SCR amplitude of the peaks per window of the totality of the doors of a complete game, for conditions A and B, B: Boxplots concerning the SCR amplitude during the interaction with the doors, throughout the sections, B1: Condition A, B2: Condition B.

5.3.3 *RSP.*

To evaluate the respiratory signal during the interaction under study, respiratory rate and signal amplitude were used as physiological indicators.

Because condition B participants had considerably high baseline respiratory frequencies, the analysis was performed between consecutive sections using ΔFR, obtained through the difference of the respiratory frequency of one section with the previous section. In Figure 10A, no evident differences can be seen between the two conditions. When comparing the means of ΔFR (Condition A: M = 0.25, SD = 4.26; Condition B: M = -0.34, SD = 4.03), using t-test, there was no statistical significance (p-value = 0.103 > 0.05).

Analyzing the game sections, individually, Figure 10B, again it is not possible to identify a criteria to distinguish between the two interaction modes, both showing identical value distributions.

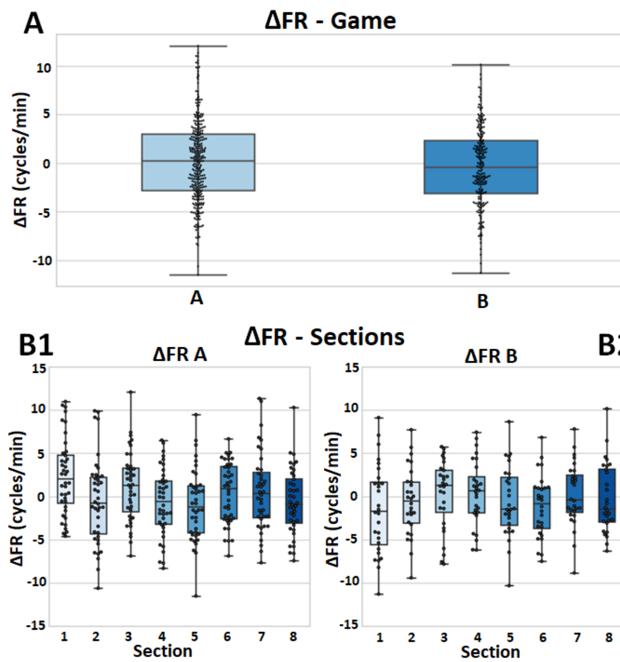

Figure 10: A: Boxplots regarding the ΔFR of all the doors of a complete game, for conditions A and B. B: Boxplots regarding the ΔFR during the interaction with the doors, throughout the sections, B1: Condition A, B2: Condition B.

Looking at the amplitude of the signal, whose results can be seen in Figure 11, we see higher values of this parameter in Condition B. Considering the contribution of all windows of a complete game (Condition A: M = 0.57, SD =1.13; Condition B: M = 2.13, SD = 1.63), the statistical test revealed significant differences between the two interactions (p-value = 8.962e-34 « 0.05), which corroborates this observation.

Similar to the differences identified, when weighting all windows, the analysis by sections revealed a tendency for the variations in respiratory signal amplitudes to be greater in condition B, which was also verified using t-tests.

Several studies indicate that rapid and shallow breathing is found during the performance of a stressful mental task and appears to be associated with anticipation, a pattern that may in turn be linked to a range of affective states such as concentration, resolution, tension, anxiety, fear, and ultimately panic. Whereas, an increase in respiratory rate combined with an increase in breathing depth seems to be associated with states of arousal/activation, often related to emotions of fear, anger, and joy [4]. Thus, the observed changes may indicate that the stress/tension felt by participants in condition A was higher, whereas condition B provided a greater degree of activation/excitement in the participants.

