POLITECNICO DI TORINO

Master's Degree in Biomedical Engineering

Master's Degree Thesis

Implementation of a virtual reality environment to elicit acute mental stress

Supervisors:
Prof. Luca Mesin
Ing. Matteo Raggi
Ing. Giulia Masi

Candidate:
Alice Marchisio

Academic Year 2024/2025

Abstract

Acute stress is a condition that arises in response to situations that are difficult to manage, manifesting in both physiological and psychological changes. Furthermore, the occurrence of stressors for a long period can result in a chronic stress condition, leading to serious health issues. To study stress in controlled conditions, research employs cognitive tests to induce acute mental stress and subsequently investigate people response and methods for stress mitigation. However, the results of cognitive tasks performed in a traditional way can be distorted by external events affecting participants, raising doubts about whether the observed effects can be attributed exclusively to the test itself. Recent studies have started implementing stress-inducing protocols in virtual reality (VR). Notably, VR-based approaches have shown better results than traditional cognitive tasks, making VR a promising tool for stress research. In this thesis, a game in VR was developed to elicit a stressful response. Notably, it does not require specific technical or social skills to be performed, making it suitable for a huge variety of people. The game is designed to stimulate universally present traits, such as reactivity to intense or unexpected stimuli and competitiveness (either against others or oneself), with the aim of generating a more uniform response. One of the purposes of this work is to assess whether stress levels measured during the VR game are as good as a mental arithmetic task (MAT) or higher. The experimental protocol, conducted on 20 subjects, included relaxation phases with calming videos and music to establish a baseline, the VR game and the MAT as a comparison. Stress during the protocol was quantified using self-report questionnaires and features of biological signals like the electrocardiogram (ECG) and the electrodermal activity (EDA). Results show significant differences in heart rate, breath rate and in the mean, median, maximum value and cumulative area of the EDA phasic component between the relaxation

phase and both stress-inducing tasks, suggesting that the proposed game elicit a consistent stressful response. These findings are supported by self-assessment questionnaires, which show no significant difference in perceived stress between the two stress-inducing tasks, but highlight a clear difference between the relaxation phase and both tasks. Questionnaire data also indicate that 95% of participants felt fully immersed in VR and disconnected from the external environment, confirming the effectiveness of the experiment. Future work will focus on techniques to reduce stress. Although this aspect was not considered in the current study, a preliminary step was addressed by developing a regression model, trained on the MAT and tested on the VR game, to predict the stress state of the user. Notably, the predictions can drive a stimulation with either auditory or mechanical stimuli to mitigate the effect of stress.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter provides a comprehensive introduction of the fundamental concepts necessary to understand the present work. It begins by introducing the concept of stress and the underlying physiological mechanisms involved, including stress response, cardiac function and skin conductance. Most of the content related to psychophysiological mechanisms of stress was taken from [1], while the material concerning anatomy and physiology of heart and skin was drawn from [2]. Then the chapter reviews the current state of the art in virtual reality–based stress research and concludes by describing the study that was carried out.

1.1 Stress: Definition and Classification

Stress is the way the human being respond to challenging, exciting or threatening situations [3][4] and it is characterized by specific physiological symptoms such as increased body activation. It can manifest itself in two main ways: Eustress and Distress.

Eustress is a form of positive stress that occurs when a person perceives a situation as stimulating and manageable. Psychologically, it is often associated with emotions such as enthusiasm, confidence, concentration and satisfaction. Typical examples include a new job, an engaging project or a challenging but rewarding sports training.

Distress, on the other hand, is a form of negative stress that arises when a situation is perceived as threatening or overwhelming. It is commonly accompanied by emotions such as anxiety, irritability, sadness or frustration. Common examples include persistent family conflicts or work overload.

However, some studies [4][5][6] emphasize how the distinction between eustress and distress is not determined by the objective characteristics of the event, but rather by the individual's subjective perception and the context where the stress response occurs. Due to the effects it produces, the most critical condition is distress. Therefore, from this point onward, only this form will be considered and the term "stress" will be used exclusively to refer to it.

Stress, like other emotions, can be classified according to three parameters [1]:

- **Valence**: it indicates whether a given emotion is categorized as positive or negative;
- **Arousal**: it represents whether the emotion is strong (high arousal) or weak (low arousal), regardless of its valence;
- **Dominance**: it describes the degree of perceived control or power that an individual feels they have while experiencing an emotion.

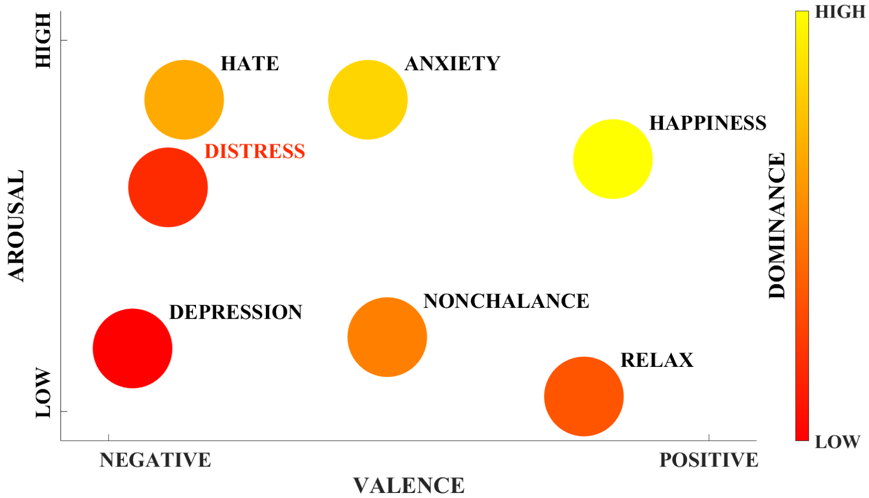

Figure 1.1: The emotional dimension of stress. Adapted from [1]. The emotions were located using the mean values reported in [7].

This classification describes stress as an emotion with negative valence, moderate arousal and low dominance. This means that stress elicits noticeable emotional activation that is perceived as negative and difficult to control, which is why stress mitigation techniques are studied to improve the individual's psychophysiological state.

Another objective and functional classification is based on the duration and frequency of the stressful response [8]. The first type is **acute stress**, an intense and short-term response to an unexpected or specific situation. It is the most common form of stress and, when limited to particular situations and periods, is not associated with significant risks. Prolonged exposure to stress-inducing stimuli, however, can lead to the development of **chronic stress**. This latter condition represents the most critical form of stress, because it can lead to physical problems (such as insomnia [9], gastrointestinal disorders [10], and hypertension [11]), psychological issues (such as anxiety, depression, and burnout [12]) and cognitive difficulties (such as concentration issues and memory problems [13]).

This thesis focuses on **acute mental stress** because it represents a transitory and safe condition for healthy individuals. Moreover, it is easily reproducible in experimental settings through cognitive tasks, and this makes it particularly suitable for studying the physiological mechanisms involved.

1.2 Neurophysiological Activation of the Stress Response

In response to stress, the human body activates a series of physiological mechanisms aimed at efficiently preparing it to respond to actual or perceived danger. This reaction is known as the "fight or flight" response and originates in the brain, where the amygdala, responsible for processing fear and emotions, detects danger and activates the hypothalamus.

From there, two main response pathways are activated:

- **Sympathetic Nervous System (SNS)** [14]: it starts an immediate response by stimulating epinephrine and norepinephrine release, so that the body is ready for physical action within seconds.

- **Hypothalamic-Pituitary-Adrenal (HPA) Axis** [15]: it provides a slower but more sustained response, leading to increased cortisol release, which maintains alertness and supports energy balance during prolonged exposure to stress.

These two systems provide a coordinated physiological response that ensures both rapid mobilization of resources for immediate action and sustained adaptation during prolonged stress exposure. This dual mechanism allows the body to effectively manage a wide range of stressors, from acute threats to ongoing challenges.

1.2.1 The Role of SNS

The sympathetic nervous system, together with the parasympathetic and enteric nervous systems, constitutes the autonomic nervous system, which is the part of the peripheral nervous system responsible for regulating involuntary functions of the body.

The sympathetic and parasympathetic systems perform distinct yet complementary functions: the sympathetic system is activated through the "fight or flight" mechanism, which prepares the body to respond to threats, while the parasympathetic system governs "rest and digest" functions, supporting recovery and homeostasis. The enteric nervous system, on the other hand, regulates the activity of the gastrointestinal tract and works differently depending on whether the sympathetic or parasympathetic system is dominant.

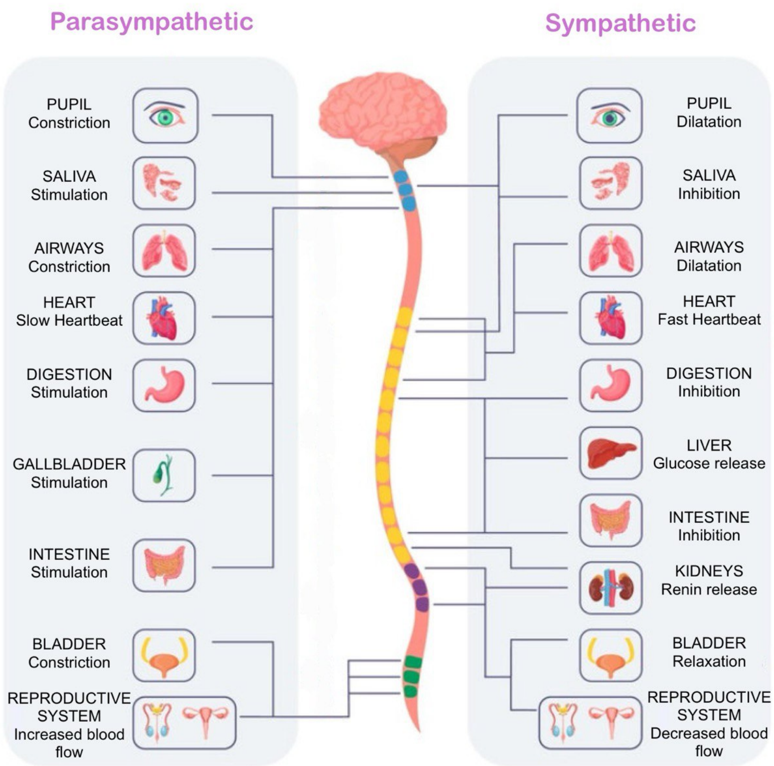

Figure 1.2: Parasympathetic vs Sympathetic Nervous System: functional comparison.

The activation of the sympathetic nervous system stimulates the adrenal medulla, the inner part of the adrenal glands, to release epinephrine and norepinephrine. These hormones trigger a range of physiological responses, including:

- Cardiovascular effects, with an increase of heart rate and blood pressure;
- Respiratory changes, including accelerated breathing and bronchial dilation to improve blood oxygenation;
- Neurological and cognitive responses, with the presence of hypervigilance and increased reactivity to stimuli;
- Muscular and metabolic alterations, characterized by increased blood and oxygen flow to skeletal muscles and reduced digestion and peripheral circulation, to prioritize physical action in case of danger;
- Cutaneous and visceral reactions, because increased sweating helps regulate body temperature during acute stress.

These changes occur automatically and adaptively, representing a fundamental evolutionary strategy for survival. However, when activated in the absence of actual danger or sustained over prolonged periods, they may contribute to the development of pathological conditions.

1.2.2 HPA Axis

The hypothalamus produces corticotropin-releasing hormone (CRH), which reaches the anterior pituitary gland through the hypothalamic-pituitary axis and stimulates the release of adrenocorticotropic hormone (ACTH) from the anterior pituitary. ACTH acts on the adrenal glands stimulating the production of corticosteroids, particularly cortisol, by the adrenal cortex (the outer and more extensive part).

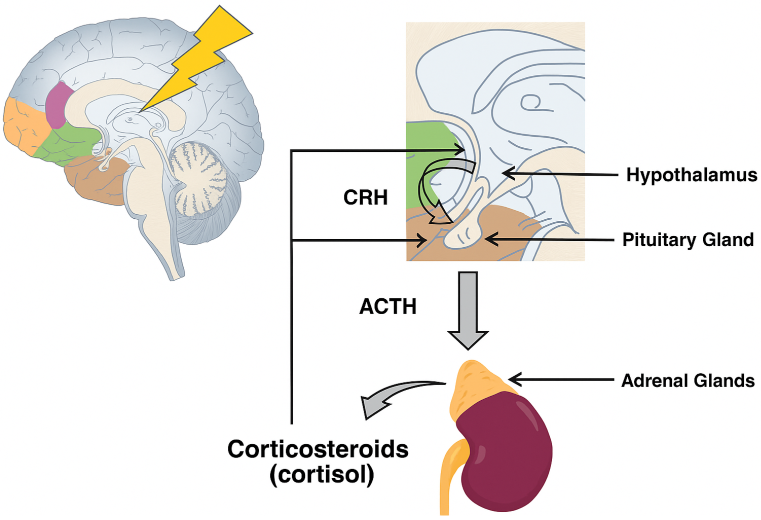

Figure 1.3: Mechanisms of the Hypothalamic–Pituitary Axis.

Cortisol does not cause immediate physical symptoms, but it modulates and prolongs those produced by the sympathetic nervous system in response to stress. In addition, it maintains elevated blood glucose levels to ensure energy availability when required and simultaneously inhibits non-essential short-term functions such as immune activity and digestion.

HPA axis has a self-regulating system: corticosteroids produce a negative feedback effect on both the pituitary gland and the hypothalamus, leading to a reduction in ACTH production as their concentration increases.

1.3 Heart and Electrocardiogram

1.3.1 Anatomical and Functional Overview

The heart is the central organ of the cardiovascular system [16] and plays a fundamental role in maintaining blood circulation throughout the body. It is an involuntary striated muscle that functions as a **pump**, propelling blood in the appropriate direction through a precise system of contractions and valves. The heart is divided into four chambers, split into two categories: atria and ventricles. These are arranged symmetrically, with one atrium and one ventricle on the right side and one atrium and one ventricle on the left, separated by a septum. The atria are responsible for receiving blood returning to the heart, while the ventricles pump blood either toward the lungs or the rest of the body.