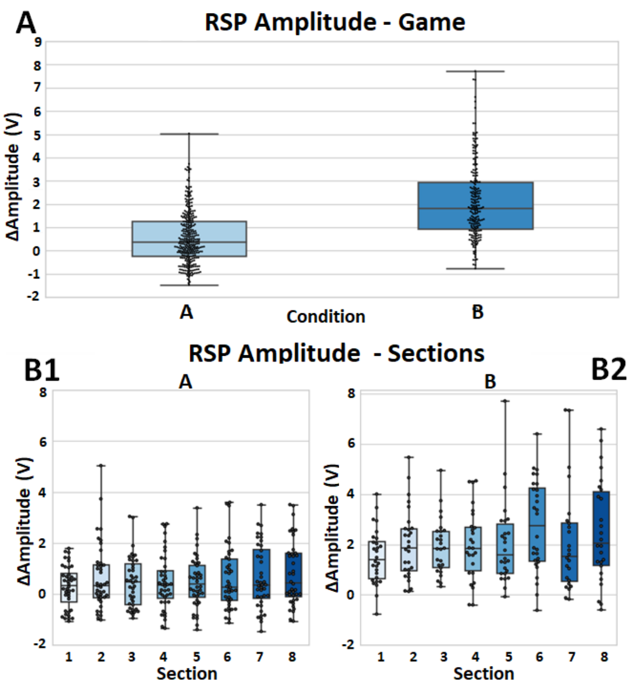

Figure 11: A: Boxplots regarding the ΔAmplitude RSP of the totality of the doors of a complete game, for conditions A and B, B: Boxplots referring to the ΔAmplitude RSP during the interaction with the doors, throughout the sections, B1: Condition A, B2: Condition B.

6 CONCLUSION

The main purpose of this study was to evaluate two different interactions in a virtual reality horror game, according to their impact on the players' stress/anxiety intensity, assessed by analyzing the behavior of several physiological indicators, obtained through EDA, ECG and RSP signals. With this, it was intended to understand which of the interactions provided a better game experience, this being understood as the one in which greater changes were observed in physiological parameters. As a hypothesis, it was initially proposed that the version that would respect these criteria would be the one in which the most realistic and closest to human actions interactions were implemented. For this study, mean HR, number and amplitude of SCR peaks, mean respiratory rate, and mean respiratory amplitude were analyzed.

The results obtained allowed us to conclude about the existence of divergences between the physiological responses to the two conditions, which may indicate different emotional reactions when playing one version or the other. More specifically, there was a greater variation in heart rate and a greater fluctuation in respiratory signal amplitude in the condition where players only had to press the button to open the doors, which indicates a more intense response to stress and a higher level of excitement/activation in the participants who played this version. Through the questionnaires, we also perceived a preference of the participants for the button interaction.

Thus, we conclude that, in opposition to the hypothesis put forward, the interaction that seemed to provide the best game experience was not the one that most closely matches the natural human movement of opening a door.

In view of the results achieved, in order to reach more robust conclusions, it would be advantageous, as a first approach, to repeat the tests and the respective analysis, similar to the one performed, for a larger number of participants. Some improvements should also be made in the interaction with the door, in order to make it more fluid, promoting a greater sense of embodiment. Additionally, changes can be introduced in the markers. On the one hand, it would be more rigorous if the marker for the opening of the door would be at a key moment of the interaction, for example, when the player starts the rotation movement of the handle, and, on the other hand, it could also be interesting the introduction of markers to verify multiple successive attempts to open the same door. Regarding the experimental protocol, for this study, we chose to alternate the different conditions among different participants, even though favorable results could be obtained for the intended study if each individual was exposed to the two types of interaction. A physiological indicator to deduce the player's immersion and sense of embodiment throughout the experience could also be a useful addition to this evaluation. Thus, in future work, EEG could be included, a signal that has already proven useful in measuring presence in virtual reality environments [3].

Finally, it should be mentioned that this type of analysis, based on games, has some limitations, as each participant ends up experiencing the events in a different and uncontrollable way, which can result in a considerable variability in the data obtained, as we have seen throughout the analysis of the different parameters. Additionally, the fact that some participants reported feeling nauseous, especially in condition B, may have had an influence on heart and respiratory rates, heart rate variability, and sweating. However, due to individual differences in autonomic regulation and variations caused by the game experience and events itself, it is difficult to predict the contribution in physiological responses directly related to the cybersickness phenomenon.

ACKNOWLEDGMENTS

This work is supported by Fundação para a Ciência e Tecnologia (FCT), under HEI-Lab R&D Unit (UIDB/05380/2020) and Project PlayersAll: media agency and empowerment (EXPL/COM-OUT/088 2/2021).

REFERENCES

[1] [n.d.]. Door of doom. https://tvtropes.org/pmwiki/pmwiki.php/Main/DoorOfDoom Accessed: 2022-09-1.
[2] Irma Andersson. 2022. *The effect of red-light sources on the player experience of a horror game level.* Bachelor's Thesis. Department of Game Design, Faculty of Arts, Uppsala University.
[3] Mohamed Athif, B. Rathnayake, S. Nagahapitiya, S. Samarasinghe, P. Samaratunga, R. Peiris, and Anjula De Silva. 2020. Using Biosignals for Objective Measurement of Presence in Virtual Reality Environments. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC).* IEEE, 3035–3039. https://doi.org/10.1109/EMBC44109.2020.9176022
[4] Frans A. Boiten, Nico H. Frijda, and Cornelis J.E. Wientjes. 1994. Emotions and respiratory patterns: review and critical analysis. *International Journal of Psychophysiology* 17 (7 1994), 103–128. Issue 2. https://doi.org/10.1016/0167-8760(94)90027-2
[5] Casper S. Boonen and Daniel Mieritz. 2018. Paralysing Fear: Player Agency Parameters in Horror Games. In *DiGRA Nordic '18: Proceedings of 2018 International*