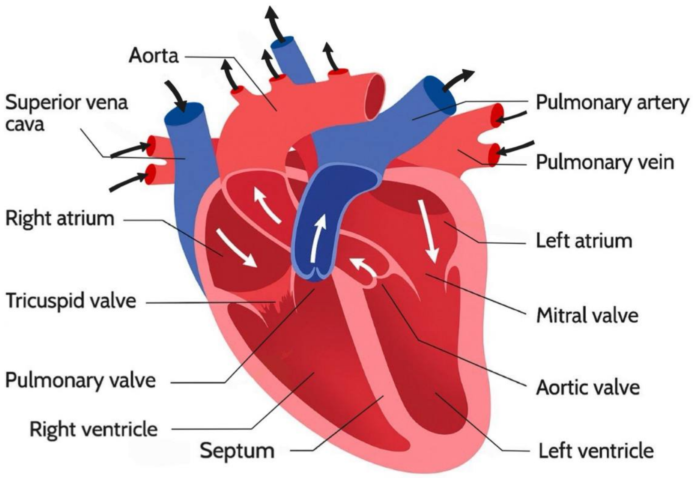

Figure 1.4: Anatomy of the heart.

Specifically, deoxygenated blood enters the right atrium and, passing through the tricuspid valve, flows into the right ventricle. From there, it is pumped through the pulmonary valve into the lungs for oxygenation: this process is known as **pulmonary circulation**. Oxygenated blood then returns to the left atrium, passes through the mitral valve into the left ventricle and is subsequently pumped into the **systemic circulation** through the aortic valve and the aorta. The principal role of cardiac valves is to ensure a unidirectional flow of blood and prevent any backflow.

1.3.2 Autonomic Regulation

Cardiac activity is regulated by the autonomic nervous system through efferent pathways, which are nerves that transmit signals from the central nervous system to the heart. As discussed in Section 1.2, the sympathetic and parasympathetic systems have opposing functions.

The sympathetic system originates in the thoracolumbar spinal cord and its fibers reach the heart trough the **cardiac sympathetic nerves**. These fibers release epinephrine and norepinephrine, with norepinephrine being the main neurotransmitter that acts on cardiac receptors to increase heart rate, contractile force and conduction velocity.

The parasympathetic system, in contrast, originates in the brainstem and influences heart activity through fibers that emerge from the **vagus nerve**. These fibers release acetylcholine, which decreases heart rate and conduction velocity, while having little effect on contractile strength. This neural regulation complements hormonal control mechanisms, such as those of the hypothalamic-pituitary-adrenal axis.

This autonomic regulation is essential because the heart is the main organ responsible for supplying blood and oxygen to all body tissues. By modulating cardiac function, the nervous system can quickly adapt the physiological conditions of the entire organism.

1.3.3 Cardiac Electrical Conduction System

The heart has a specialized electrical conduction system that coordinates the rhythmic contraction of the cardiac muscle. Each heartbeat follows a precise sequence of phases [13].

The process begins in the **sinoatrial (SA) node**, located in the upper wall of the right atrium, which spontaneously generates an action potential that triggers a change in membrane potential and initiates the depolarization of atrial cells. The impulse then spreads through the atrial muscle fibers, causing complete depolarization of the atria and their contraction, known as atrial systole. After this contraction, atrial cells enter the repolarization phase, returning to their resting membrane potential. When the impulse reaches the **atrioventricular (AV) node**, conduction slows down creating a delay that allows atrial diastole, a complete emptying of the atria before ventricular contraction begins. The electrical signal subsequently travels along the His bundle and is distributed through the Purkinje fibers, producing rapid depolarization of ventricular cells and resulting in ventricular contraction, called ventricular systole. Finally, the ventricular cells repolarize, restoring their resting membrane potential and allowing ventricular diastole, the relaxation phase of the ventricles.

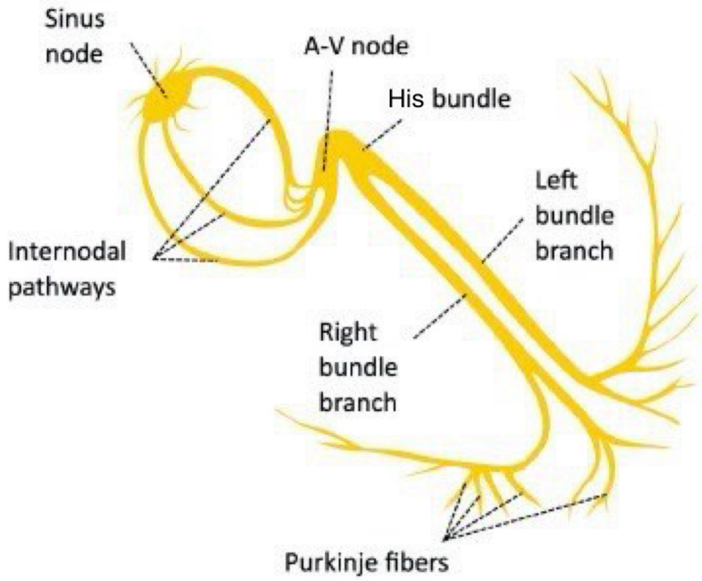

Figure 1.5: Cardiac electrical conduction system.

The duration of these electrical phases is not constant but may vary in response to multiple factors. In particular, the activation of the autonomic nervous system during stress events alters the frequency and regularity of heartbeats. This phenomenon is known as **heart rate variability (HRV)**.

1.3.4 ECG Signal

The electrocardiogram (ECG) is the graphical representation of the heart's electrical activity. It reflects the variations in electrical potential associated with the depolarization and repolarization of myocardial cells during each cardiac cycle and includes several characteristic waves, each of them with a particular duration range [17]:

- **P wave**: it represents atrial depolarization, which precedes atrial contraction (atrial systole). It has an average duration of 60–120 ms;
- **QRS complex**: it corresponds to ventricular depolarization, which initiates ventricular contraction (ventricular systole). Atrial repolarization occurs

simultaneously but it is masked by this complex. The QRS duration is approximately 60–90 ms;

- **T wave**: it represents ventricular repolarization and the return of ventricular cells to their resting potential. This event coincides with the beginning of ventricular diastole, the relaxation and filling phase. It has an average duration of 100–250 ms.

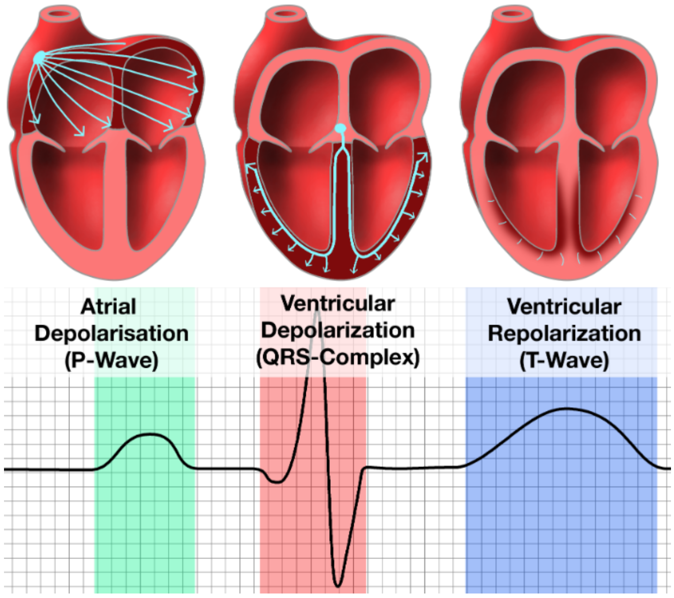

Figure 1.6: ECG signal: characteristic waves and corresponding electrical activity.

These are the main waves, representing the key events in cardiac electrical conduction. It is also important to analyze the intervals between these waves to accurately estimate the total duration of the cardiac cycle:

- **PQ interval**: measures the time taken for the electrical impulse to travel from the atria to the ventricles. It includes atrial depolarization (P wave), the delay introduced by the atrioventricular node and the onset of ventricular depolarization. The average duration is approximately 120–200 ms;

- **QT interval**: represents the total duration of ventricular systole, from the start of the QRS complex to the end of the T wave, and normally ranges from 320 to 470 ms;

- **TP interval**: corresponds to the period of complete diastole, during which both atria and ventricles are at rest, and generally varies between 200 and 420 ms.

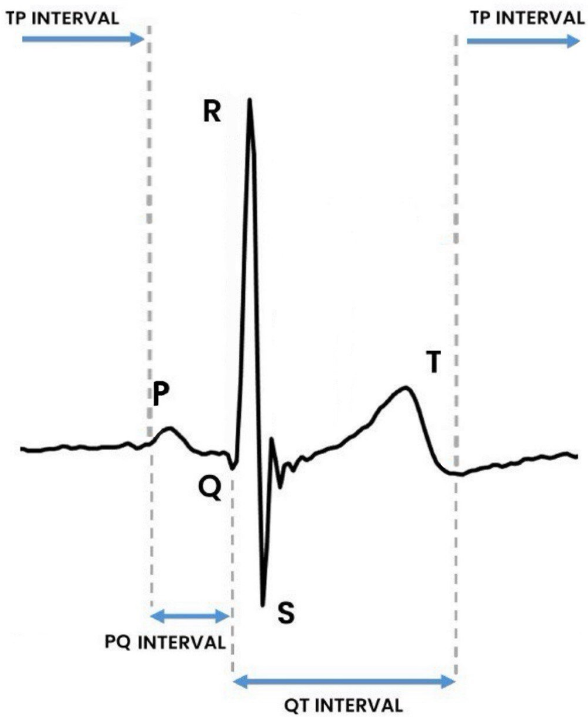

Figure 1.7: Example of ECG signal.

By summing the PQ interval (including the P wave), the QT interval (from the beginning of the QRS complex to the end of the T wave) and the TP interval (representing the period of complete diastole), an estimate of the average duration of a cardiac cycle can be obtained. These values are indicative and may vary depending on an individual's physical condition or the presence of cardiovascular pathologies. The instantaneous heart rate can then be calculated as the reciprocal of the duration of a single cardiac cycle, which corresponds to the RR interval. Since heart rate is normally expressed in beats per minute (bpm), the result must be multiplied accordingly to convert from milliseconds to minutes:

$$\text{Heart Rate (bpm)} = \frac{1}{\text{RR interval (ms)}} \times 1000 \times 60$$

1.3.5 EDR Signal

ECG-Derived Respiration (EDR) refers to the **estimation of the respiratory signal** indirectly from the ECG. It is widely used in both clinical and research contexts when a direct measurement of respiration is unavailable. EDR is based on the physiological and anatomical effects that the respiratory cycle induces on the ECG signal. As highlighted in Subsection 1.2.1, respiration itself changes under stress, making EDR potentially valuable for stress assessment.

EDR is based on the idea that breathing changes the heart's position and orientation relative to the ECG electrodes. During inspiration, the diaphragm contracts and shifts down, allowing the lungs to expand. Consequently, the heart is slightly displaced and rotated, which modifies its electrical projection onto the leads and produces observable variations in the ECG signal. In particular, the amplitudes of the R and T waves decrease during inspiration because the air-filled lungs have a lower electrical conductivity compared to body tissues, and the time required for the R wave to reach its maximum value may shift.

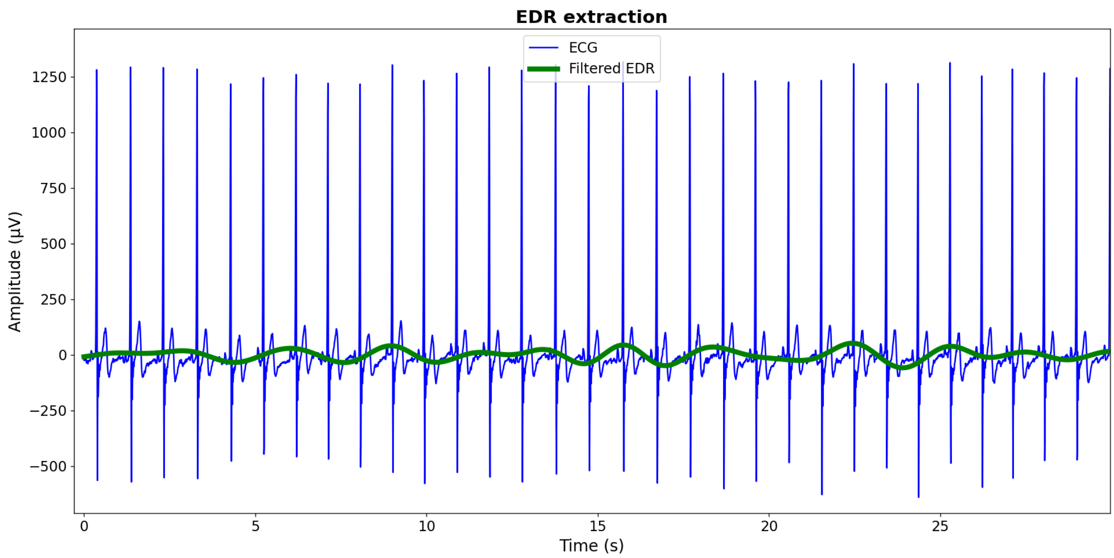

Figure 1.8: ECG, R peaks interpolation and filtered EDR.

1.4 Skin and Electrodermal Activity

1.4.1 Anatomical and Functional Overview

The skin is the human body's external surface and it is composed of three layers [18]. **Epidermis** is the most superficial layer, mainly composed of keratinocytes. It contains neither blood vessels nor sweat glands and its primary function is to act as a protective barrier. **Dermis** is the intermediate layer, essential for the support and nourishment of the skin. It contains collagen and elastin fibers to ensure the skin's strength and elasticity, capillaries to provide oxygen and nutrients, nerve endings for tactile sensitivity, sweat glands to regulate temperature and sebaceous glands to keep the skin hydrated. **Hypodermis**, also known as subcutaneous tissue, is the deepest layer and it's composed of adipose and connective tissue. It functions as thermal insulation and energy reserve.

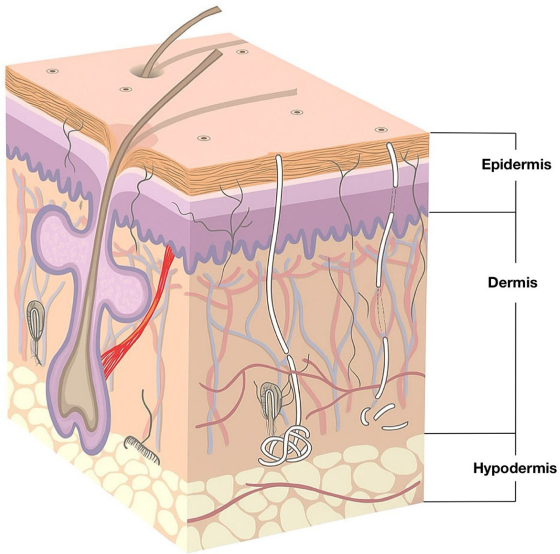

Figure 1.9: Skin layers: epidermis, dermis and hypodermis.

Sweating is controlled by sweat glands, which are divided into two main categories [19]. **Eccrine glands**, which produce a watery sweat (containing only water and salts) that is released directly through pores. They are distributed throughout the body but especially concentrated on the palms of the hands, soles of the feet and forehead. **Apocrine glands**, instead, produce a thicker sweat (containing proteins and lipids) that is secreted through hair follicles. They are located in specific areas such as the armpits and genital region.

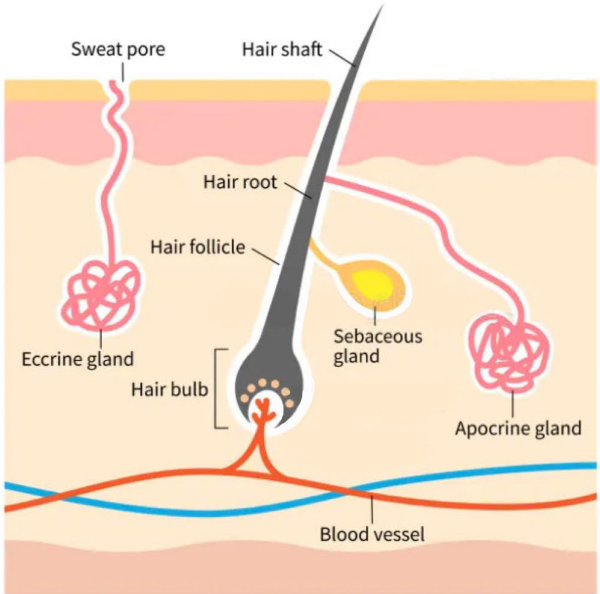

Figure 1.10: Sweat glands: Eccrine and Apocrine.

1.4.2 Electrodermal Activity and SNS Response

Skin conductance represents the skin's ability to conduct electricity and depends primarily on the level of skin wetness: the more humid the skin, the higher the conductance.

Electrodermal activity (EDA), also known as Galvanic Skin Response (GSR), measures variations in the skin's electrical conductance. This signal depends mainly on eccrine sweating, which occurs in much greater quantities than apocrine sweating and it is highly conductive due to its high salt content. For this reason, to obtain more accurate measurements, EDA sensors are typically placed in areas with a high density of eccrine glands, such as the palms of the hands or the fingers.

Eccrine sweat glands are controlled by the Sympathetic Nervous System, specifically by Sympathetic Sudomotor Nerves, so their activity is considered a direct indicator of stress and emotional activation.

1.4.3 EDA Signal

The EDA signal by itself does not provide useful information, it must be decomposed into its components because each one reflects different aspects of sympathetic activity.

The two main components are:

- **Skin Conductance Level (SCL)**: also known as tonic component, it is the baseline level of skin conductance in absence of stimuli. It varies slowly (over minutes) and reflects general sympathetic activity. It depends on external factors such as temperature, hydration and prolonged background stress. It is used to monitor sustained attention states.

- **Skin Conductance Response (SCR)**: this is characterized by rapid and transient variations associated with specific stimuli, called phasic activity. It has a very quick onset (1–3 seconds after the stimulus) and a short duration (a few seconds). It can be used to quantify stress and isolate stressful stimuli.

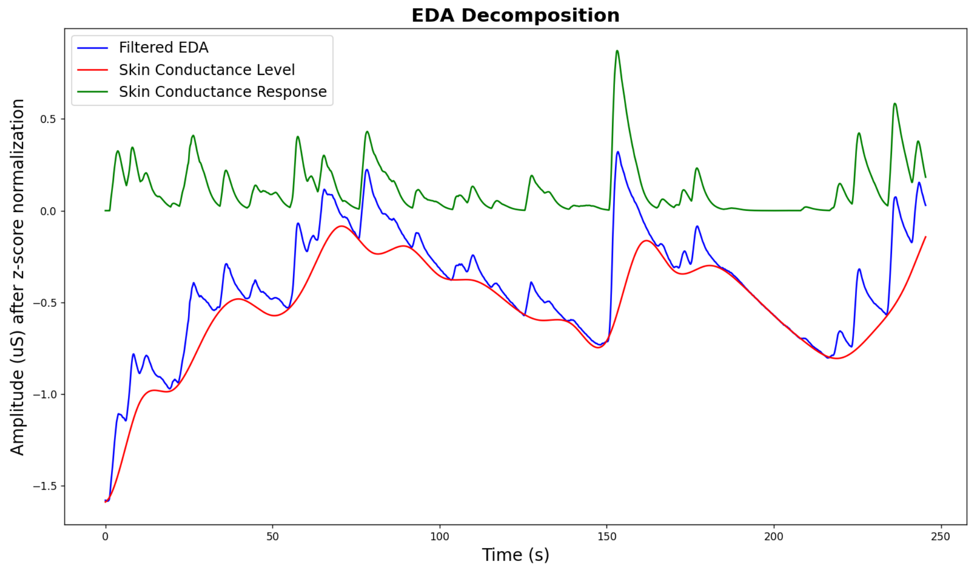

Figure 1.11: EDA decomposition during a stress-inducing task. Data were recorded using a Shimmer sensor at 64 Hz and preprocessed with z-score normalization and a 2 Hz low-pass filter.

In relaxation conditions, EDA typically shows a low and stable SCL with little SCR. This indicates minimal sympathetic nervous system activity, typical of situations without significant emotional or cognitive stimuli.

In contrast, stressful situations cause an increase in SCL, while SCR exhibits a higher number of responses per minute and amplitudes greater than in resting state.

Numerically [20], the mean SCL ranges from about 0.5–2 µS up to values that can reach 5 µS, while the number of SCRs can increase from 1–2 to about 7–10 per minute. These changes indicate elevated sympathetic activation, reacting to perceived challenging or threatening stimuli.

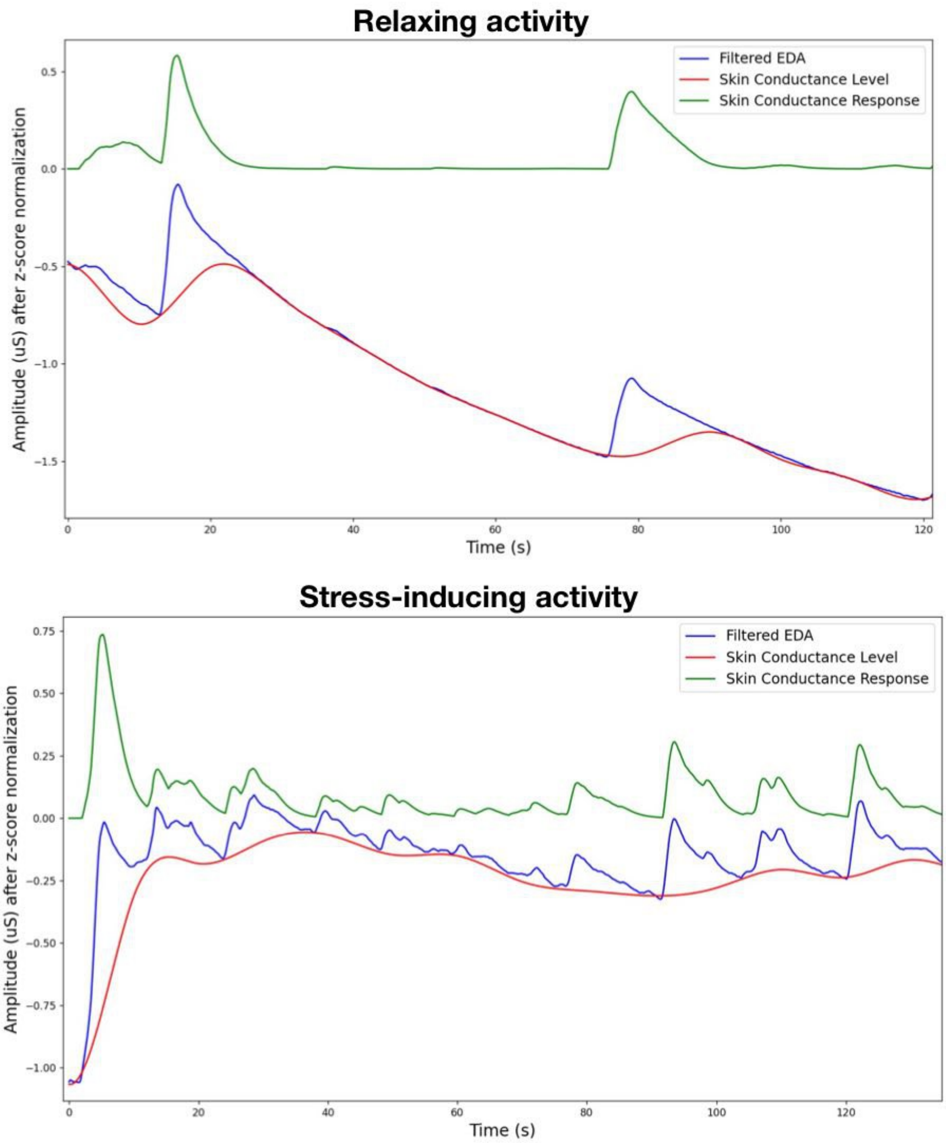

Figure 1.12: Comparison of EDA signal during relaxing and stress-inducing activity. Data were collected with a Shimmer sensor at 64 Hz, preprocessed with z-score normalization and a 2 Hz low-pass filter. SCL ranges cannot be observed due to normalization, but a higher mean value can be noted during the stress-inducing activity.

1.5 VR in Stress Research: State of the Art

1.5.1 Definition of Extended Reality

The term **Extended Reality (XR)** is used to describe all technologies that modify our perception of reality through the use of digital tools. XR incorporates three main subcategories:

- Virtual Reality (VR);
- Augmented Reality (AR);
- Mixed Reality (MR).

VR creates a completely digital environment that replaces the real world, allowing the user to immerse themselves in a simulated experience and interact with the surrounding environment. Examples include video games played with headsets, where users can explore new worlds while remaining physically stationary.

AR creates an environment that overlays digital elements onto the real world, without disconnecting from it but rather enriching it (thus the term "augmented"). However, this is merely a graphical superimposition, meaning that the virtual objects added to the scene do not interact with the real-world environment. Examples include smartphone applications (or specialized AR glasses) that, when framing or observing a historical building, display information such as its name and history.

MR creates a more complex environment where real and virtual objects coexist and interact in a realistic manner. In some cases it is even possible to interact with virtual objects directly using hands, without the need for joysticks or gloves. Examples include applications that allow users to visualize how pieces of furniture would fit within their own living spaces.

For the purposes of the proposed study, the focus will be exclusively on VR.

1.5.2 The Role of Virtual Reality in Healthcare

In recent years, VR has become an essential tool in the medical field. It is now applied in a variety of areas, including:

- **Medical training and education** [21][22]
  Virtual environments are created to faithfully replicate the hospital environment, allowing healthcare professionals to learn and refine techniques in complete safety, thereby reducing the risk of errors during real interventions. Compared to traditional mannequins, VR enables the simulation of a wider range of complex clinical scenarios, making the training experience more dynamic and adaptable.

- **Motor rehabilitation** [23][24]
  Immersive programs are developed to assist patients in performing specific exercises in a more engaging manner, supported by visual and auditory feedback that facilitates the recovery of motor functions.

- **Surgical support** [25][26]
  VR systems are designed to assist surgeons during real operations by providing, for example, detailed 3D visualisations of the patient's anatomy. This can offer significant advantages, such as improved precision and reduced operative times.

- **Research** [27][28]
  It is possible to create immersive and controlled environments customized to the specific needs of any research field. In these settings subjects are isolated from external stimuli, allowing researchers to minimize unwanted variables and thus ensure greater standardization of experimental conditions. This approach enables a more accurate observation of human physiological and cognitive responses to specific stimuli, resulting in more reliable and reproducible experimental outcomes.

The following section will explore the use of VR in stress research, highlighting how this technology creates controlled environments where psychophysiological responses are directly influenced by the surrounding conditions.

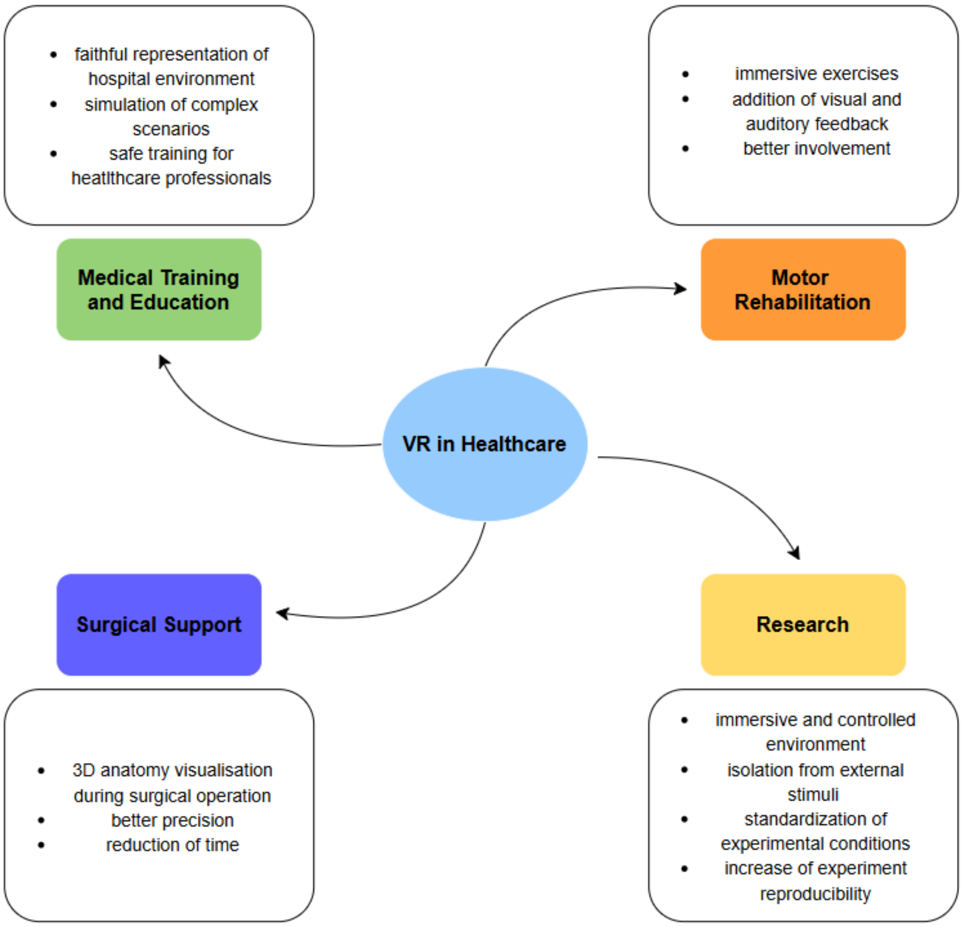

Figure 1.13: The role of VR in Healthcare.