DiGRA Nordic Conference. http://www.digra.org/wp-content/uploads/digital-library/DiGRA_Nordic_2018_paper_19.pdf
[6] John Brooke. 1995. SUS: A quick and dirty usability scale. *Usability Eval. Ind.* 189 (11 1995).
[7] Mathias Clasen, Jens Kjeldgaard-Christiansen, and John A. Johnson. 2020. Horror, personality, and threat simulation: A survey on the psychology of scary media. *Evolutionary Behavioral Sciences* 14 (7 2020), 213–230. Issue 3. https://doi.org/10.1037/ebs0000152
[8] Marlon Dalmoro and Kelmara Mendes Vieira. 2013. Dilemas na construção de escalas Tipo Likert: o número de itens e a disposição influenciam nos resultados? *Revista Gestão Organizacional* 6 (2013), 161–174. Issue 3.
[9] Anders Drachen, Lennart E. Nacke, Georgios Yannakakis, and Anja Lee Pedersen. 2010. Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. In *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games - Sandbox '10* (New York, New York, USA). Association for Computing Machinery, 49–54. https://doi.org/10.1145/1836135.1836143
[10] Dominik Gall, Daniel Roth, Jan-Philipp Stauffert, Julian Zarges, and Marc Erich Latoschik. 2021. Embodiment in Virtual Reality Intensifies Emotional Responses to Virtual Stimuli. *Frontiers in Psychology* 12 (9 2021). https://doi.org/10.3389/fpsyg.2021.674179
[11] Eva Ganglbauer, Johann Schrammel, Stephanie Schwarz, and Manfred Tscheligi. 2009. Applying Psychophysiological Methods for Measuring User Experience: Possibilities, Challenges and Feasibility. In *Workshop on user experience evaluation methods in product development*. Uppsala, Sweden.
[12] Tom Garner, Mark Grimshaw, and Debbie Abdel Nabi. 2010. 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* (Piteå, Sweden) *(AM '10)*. Association for Computing Machinery, New York, NY, USA, Article 10, 9 pages. https://doi.org/10.1145/1859799.1859809
[13] Sarra Graja, Phil Lopes, and Guillaume Chanel. 2021. Impact of Visual and Sound Orchestration on Physiological Arousal and Tension in a Horror Game. *IEEE Transactions on Games* 13 (9 2021), 287–299. Issue 3. https://doi.org/10.1109/TG.2020.3006053
[14] Rabab A. Hameed, Mohannad K. Sabir, Mohammed A. Fadhel, Omran Al-Shamma, and Laith Alzubaidi. 2019. Human emotion classification based on respiration signal. In *Proceedings of the International Conference on Information and Communication Technology* (Baghdad, Iraq) *(ICICT '19)*. Association for Computing Machinery, New York, NY, USA, 239–245. https://doi.org/10.1145/3321289.3321315
[15] Johan Hjern. 2015. *The Horror of Doors*. Bachelor's Thesis. Uppsala University, Disciplinary Domain of Humanities and Social Sciences, Faculty of Arts, Department of Game Design.
[16] C. Kothe, D. Medine, C. Boulay, and Grivich M.and Stenner T. 2012. (2012) Lab Streaming Layer (LSL)-A software framework for synchronizing a large array of data collection and stimulation devices. The Swartz Center for Computational Neuroscience Software. Available online: https://code. google.com/archive/p/labstreaminglayer/(accessed on 15 April 2022). https://code.google.com/archive/p/labstreaminglayer/
[17] Raúl Lara-Cabrera and David Camacho. 2019. A taxonomy and state of the art revision on affective games. *Future Generation Computer Systems* 92 (3 2019), 516–525. https://doi.org/10.1016/j.future.2017.12.056
[18] Paul Lehrer. 2013. How Does Heart Rate Variability Biofeedback Work? Resonance, the Baroreflex, and Other Mechanisms. *Biofeedback* 41 (3 2013), 26–31. Issue 1. https://doi.org/10.5298/1081-5937-41.1.02
[19] Rodrigo Lima, Daniel Osório, and Hugo Gamboa. 2019. Heart rate variability and electrodermal activity in mental stress aloud: Predicting the outcome. In *BIOSIGNALS 2019 - 12th International Conference on Bio-Inspired Systems and Signal Processing, Proceedings; Part of 12th International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2019*. SciTePress, 42–51. https://doi.org/10.5220/0007355200420051