1.5.3 Use of Virtual Reality in Stress Study

In stress studies using virtual reality, traditional tasks already found in the literature are often repurposed, adapted and enriched with specific elements that increase their effectiveness. VR also allows precise control of the environment, isolating from the outside world and limiting confounding effects. Such features can promote results that are comparable or superior to traditional protocols.

Tasks designed for this purpose are generally divided into two categories: **passive stimulation**, in which the subject is exposed to sensory stimuli without actively responding or interacting, and **active stimulation**, in which the subject is actively engaged and must interact with the environment.

A case study [29] attributable to passive stimulation consisted of participants **experiencing a 360° video** of a roller coaster ride through a headset, featuring high speeds, sharp turns and steep slopes. Two rest sessions of 180 seconds, one before and one after exposure to the stimulus, involved participants remaining still and relaxed while watching a video of a forest with a river accompanied by calming music. These periods served as a baseline, a reference state of rest and relaxation, used to assess how physiological parameters changed in response to the stimulus. The results showed that, even in the absence of active interaction, intense and unpredictable sensory stimuli can generate a significant physiological response. However, it is important that the exposure occur only once for each subject, as any repetition could trigger a learning effect, reducing the effectiveness of the stimulus and nullifying differences from baseline.

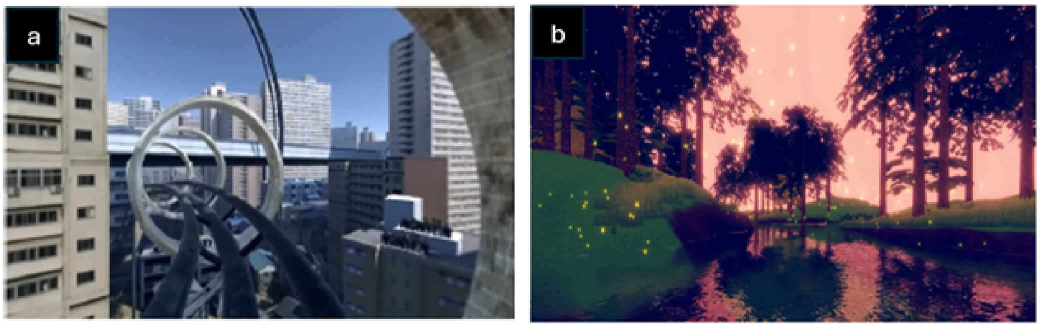

Figure 1.14: The 360° video of the roller coaster (a) and a forest with a river (b) ([29]).

Another study [30], belonging instead to active stimulation, employed virtual reality to recreate a **battlefield scenario** where the subject can move freely in space. The main objective was to test whether the VR environment could elicit stress levels similar to those experienced by veterans during war. The key distinction from the previous case lies in the alternating presence of highly stressful zones and neutral or stimulus-free zones. It was observed that significant stress levels were generated specifically when transitioning between safe zones, where the environment was calm and free from danger, and ambush zones, where sounds of gunfire and visible bullet traces created a tense atmosphere.

Despite the differences between the two studies, in terms of type of stimulation and scenario, both highlight how the unpredictability of the situation plays a crucial role in triggering a stress response that can be measured by physiological parameters. The concept of unpredictability is also one of the eight situational characteristics described by [31] that influence how a person appraises the stressfulness of a situation.

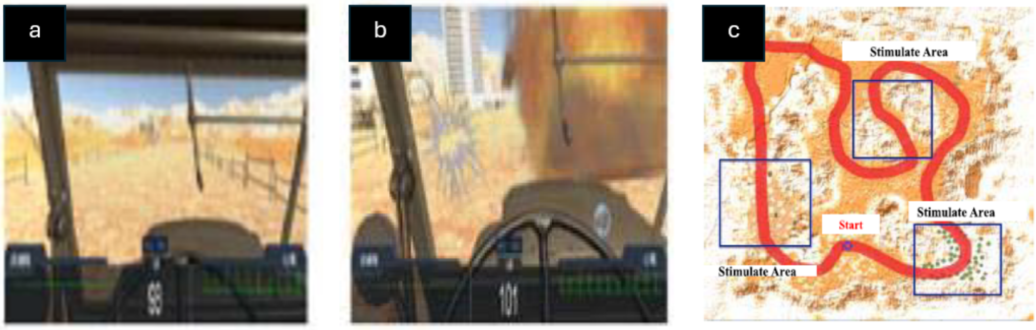

Figure 1.15: View of the player in a safe area (a), in an ambush area (b) and the distribution of ambush areas along the path (c) ([30]).

The category of active stimulation also includes **mental rotation task (MRT)** and **mental arithmetic task (MAT)** which, despite not involving movement through virtual space, still require information processing, decision making and response to cognitive stimuli. In cases where physical interaction with the environment is missing, there is a risk of reduced immersive quality. To mitigate this limitation, one study [32] introduced a preliminary interaction task, where the participant had to pick up and manipulate objects within a specially designed virtual scene. This introductory phase was intended to familiarize the subject with the environment and increase the sense of presence, especially for those approaching virtual reality for the first time. The same study also compared three modes of instruction delivery: first-person audio, interaction with a virtual character who spoke in second-person and written text. The results showed that instructions delivered in the second person via an external character tended to generate the highest level of stress, although the difference was not statistically significant.

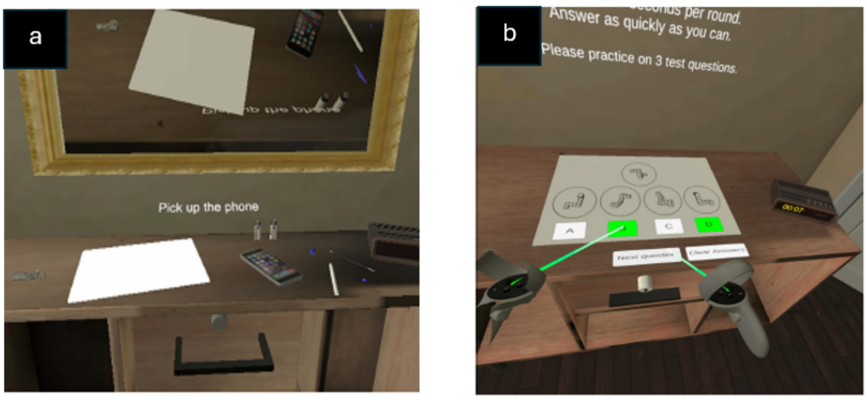

Figure 1.16: Preliminary interaction task (a) and rotation task (b) ([32]).

An alternative strategy designed to enhance the sense of immersion was proposed by a study [33] based on the **Stroop test**, a cognitive task used to assess attention control and the ability to resist cognitive interference. In the most common version of the test, participants are asked to name the color of a word while disregarding the meaning, which typically denotes a different color, thereby creating a cognitive conflict. The study compared a classic 2D version on PC with an immersive virtual reality version, set in a video game context simulating a desert war conflict. In the latter, the subject is driving a vehicle while words appear superimposed on the scenario, which alternates between safe areas (with no special stimuli) and ambush zones, where stressful events such as explosions, gunfire and shouting occur. Results showed that, for assessing cognitive decline, the 2D version is more effective, as it is closer to real-world conditions. However, the VR version showed a higher level of immersive nature, confirming the interest for the use of VR in studies aimed at inducing or measuring stress.

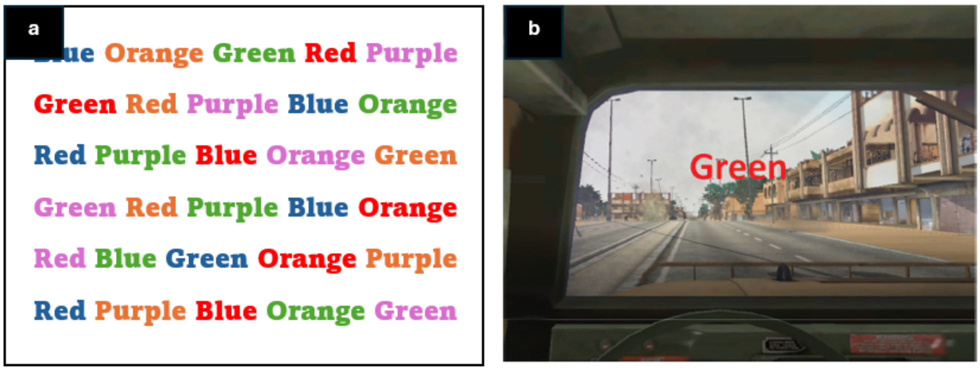

Figure 1.17: Traditional Stroop Test (a) and virtual reality version (b) ([33]).

The most frequently reproduced type of scenario in the VR environment is the **Trier Social Stress Test (TSST)** [34][35][36][37], one of the most established protocols for standardized stress induction in the laboratory. The TSST consists of two main phases: anticipation phase, a short period of time where participants can prepare a speech to introduce themselves and highlight their skills, with the knowledge that they will be observed and judged, and performance phase, where participants deliver their speech in front of judges, knowing that they will be recorded. This presentation is often followed by an arithmetic task (for example counting backwards quickly and accurately), also performed under observation.

Exposure to a judging context by strangers, along with the requirement to perform cognitive tasks under observation, can trigger a physiological stress response.

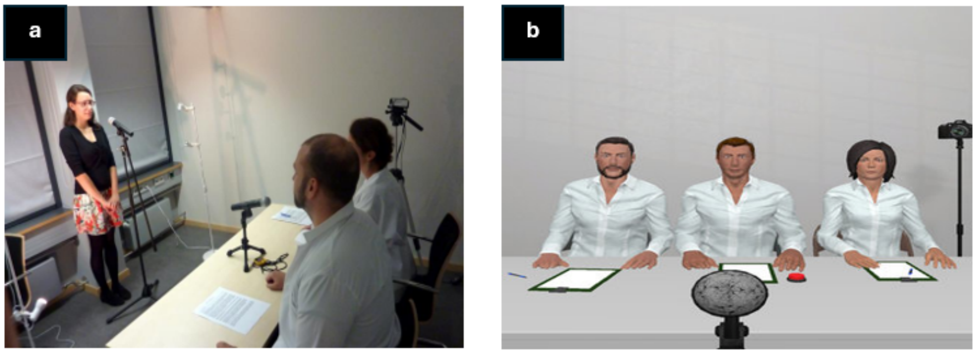

Figure 1.18: Traditional TSST (a) and virtual reality version (b) ([38], [39]).

Measurement methods adopted

The main signals monitored to assess stress response include:

- **Cardiac activity**, detected by ECG tracing, a direct indicator of autonomic nervous system activation;
- **Respiratory rate**, estimable from the ECG tracing, which varies in states of anxiety or physiological activation, becoming less predictable;
- **Skin conductance**, which reflects sweat glands activity and varies in response to intense emotions, measured trough skin-contact sensors;
- **Level of cortisol**, a stress hormone, measured through salivary samples and useful for assessing the endocrine response to acute stress;
- **Eye movements** [30], measured by eye-tracking (implemented in some headseats) and correlated with cognitive engagement and emotional state.
- **Self-assessment**, collected by questionnaires investigating perceptions, emotions and psychological states associated with the stressful experience.

Among the parameters considered, most indicators generally showed sharper differences between resting and stressful conditions in VR tasks compared to traditional tasks, whereas cortisol was the only measure to show less marked changes [37]. The authors suggest that VR tasks may enhance stress-related arousal due to novelty and immersion, which consequently increases sympathetic activity and related physiological parameters. In contrast, because the brain perceives the threat as less 'real', the activation of the HPA axis is reduced, resulting in attenuated cortisol responses. However, it is important to note that these observed differences between stress conditions in VR settings compared to traditional environments were not statistically significant.

Recommended practices for using virtual reality

VR has shown to be a particularly interesting tool for stress research, as it allows the simulation of realistic environments while maintaining full control over the stimuli administered, avoiding external interference such as environmental noise or the presence of observers, which can unintentionally influence the subject's experience.

Based on the VR studies discussed above, the literature highlights several best practices to achieve reliable results and take full advantage of the benefits of this tool:

- to introduce **familiarization sessions** for less experienced subjects, so that they can subsequently perform the task without difficulty;
- to provide **clear instructions** to ensure that the task is performed in the correct manner, keeping in mind that instructions delivered as commands by virtual characters may elevate the perceived level of stress;
- to make the **environment immersive and believable**, as the absence of such features could negate the effectiveness of using VR;
- to add **unexpected elements** such as sounds or lights to elicit stress;
- to expose subjects to **relaxing videos** in order to measure the baseline;
- to **dynamically adapt** the virtual environment based on the subject's individual responses, making the experience more personalized and effective.

However, the use of VR is not free from **side effects**: the most prominent is **cybersickness** [29], a disorder related to the discrepancy between visual input and vestibular and proprioceptive signals, which can manifest as dizziness, headache, blurred vision and difficulty in concentration. To reduce the risk of occurrence of such symptoms, it is advisable to keep task duration relatively short.

1.6 Relationship between Competitiveness and Stress

One topic that has been widely examined in the scientific literature is the relationship between competitiveness and stress response [40][41]. Excessive competitiveness can significantly amplify stress responses in various areas of daily life, particularly in environments such as school and sport, where pressure to achieve high results and constant comparison with others can compromise individuals' psychological and physical well-being.

A scientific experimental study [42] explored this relationship through two experiments that employed some mathematical tasks and the TSST, one of the most widely used paradigms in stress research, as already mentioned in Section 1.5.3.

In the first experiment, participants were asked to complete three mathematical tasks:

- the first task was carried out under a piece-rate scheme, where participants received a fixed reward for each correct answer, without a direct competition;
- the second task was conducted in tournament mode, where participants competed against each other and only the highest-scoring individual received the reward, but it is considerably larger than in the previous case;
- in the third task participants could choose which of the two compensation schemes they preferred.

This experimental design enabled a comparison of the physiological stress responses between a neutral context (the first one) and a highly competitive one (the second one). Moreover, the choice regarding the compensation method in the third task allows participants to be classified as either "more competitive" or "less competitive", allowing the researchers to investigate potential differences between these groups.

The second experiment aimed to explore whether the induction of acute mental stress could influence individuals' tendency to engage in competitive situations. For this purpose, participants were first exposed to the TSST and were then asked to choose whether to complete a mathematical task with a piece-rate scheme or in tournament mode.

The main findings of the study were:

- Competition increases the stress response, as cortisol levels were significantly higher during the tournament condition compared to the piece-rate condition.
- More competitive individuals exhibit stronger stress reactivity, as participants who voluntarily chose the tournament mode in the third task of the first experiment also showed a higher physiological stress response in all the other tasks, indicating a stable individual trait.
- Stress does not increase competitiveness, as inducing acute stress through the TSST did not enhance participants' likelihood of choosing the competitive condition in the subsequent task.

1.7 Current Study

As previously highlighted in Section 1.5.3, the literature presents a wide range of studies that employ VR as a research tool. In particular, VR is used to create stressful environments under controlled conditions, allowing the subject to be **isolated from the outside world** and reducing confounding effects.

In this study, a virtual game was developed with the aim of integrating various stress-inducing criteria identified in previous research. Specifically, the designed environment includes the following features:

- **Competitiveness**: the game was designed as a race to create a competitive environment that stimulates participants ambition to excel [42].
- **Simultaneous and highly intense stimuli**: the virtual environment presents simultaneous and overlapping stimuli that are difficult to ignore due to their high intensity, with the goal of strongly engaging the subject's attention and cognitive resources [33].
- **Simple second-person instructions**: the use of the second person was intentionally chosen to make the instructions perceived as commands, thus increasing the level of stress [32]. However, they were written clearly to facilitate understanding and avoid further complications, especially for less experienced subjects.
- **Familiarization phase**: an initial phase was included to allow less experienced users to become familiar with the virtual environment, reducing anxiety related to uncertainty and facilitating the correct execution of the task. This approach enables a more accurate isolation of stress induced exclusively by stimuli designed in the virtual context [32].
- **Unexpected situations**: some stimuli are presented unpredictably to prevent the subject from anticipating their timing, thereby increasing uncertainty and minimizing the risk of learning effects [29][30].

Therefore, the current study aims to create an environment designed to stimulate universally present traits, such as reactivity to intense or unexpected stimuli and competitiveness (either against others or oneself), without requiring specific technical or social skills, with the goal of eliciting a more uniform response.

Chapter 2

Materials and Methods

This chapter provides a detailed description of the equipment used for delivering the procedure and collecting physiological signals, the sample selected for the study, the protocol followed during the acquisition phase, the questionnaires administered, the experimental environment adopted and the methods employed for data analysis.

2.1 Materials

To carry out the selected protocol, an experimental setup was implemented consisting of the different components.

POLAR H10 CHEST STRAP
It includes electrodes integrated into the inner surface of the strap that detect the ECG signal and thoracic movement. It must be worn on the chest just below the sternum in direct contact with the skin, with the conductive areas moistened to optimize signal acquisition. It connects via Bluetooth to a PC and collects data at a sampling rate of 130 Hz.

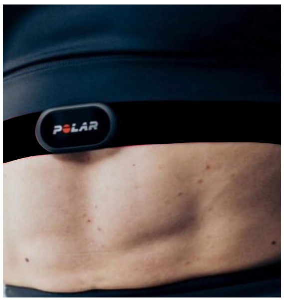

Figure 2.1: Polar H10 chest strap.

SHIMMER SENSOR

This device is used to measure EDA and is equipped with an accelerometer to detect movement of the limb. It is applied to the base of the index and middle fingers, on the side closest to the palm. It operates as a stand-alone device, storing data internally. These data are later downloaded using Consensys, the official software developed by Shimmer Sensing to configure, acquire, visualize and export sensor data. The sampling rate is 64 Hz.

Figure 2.2: Shimmer Sensor.

META QUEST 2 HEADSET

Used for the delivery of a stage of the protocol, it is worn over the participant's eyes, who holds a controller in the left hand. The right hand remains still, as several sensors are attached to it; any movement could introduce motion artifacts that would compromise signal quality.

Figure 2.3: Meta Quest 2 Headset.

ZOTAC ZBOX E SERIES MAGNUS EN52060V COMPUTER

It is a compact desktop PC equipped with a 9th generation Intel Core i5 processor, an NVIDIA GeForce RTX 2060 compatible with the VR headset used, and 16 GB of RAM. This device was used to display another stage of the protocol on screen, to connect the headset via cable during the second task and to acquire the data transmitted from the Polar H10 chest strap.

2.2 Participants

Initially, a pilot study was conducted with 10 participants, during which only the ECG signal and chest acceleration were recorded. A subsequent data collection involved another 10 participants, and in this session, both the ECG and EDA signals were recorded. Since the experimental protocol remained unchanged, all ECG data were later processed together.

The twenty subjects of the study include 12 males and 8 females, with an average age of 26.4 $\pm$ 2.7 years and an educational level of bachelor's degree or higher. The decision of selecting a **homogeneous sample** for age and education had the aim of reducing internal variability. This approach is particularly useful when there is a numerically limited sample, as it makes it easier to obtain meaningful results. It should be emphasized, however, that such homogeneity circumscribes the relevance of the results to the specific population segment under consideration; since this is a preliminary study, this limitation should be considered acceptable.

The entire experimental protocol was performed with the left hand only, while the right hand was kept still to avoid artifacts in the signals caused by muscle movements. The choice to use the left hand stemmed from the fact that it was intended to have the **non-dominant hand used**, and generally the majority of people are right-handed, which proved to be true in this case as well (19 right-handed, 1 left-handed). The decision was motivated by the need to reduce differences between participants with different levels of familiarity with video games, and it was thought that the use of the non-dominant hand could help to align the level of perceived difficulty and the experience of use, making the sample more homogeneous from a performance point of view.

Since cortisol levels are typically higher in the early morning and late evening due to circadian rhythms, data were collected between 10 a.m. and 4 p.m.

2.3 Experimental Protocol

This study was developed starting from a well-defined experimental question: **is it possible to induce acute mental stress in participants during the execution of a virtual reality game?** Although the proposed tasks differ in nature, the objective is to compare the participant's psychophysiological response to that typically observed during the MAT, a widely used technique for the controlled induction of stress

To address this question, an experimental protocol was structured in two parts, each involving the execution of a specific task: a computer-based MAT, hereafter referred to as Math Task, and a VR game, from this point onward referred to as Game Task. Both tasks are preceded by a baseline phase and followed by a recovery phase. The recovery phase is intended to allow more slowly-varying physiological parameters (such as EDA) to return to baseline, avoiding strong summation effects that could undermine the validity of subsequent results.

The procedure includes the following phases, summarized below:

- **Instrumentation Setup**
  This phase lasts 10 minutes and involves positioning and activating all the sensors required for monitoring physiological parameters.

- **Administration of Pre-Test Questionnaires**
  Participants compile questionnaires related to general health status, personal habits and stress condition. The aim is to collect baseline data unaffected by subsequent experiences. This phase takes approximately 10 minutes.

- **Math Task**
  This session includes:

  - an introductory page with instructions and optional access to a practice environment similar to the main task, allowing participants to familiarize themselves with it (4 minutes);
  - a baseline phase featuring a relaxing video and water sounds (2 minutes);
  - the MAT (4 minutes);
  - a recovery phase with a similar video and water sounds (2 minutes).

  During this phase, three **during-task questionnaires** are collected: one at the end of the baseline, one after the task and the last at the end of the recovery phase.

- **Administration of Post-Math Task Questionnaire**
  The participant evaluates the perceived workload during the Math Task. This phase lasts approximately 3 minutes.

- **Game Task**
  This phase includes:
  - an introductory session with instructions on how to use the controller, a practice session in a simplified environment and new instructions about the competitive gameplay (approximately 4 minutes);
  - a baseline phase with a blank screen and relaxing music (1.5 minutes);
  - the VR game (3-5 minutes);
  - a recovery phase with a blank screen and relaxing music (1.5 minutes).

  As in the previous task, three **during-task questionnaires** are collected: one at the end of the baseline, one after the task and one at the end of the recovery phase.

- **Administration of Post-Game Task Questionnaires**
  The participant completes both a perceived workload questionnaire and a questionnaire on the VR environment. This phase lasts about 6 minutes.

- **Instrumentation Removal**
  The session concludes with the removal and deactivation of the monitoring equipment, which takes around 5 minutes.

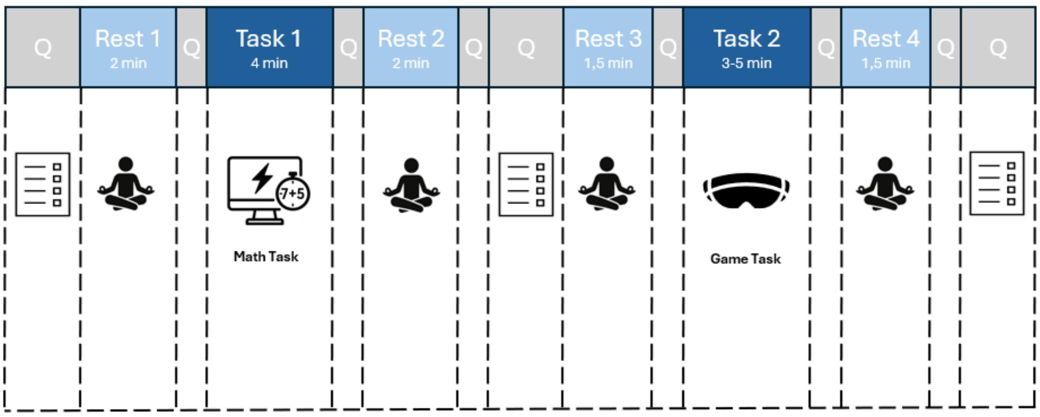

Figure 2.4: Overview of the experimental protocol. Detailed information on the questionnaires and the environments is provided in Sections 2.4 and 2.5, respectively.

2.4 Questionnaires

In addition to objective assessments of stress carried out through the measurement of physiological parameters, a subjective evaluation was also conducted by having the participants complete various questionnaires before, during and after the execution of the task. The questionnaires were administered entirely in Italian.

2.4.1 Pre-task Questionnaires

Before starting the test, participants completed three different questionnaires:

- **General questionnaire** [43]: originally called like this because it was used before performing a TSST, this questionnaire is designed to collect baseline information about participants and their activities over the 24 hours prior to the stress test. The aim is to monitor and document variables that may influence the physiological and psychological response to stress, and to potentially exclude individuals who may not provide reliable results (for example subjects under psychotropic medication).

- **PSS (Perceived Stress Scale)** [44]: a questionnaire used to measure the level of perceived stress experienced over the past 4 weeks. The selected version includes 10 items with responses based on a Likert scale, a rating tool used to assess attitudes, opinions or perceptions by asking subjects to express their degree of agreement or disagreement with a series of statements. In this case, the scale used ranges from 0 (never) to 4 (very often). The resulting score provides an indicative interpretation of stress level: scores between 0 and 13 indicate low stress, 14 to 26 moderate stress and 27 to 40 high stress. The scores for positively worded items (items 4, 5, 7 and 8) must be reversed (0 becomes 4, 1 becomes 3, 2 remains unchanged) during the scoring process.

- **STAI (State-Trait Anxiety Inventory)** [45]: a 40-item questionnaire used to assess anxiety in two distinct forms:

  - 20 items for state anxiety, which evaluate transitory anxiety, that means how the subject feels at the current moment;
  - 20 items for trait anxiety, which assess dispositional anxiety, like the subject's general tendency to experience anxiety.

  Responses are rated on a 4-point Likert scale, with some reverse-scored items that must be adjusted before computing the final score. Although there are no universally established thresholds, higher scores reflect greater levels of perceived anxiety, both as a state and as a trait. Analyzing both types of anxiety is crucial, as it provides a more comprehensive understanding of how anxiety manifests in the short term and how it tends to persist in individuals over the long term.

2.4.2 During-task Questionnaires

During the execution of the task, subjective stress assessments were also collected from the participants. Specifically, the **VAS (Visual Analogue Scale)** [46] was used, which displayed on-screen five buttons with values ranging from 1 (low or absent stress) to 5 (high stress).

Figure 2.5: Visual Analogue Scale. A value of 0 represents the absence of stress, while a value of 1 indicates high stress.

This questionnaire was administered multiple times (after each task and each rest) to obtain real-time evaluations of stress, thereby allowing for a more comprehensive mapping of the participant's stress levels.

2.4.3 Post-task Questionnaires

After the test, participants were asked to complete the final two questionnaires:

- **NASA-TLX (Task Load Index)** [47]: a questionnaire used to assess the perceived workload during the execution of a task. It comprises 6 items rated on a Likert scale from 0 to 4, which evaluate different aspects of workload: mental demand, physical demand, temporal demand, performance, effort and frustration. The analysis of these results helps determine whether the administered task achieved the intended effect.
- **Unity Environment Feedback**: a custom-developed questionnaire designed to assess whether the elements present in the game interface produced the desired effect on participants. Most of the responses were binary (YES/NO), although space was provided for suggestions and additional comments from the participants.