[20] Phil Lopes and Maria Fonseca. 2021. Towards Effectively Adapting Games: What needs to be Conquered to Achieve Adaptation. In *Eurosis Game-On 2021*.
[21] Erika Lutin, Ryuga Hashimoto, Walter De Raedt, and Chris Van Hoof. 2021. Feature Extraction for Stress Detection in Electrodermal Activity. In *Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies*. SCITEPRESS - Science and Technology Publications, 177–185. https://doi.org/10.5220/0010244601770185
[22] Dominique Makowski, Tam Pham, Zen J. Lau, Jan C. Brammer, François Lespinasse, Hung Pham, Christopher Schölzel, and S. H. Annabel Chen. 2021. NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behavior Research Methods* 53 (8 2021), 1689–1696. Issue 4. https://doi.org/10.3758/s13428-020-01516-y
[23] G. Neil Martin. 2019. (Why) Do You Like Scary Movies? A Review of the Empirical Research on Psychological Responses to Horror Films. *Frontiers in Psychology* 10 (10 2019). https://doi.org/10.3389/fpsyg.2019.02298
[24] Federica Pallavicini, Alessandro Pepe, and Maria Eleonora Minissi. 2019. Gaming in Virtual Reality: What Changes in Terms of Usability, Emotional Response and Sense of Presence Compared to Non-Immersive Video Games? *Simulation & Gaming* 50 (4 2019), 136–159. Issue 2. https://doi.org/10.1177/1046878119831420
[25] Giuseppe Riva, Fabrizia Mantovani, Claret Samantha Capideville, Alessandra Preziosa, Francesca Morganti, Daniela Villani, Andrea Gaggioli, Cristina Botella, and Mariano Alcañiz. 2007. Affective interactions using virtual reality: The link between presence and emotions. *Cyberpsychology and Behavior* 10 (2 2007), 45–56. Issue 1. https://doi.org/10.1089/cpb.2006.9993
[26] Raquel Robinson, Katelyn Wiley, Amir Rezaeivahdati, Madison Klarkowski, and Regan L. Mandryk. 2020. "Let's Get Physiological, Physiological!": A Systematic Review of Affective Gaming. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (New York, NY, USA). ACM, 132–147. https://doi.org/10.1145/3410404.3414227
[27] Mel Slater, Martin Usoh, and Anthony Steed. 1994. Depth of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments* 3 (1 1994), 130–144. Issue 2. https://doi.org/10.1162/pres.1994.3.2.130
[28] Paul Toprac and Ahmed Abdel-Meguid. 2010. *Causing Fear, Suspense, and Anxiety Using Sound Design in Computer Games*. IGI Global, 176–191. https://doi.org/10.4018/978-1-61692-828-5.ch009
[29] Vanus Vachiratamporn, Roberto Legaspi, Koichi Moriyama, and Masayuki Numao. 2013. Towards the Design of Affective Survival Horror Games: An Investigation on Player Affect. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*. IEEE, 576–581. https://doi.org/10.1109/ACII.2013.101
[30] David Watson, Lee Anna Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology* 54 (1988), 1063–1070. Issue 6. https://doi.org/10.1037/0022-3514.54.6.1063
[31] Bob G. Witmer and Michael J. Singer. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments* 7 (6 1998), 225–240. Issue 3. https://doi.org/10.1162/105474698565686
[32] Wenlu Yang, Maria Rifqi, Christophe Marsala, and Andrea Pinna. 2018. Physiological-Based Emotion Detection and Recognition in a Video Game Context. In *2018 International Joint Conference on Neural Networks (IJCNN)*. 1–8. https://doi.org/10.1109/IJCNN.2018.8489125
[33] Georgios Yannakakis, Hector Martinez, and Maurizio Garbarino. 2016. *Psychophysiology in Games*. Vol. 4. Springer International Publishing, Cham, 119–137. https://doi.org/10.1007/978-3-319-41316-7_7
[34] M. Abdullah Zafar, Beena Ahmed, and Ricardo Gutierrez-Osuna. 2017. Playing with and without Biofeedback. In *2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH)*. IEEE, 1–7. https://doi.org/10.1109/SeGAH.2017.7939272
[35] Nelson Zagalo. 2007. A Diversidade Emocional nos Videojogos. *VI Brazilian Symposium on Computer Games and Digital Entertainment* (2007). http://www.microsoft.com/presspass/exec