2.5 Environment

2.5.1 Math Task

This is a timed task that requires the rapid execution, using the PC keyboard, of basic mathematical operations (addition, subtraction, multiplication and division) involving two-digit numbers of varying difficulty. Each problem appears on a gray background, accompanied by a visible timer that imposes a **strict 4-second time limit** to provide an answer. Once the time expires, the problem is automatically replaced by the next one, regardless of whether a response was given. Only one attempt is allowed per question: if a wrong answer is submitted, it cannot be corrected and the participant must wait for the next question, which may cause frustration and a sense of helplessness.

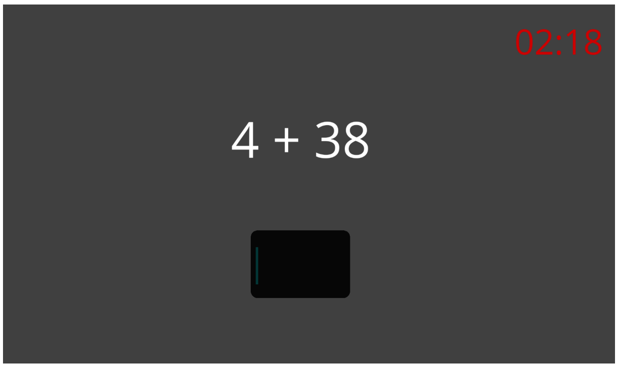

Figure 2.6: Math Task Environment.

Time pressure, the inability to correct mistakes and error notification increase stress levels, creating a sense of urgency and insecurity.

2.5.2 Game Task

This task involves a video game where the participant drives a car around a circuit, aiming to complete **three laps in the shortest possible time**. The track is full of obstacles to avoid and bonuses to collect, each lap is only validated if a minimum number of bonuses is collected and a maximum number of obstacles hit is not exceeded.

The task is deliberately designed with **multiple stress-inducing elements**:

- a constantly visible clock imposes a maximum limit of 5 minutes, increasing pressure;

- a real-time counter of penalties and bonuses is always visible on screen, allowing the participant to predict whether a lap will count, thus anticipating potential failure;
- every 90 seconds, manipulated feedback on ranking is displayed, always suggesting a worsening position (third-to-last, second-to-last, last), thereby generating a false sense of personal failure;
- the intense background music during gameplay increases tension, in sharp contrast to the relaxing music played only during the baseline and recovery phases;
- colliding with obstacles, difficult to avoid due to their random movements, triggers a loud explosion-like sound, designed as an unexpected auditory stimulus to provoke acute stress spikes;
- the track is surrounded by muddy areas which, while appearing to offer an escape route, significantly slow down the vehicle and contain neither bonuses nor obstacles. This forces participants to constantly weight risk and strategy, contributing to cognitive load.

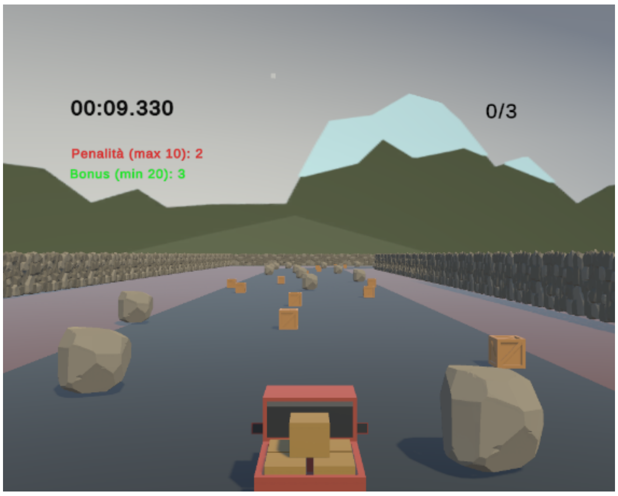

Figure 2.7: Game Task Environment.

These elements make the environment stressful due to its competitive nature and because participants must react to both visual and auditory stimuli, while continuously evaluating tactical decisions and personal performance.

2.6 Methods

In addition to the self-assessment questionnaires described in Section 3.1 and according to the literature, a set of physiological parameters were examined to assess variations in participants' stress levels:

- **ECG and thoracic acceleration**: these parameters were measured using a Polar H10 chest strap; ECG was analyzed as an objective indicator of stress level variations, thoracic acceleration was used to detect motion artifacts and enabling data cleaning.

- **EDA and arm acceleration**: both parameters were recorded using a sensor called Shimmer. EDA was used as an additional objective measure of physiological arousal related to stress, while arm acceleration was employed to detect motion artifacts and support data cleaning during preprocessing.

The signals were first segmented to analyze the different phases of the protocol separately and to exclude the periods of questionnaire completion. **Segmentation** was made possible by the timestamps of significant events, recorded during both the first trial (baseline + math task + parameter recovery) and the second trial (baseline + VR game + parameter recovery). Event timestamps were compared with those of the signals, and cuts were made at the closest indices, since the sampling frequency of the devices may not guarantee the presence of the exact timestamp corresponding to the event.

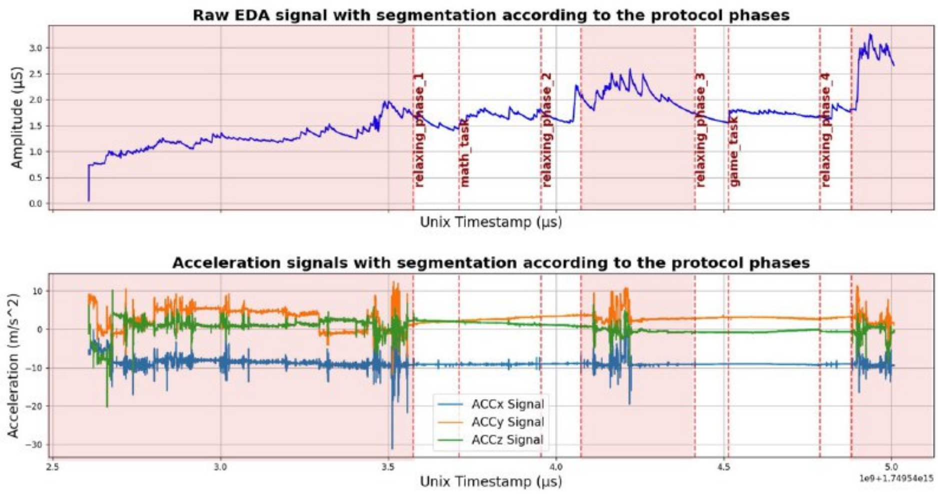

Figure 2.8: Segmentation of the EDA signal according to the protocol phases. The red segments correspond to the periods in which participants were completing the questionnaires, as indicated by a marked increase in arm acceleration due to writing. The same procedure was applied to the ECG signal.

Both signals were continuously monitored from 15 minutes before the first relaxation period until the completion of the last questionnaire. This initial interval allowed participants to reach a stable baseline of skin conductance, thus preventing increases in the SCL due to adaptation from being misinterpreted as stress indicators.

Once the signals corresponding to the different phases of the protocol were obtained, the analysis of ECG and EDA proceeded separately.

2.6.1 ECG Analysis

Two ECG recordings were excluded from the analysis sample because the Polar H10 chest strap lost connection before the end of the protocol. Since the data were incomplete, the entire recording was discarded.

Five main metrics were analyzed: heart rate, breath rate and three relative to the HRV, specifically the standard deviation of NN intervals (SDNN), the root mean square of successive differences (RMSSD) and the percentage of NN intervals differing by more than 50 ms (pNN50).

The Polar H10 chest strap applies internal filtering to the signal, removing both high-frequency noise and the DC component. Power spectral density analysis using Welch method, Hann window with a segment length of 1024 samples and 50% of overlap confirmed that the signal contained no components outside the ECG frequency band of interest. Therefore, **no additional filtering** was applied prior to analysis.

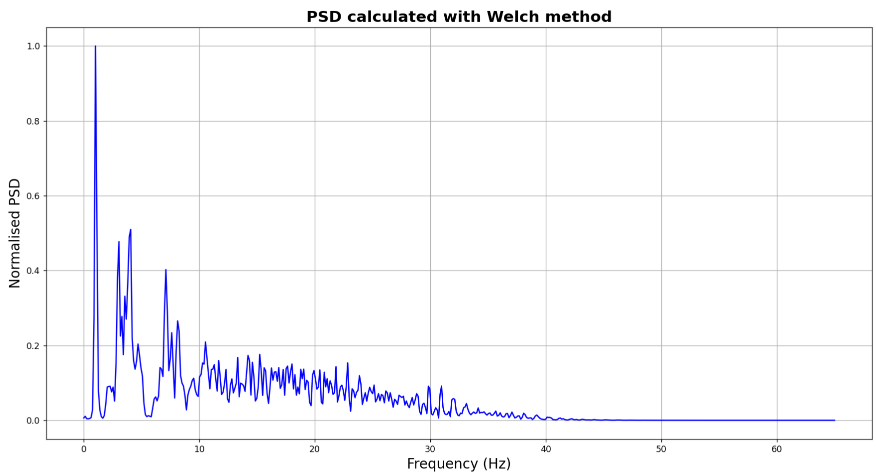

Figure 2.9: PSD estimated using Welch method on the zero-mean ECG signal.

The **heart rate (HR)** was calculated through a multi-step procedure. First, the R-peak indices were detected in the ECG signal, assuming a minimum amplitude of 40% of the maximum peak amplitude. Successive differences between R-peaks were then computed and converted into temporal differences dividing the index values by the sampling frequency, resulting in a sequence of RR intervals expressed in seconds. To ensure accuracy, incorrect peaks were excluded by retaining only those intervals no longer than 1.5 times the subject-specific mean of the RR intervals and never shorter than 400 ms (corresponding to a heart rate of 150 bpm), while invalid intervals were assigned a NaN value. Based on the valid RR intervals, the instantaneous heart rate in beats per minute (bpm) was obtained as the reciprocal of the RR interval in seconds multiplied by 60, as previously mentioned in the formula in Section 1.3.4. Finally, the average heart rate across the entire signal was calculated to enable comparisons across phases for inter-phase variability analysis.

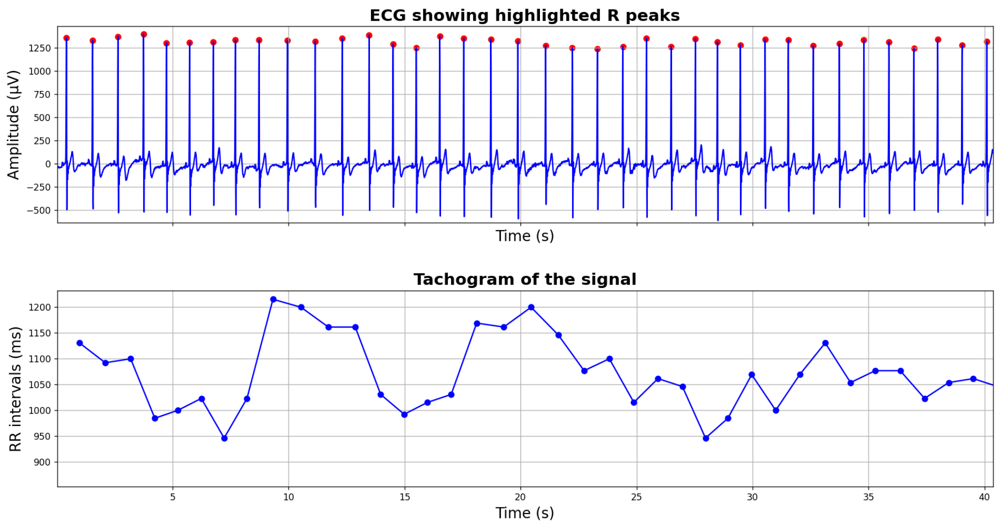

Figure 2.10: Identification of R peaks and the corresponding tachogram showing the time intervals between successive beats.

The **breath rate (BR)** was computed using variations in R-peaks, whose amplitude oscillations are known to reflect respiratory activity due to the displacement and rotation of the heart during the breathing cycle. First, the amplitudes and time instants associated with the previously detected R-peaks were extracted. Then, to ensure continuity, both amplitude and time vectors were linearly interpolated to replace any NaN values, after which the initial raw EDR signal was generated through cubic interpolation. This raw EDR signal was then band-pass filtered within the respiratory range of 0.15–0.5 Hz (corresponding to 9–30 cycles per minute) using a second-order Butterworth filter, which provides approximately

12 dB of attenuation per octave outside the passband.. Peaks were subsequently detected in the filtered EDR signal, and the average respiratory rate, expressed in cycles per minute, was finally computed by dividing the number of detected peaks by the duration of the filtered signal in minutes.

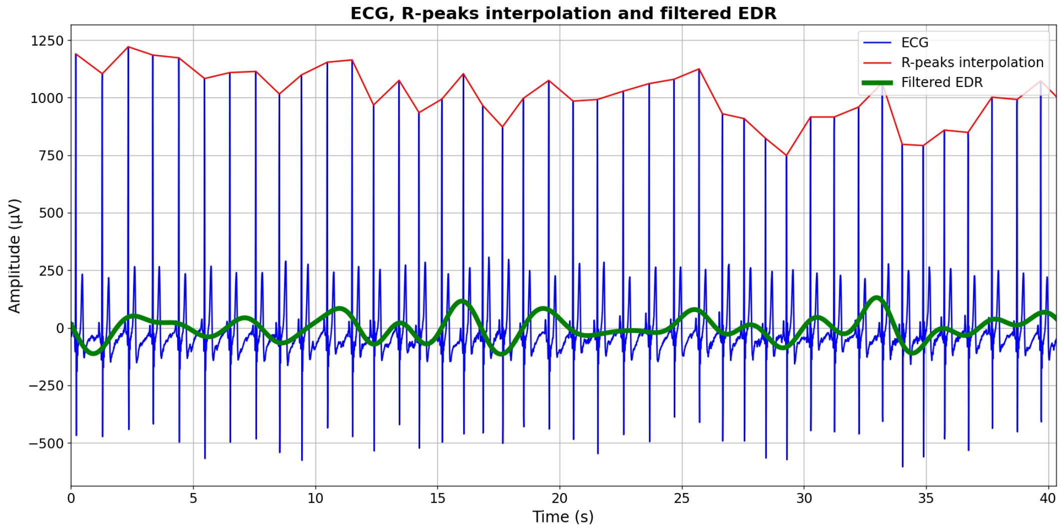

Figure 2.11: ECG, R-peak interpolation and filtered EDR. The blue trace shows the ECG signal, the red trace represents the interpolated R peaks and the green trace corresponds to the EDR signal filtered within the respiratory band.

The three metrics used for HRV analysis are easy to be implemented:

- **SDNN**: measures the standard deviation of NN intervals, thus RR intervals excluding ectopic beats (the ones not generated by the sinoatrial node). Since anomalous RR intervals were excluded in our analysis, this metric is fully applicable. It is represented by a single value summarizing the overall variability of RR intervals.
- **RMSSD**: similar to SDNN, it is particularly suitable for short recordings, such as in our study, and it is minimally influenced by respiration.
- **pNN50**: the differences between RR intervals excluding ectopic beats are calculated, the number of differences greater than 50 ms is counted, then divided by the total number of intervals and finally multiplied by 100 to obtain a percentage. Unlike the NN50 metric, which provides only the raw count, pNN50 allows comparison between signals of different durations.

2.6.2 EDA Analysis

One EDA recording, likely due to incorrect sensor placement, exhibited a very noisy signal. To avoid distorting the overall comparisons across participants, this trace was directly discarded.

The first step in EDA signal analysis actually consists of applying **z-score standardization**: for each sample of the original signal, the mean of the entire signal is subtracted and the result is divided by the signal's overall standard deviation. This procedure removes the initial sensor adaptation effect, making comparisons between different subjects more reliable.

The next step focused on the **elimination of motion artifacts** in the EDA recordings. EDA traces were observed together with wrist acceleration data, and segments with high acceleration were identified as containing motion artifacts. These affected segments were then removed and reconstructed using linear interpolation between their endpoints.

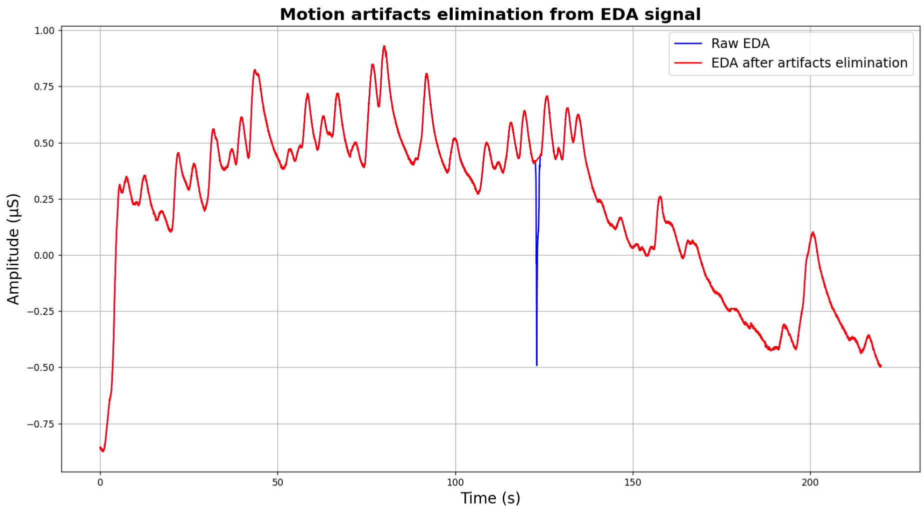

Figure 2.12: Motion artifact elimination from EDA signals, reconstructing the affected segments by linear interpolation between their endpoints.

The signal acquired from the Shimmer sensor was sampled at 64 Hz; however, it is a slowly varying signal with a frequency content reaching at most 1 Hz. Therefore, a sixth-order low-pass Butterworth **filter** with a **2 Hz** cutoff frequency was applied.

Subsequently, the **tonic and phasic components** were extracted using cvxEDA, an advanced Python algorithm for EDA signal decomposition. This allowed the separation of the tonic component, which reflects the general activity of the sympathetic nervous system, from the phasic component, which varies in response to specific stress-inducing events [48]. The output of this algorithm was already shown in Figure 1.11.

Finally, several features of the previously extracted components were calculated to provide an immediate representation of the signals. In particular, the **mean and median** of both components seemed especially useful, as they clearly distinguish between periods of relaxation and periods of stress.

Additionally, the **maximum value** and the **cumulative sum** of the phasic component were considered: the maximum allows determination of the peak heights, while the cumulative sum provides an indirect indication of the number of peaks present, since the appearance of a new peak can interrupt the growth of the previous one, blocking it from reaching its maximum value.

2.6.3 Regressor

The final step of data processing was the development of a **supervised machine learning model** capable of recognizing a state of stress during the execution of the VR Game task, with the goal of distinguishing stress from relaxation and potentially using it in the future as a real-time stress predictor.

Model Selection

The chosen model was **Random Forest**, an algorithm that combines the results of multiple decision trees to handle both classification and regression problems. In the case of classification, each tree selects a class and the most voted class becomes the final prediction; in the case of regression, the mean of the values predicted by all trees is calculated. In this study, **regression** was chosen, since stress does not only appear as "present" or "absent", but also includes intermediate states [49]. A schematic of the algorithm's operation is shown in the Figure 2.13.

The name Random Forest is self-explanatory: it is called forest because it applies the principle of **ensemble learning**, a mechanism that uses many simple and weak models (decision trees) to obtain a more stable and accurate prediction; the word random refers to the fact that each tree is trained on a random subset of the data. In particular, the **bootstrap sampling** technique is used: to build each tree, samples are randomly selected from the original data, allowing the repetition of some examples and the exclusion of others.

Another descriptive characteristic of Random Forest is **feature subsampling**: in addition to selecting random samples of data, each tree can also select random subsets of features, with the goal of reducing correlation among the trees and increasing the diversity of the forest.

Thanks to its approach, this algorithm has two main advantages. First, it **reduces overfitting** because each tree sees different subsets of data, allowing the overall model to generalize better. Second, it **improves stability**, as combining predictions from different trees makes the model less affected by noise and outliers.

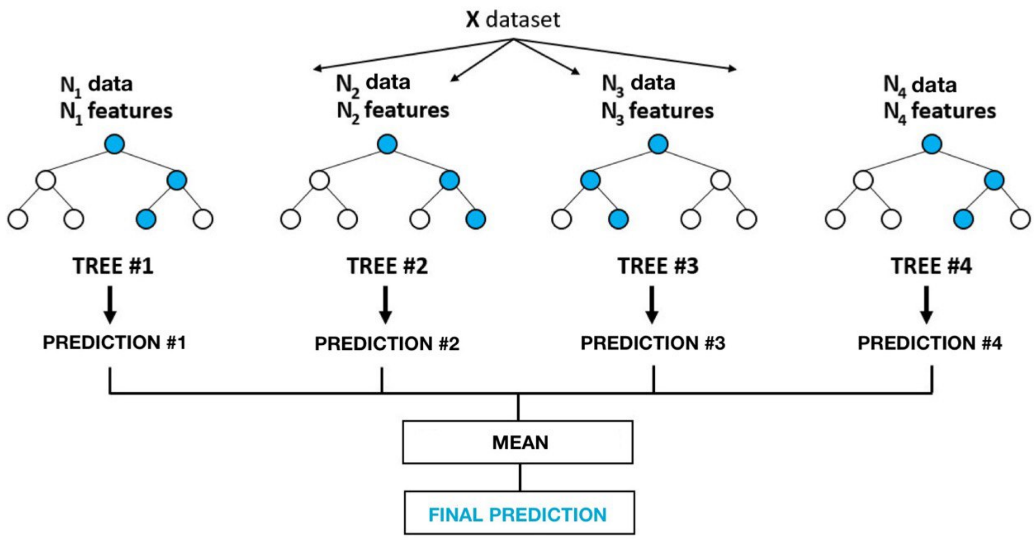

Figure 2.13: Working principle of the Random Forest algorithm for regression.

This algorithm was chosen because it does not require precise parameter setting and represents a good balance between accuracy and robustness. Additionally, it provides a **reliable estimate of feature importance**, helping to identify which variables have the greatest influence on predictions.

Model Parameters Definition

The algorithm used comes from the Scikit-Learn library in Python and allows the customization of several parameters:

- Number of trees: indicates how many independent models will provide a separate prediction;
- Maximum number of features per tree: determines how many variables each tree considers when making a split;
- Maximum depth of the trees: defines how many times a tree can continue splitting its nodes;
- Minimum number of samples required to split a node: prevents splitting nodes that are too small;
- Minimum number of samples per leaf: ensures that terminal nodes contain at least a certain number of samples;
- Use of bootstrap: there is a boolean parameter indicating whether to use bootstrap sampling. In this library, bootstrap sampling creates for each tree a training sample of the **same size as the original dataset**, but different from it because it is randomly selected.

For our case, the chosen parameters were:

- **400 trees**: a number sufficient to ensure stability and accuracy, but limited enough to avoid overly similar trees (and thus overfitting) and to prevent slowing down training, with the goal of future real-time use;
- **10 features per tree**: the typical value recommended for regressors in order to increase tree diversity is about one third of the total features (in this case 5). However, since the available features were few, a higler number was chosen to preserve prediction accuracy;
- **Unlimited tree depth**: each tree will continue to split until the minimum number of samples required for a split is reached;
- **Minimum 2 samples to split** a node: higher values would reduce tree depth and help prevent overfitting, but with limited data and few features overly constraining the model could decrease its accuracy;
- **Minimum 2 samples per leaf**: similarly, higher values would reduce overfitting but could penalize accuracy with limited data;
- **Bootstrap sampling enabled**: chosen because it ensures the model's robustness.

ECG Features Selection

The features provided to the algorithm were extracted using the Python library NeuroKit2.

Since the future goal is to implement a real-time regressor, features were selected to be valid even over short time windows, obtained from the different waves of the ECG signal. **Eight primary** and independent features were identified, as shown in Table 2.1, along with **eight derived** features representing the temporal variation of each primary feature, shown in Table 2.2. The number of features was deliberately increased, even including correlated ones, because the Random Forest algorithm is not affected by feature correlation and can rank features by importance. Therefore, adding more features may potentially improve the model's accuracy.

FEATURE	DESCRIPTION
RR Interval	interval between two consecutive R peaks
QT Interval	interval between the the Q peak and the T peak within a single waveform
QTc Interval	ratio between the QT interval and the square root of the RR interval
RT Ratio	ratio between the R peak and the successive T peak
RP Ratio	ratio between the R peak and the successive P peak
BR Ratio	ratio between two consecutive peaks of the EDR signal
BR Width	duration of a respiratory cycle
BR pp	ratio between a peak and the following valley of the EDR signal

Table 2.1: Chosen features for Regressor Training: Main Features.

FEATURE	DESCRIPTION
RR comp	ratio of the RR Interval in the third respiratory cycle to those in the second and first respiratory cycles
QT comp	ratio of the QT Interval in the third respiratory cycle to those in the second and first respiratory cycles
QTc comp	ratio of the QTc Interval in the third respiratory cycle to those in the second and first respiratory cycles
RT comp	ratio of the RT Ratio in the third respiratory cycle to those in the second and first respiratory cycles
RP comp	ratio of the RP Ratio in the third respiratory cycle to those in the second and first respiratory cycles
BRr comp	ratio of the BR Ratio in the third respiratory cycle to those in the second and first respiratory cycles
BRw comp	ratio of the BR Width in the third respiratory cycle to those in the second and first respiratory cycles
BRp comp	ratio of the BR pp in the third respiratory cycle to those in the second and first respiratory cycles

Table 2.2: Chosen features for Regressor Training: Derived Features.

Before being passed to the machine learning algorithm, the extracted features were normalized using the **StandardScaler method** from the Scikit-Learn library. This method performs a z-score normalization on the baseline, setting the mean to 0 and the standard deviation to 1, and subsequently scales all extracted features relative to the baseline ones.

Model Implementation

The ultimate goal is to implement a real-time regressor. Therefore, it is desirable for the regressor to output a value over a relatively short time window, which has been defined as **three respiratory cycles**. The choice of a **variable-length window** is driven by the need to have the same number of features for each segment, and three cycles were selected as the minimum length that allows the extraction of the various ECG waves.

First, the EDR signal was extracted from the ECG signal and automatically filtered by the default NeuroKit2 function between 0.1 and 0.4 Hz. Then, the indices of the respiratory peaks were identified and used to divide both the ECG and EDR signals into windows, each covering three respiratory cycles. Each segment was analyzed and subsequently the start and the end of the window were shifted by one respiratory peak, resulting in an overlap of two cycles between consecutive analyses.

For each window, three values per feature were obtained by calculating the mean of each feature across each respiratory cycle within the window.

The model was evaluated using **5-fold cross-validation** method. The data were split into five segments, with the model trained on four parts and tested on the remaining one. This process was repeated five times so that each segment served as a test set once. To measure performance, the **coefficient of determination (R^2)** was calculated for each fold, and the average of the five R^2 values was used as an indicator of the model's fit. Finally, the model with the highest average R^2 was selected.

Model Training and Testing

The regressor was **subject-specific**: each user had their own model. It was trained using data from the baseline condition prior to the math task, labeled as 0 (complete relaxation), and data from the math task, labeled as 1 (acute stress). The model was first trained using all features, then retrained using the four best features.

Once trained, the model was applied to the signal from the Game Task as well as the segment corresponding to the immediately preceding relaxation period, in order to assess whether the regressor could identify a stress state and distinguish it from a moment of relaxation. The same training structure was repeated: each window covered three respiratory cycles, overlapped by two cycles and produced a single regressor output.

Since training the regressor on a task so different from the one used for testing might seem inappropriate, the **reverse procedure** was also carried out: the regressor was trained on the Game Task and tested on the Math Task. The goal of this approach was to verify whether the mean stress levels detected in the Game Task using a regressor trained on the Math Task were similar to those detected in the Math Task using a regressor trained on the Game Task. If the results were comparable, this would indicate that the two stressful tasks elicit similar physiological responses, so a regressor trained on the Math Task could be applied to the VR Game.

2.6.4 Statistical Analysis

Each extracted variable was subjected to statistical analysis, which was divided into three phases:

- **2-way ANOVA** to investigate differences among distributions;
- **paired tests** to examine differences between specific pairs of distributions;
- **correction factor** to adjust multiple comparisons.

In both the 2-way ANOVA and paired tests, non-parametric methods were chosen. This decision was motivated by the small sample sizes (9 subjects for EDA variables and 18 for ECG variables), which would have made normality tests such as Shapiro–Wilk unreliable. **Non-parametric tests**, which do not require assumptions about the data distribution, were therefore considered a more robust alternative even if they are less powerful than parametric tests.

The **Scheirer-Ray-Hare (SRH) test**, excluding interaction terms, was applied as a non-parametric alternative to a 2-way ANOVA. Subjects and conditions were included as the two factors, and significant differences between conditions were investigated. Then the **Wilcoxon signed-rank test** was conducted for all pairs of variables for which the SRH test revealed a global significant difference.

The risk of a Type I error, which means that a result may appear significant simply because many tests are conducted, increases in case of paired tests. To address this, a **Bonferroni correction factor** was applied, which consists of multiplying each p-value by the number of comparisons. This procedure substantially lowers the p-values; however, when a difference is found to be statistically significant, it can be reasonably concluded that it reflects a true effect.

Chapter 3

Results

3.1 Questionnaires

As described in in Section 3.1, in addition to measuring physiological parameters, participants were asked to complete questionnaires to subjectively assess their stress levels, in order to compare their perception with the collected physiological data.

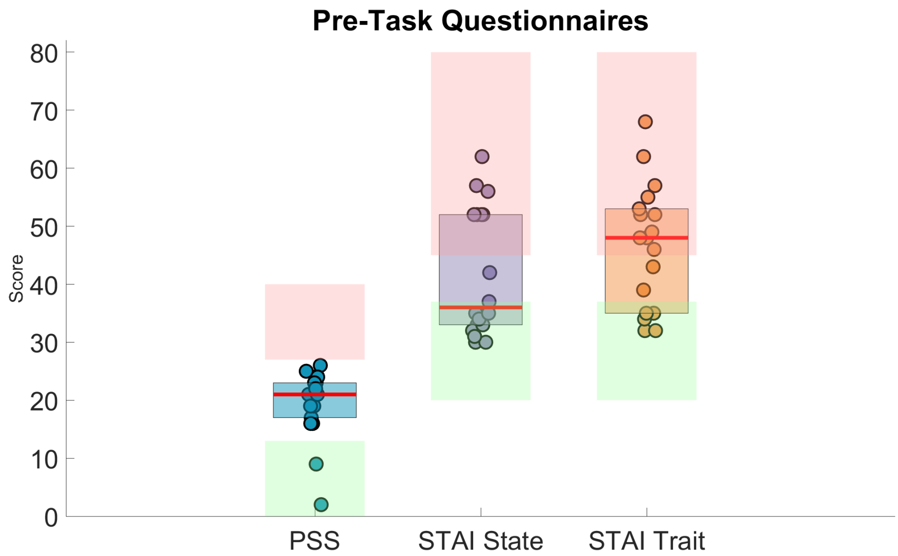

Figure 3.1: Pre-task questionnaire scores: they indicates levels of stress or anxiety. Green indicates low range, white indicates moderate range and red indicates high range. Values were taken from [1].

Figure 3.1 shows the results of questionnaires administered at the beginning of the protocol: the PSS to measure perceived stress over the past weeks and the STAI to assess state and trait anxiety. For clarity, the STAI evaluations are presented in two separate parts. Score ranges highlighted in green correspond to low range (both for stress and anxiety), those in red indicate high range and the central zone represents moderate range, where the majority of responses is expected. In the literature there are no universally valid ranges; in this study the ranges used were drawn from [1]. For the PSS, all participants scored outside the high-stress range. In contrast, both in the State and Trait subscales of the STAI, some participants fell within the high-anxiety range.

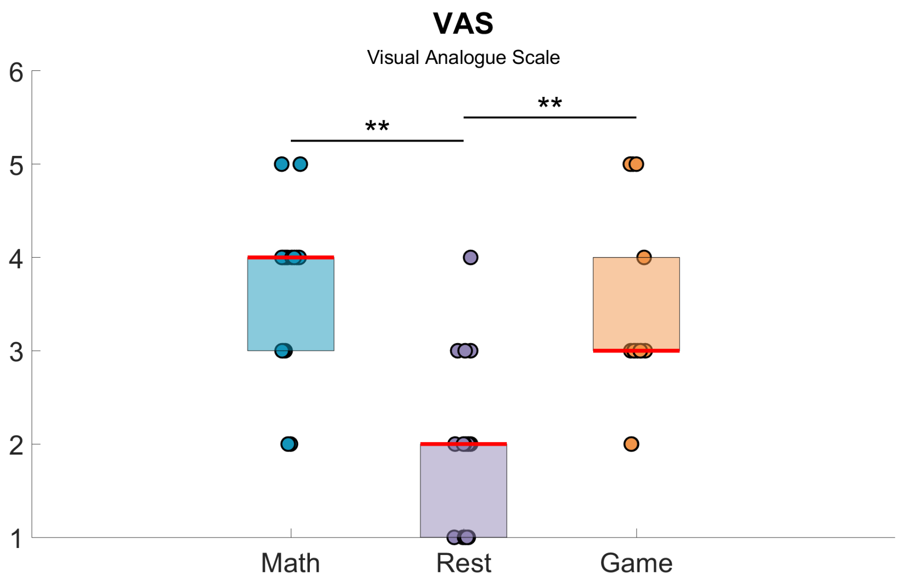

Figure 3.2: During-task questionnaire scores: self-evaluation after Math Task, Rest and Game Task. * indicates p between 0.05 and 0.01, ** between 0.01 and 0.001 and *** $<$ 0.001.

Figure 3.2 shows the VAS scores of all participants immediately after the two stress-inducing tasks, the mathematical task and the VR game, compared to the relaxation period preceding the VR task, which was chosen as the rest reference. This relaxation period was preferred over the one preceding the mathematical task because the latter was characterized by higher variability: many participants became restless before the end of that rest phase, thus introducing noise into the data.

Both the VR game and the mathematical task show a significant difference compared to the rest condition, but no significant difference between each other.

This absence of significant differences between the VAS scores of the mathematical task and the VR game was also confirmed by the NASA-TLX questionnaire, administered at the end of the respective recovery phases following each task and shown in Figure 3.3. In both questionnaires, VAS and NASA-TLX, the median stress ratings for the VR game were sligthly lower than those for the mathematical task, although the difference was not statistically significant.

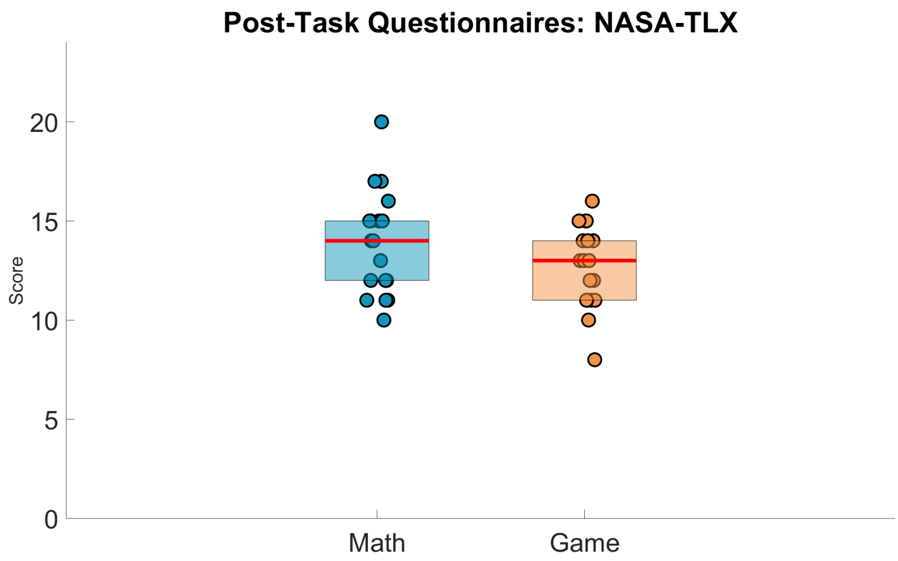

Figure 3.3: Post-task questionnaire scores: Math Task vs Game Task.

The protocol also included the completion of a questionnaire aimed at investigating the participants' subjective perception of the virtual environment, with its questions presented in Appendix A. One section focused on the sense of immersion, since isolation from the surrounding environment was considered a major strength: it ensured that the stress experienced by participants was exclusively due to the task elements and not to external factors. A total of 95% of participants described the environment as immersive, indicating that this objective was achieved.

Subsequently, participants were asked to evaluate each task element as stressful or not with binary responses (Yes/No). An additional section allowed them to report any other elements perceived as stressful. Figure 3.4 shows the percentage of participants who considered each element stressful: in blue are the elements intentionally designed to induce stress, while in red are those reported by participants as stressful.

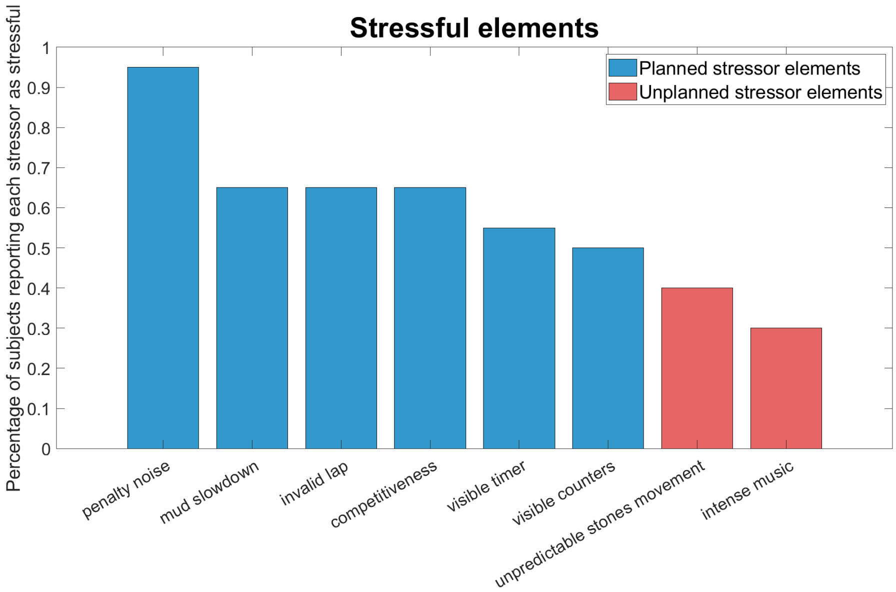

Figure 3.4: Stressful elements in Game Task.

One critical issue identified was the participants' possible lack of experience with headsets. To address this, a preliminary familiarization trial with the virtual environment was introduced. As shown in Figure 3.5, 55% of participants had never used an headset before but 85% found the familiarization trial useful, thus confirming the effectiveness of this measure.

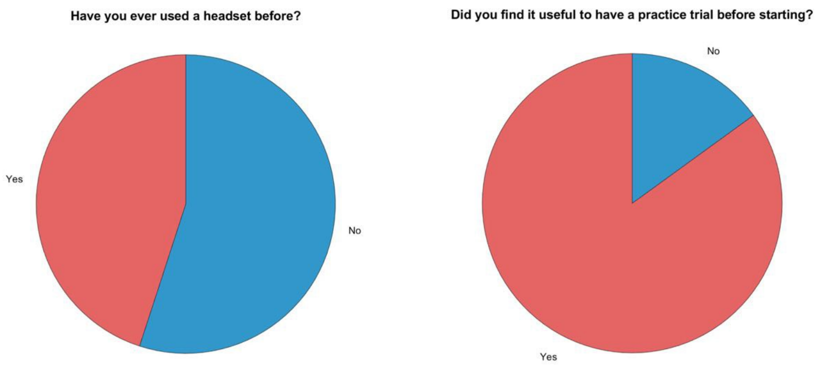

Figure 3.5: Pie charts showing prior use of a headset vs the perceived usefulness of the familiarization session.

Another anticipated issue was related to participants' competitiveness: less competitive individuals might have perceived the task as less stressful. With the aim of identifying strategies to make the game more stressful for all participants, participants were asked whether playing against a friend would have increased their stress levels. As shown in Figure 3.6, the percentage of participants who considered competitiveness stressful increased from 65% to 85% in the case of playing against a friend.

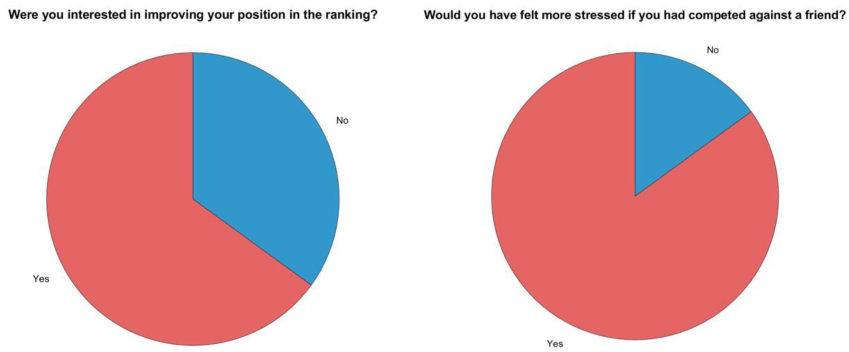

Figure 3.6: Pie charts showing average sense of competition vs sense of competition when playing with a friend.

Finally, the last section of the questionnaire collected feedback and suggestions on how to make the game more stressful.

3.2 ECG Analysis

As highlighted in Section 2.6.1, the ECG analysis was carried out using two basic metrics, Heart Rate and Breath Rate, and three HRV indices, such as SDNN, RMSSD and pNN50. **All ECG metrics were normalized** by dividing the mean value of each segment for each metric by its corresponding mean value during the baseline prior to the Math task (this condition was therefore excluded from comparisons, as it was always equal to 1). This normalization allowed for greater comparability across participants and focused the analysis on relative variations only.

Figure 3.7 shows that the mean heart rate during both the Math task and the VR game was significantly higher compared to the baseline prior the VR Game. However, no significant differences were observed between the two stress-inducing tasks.

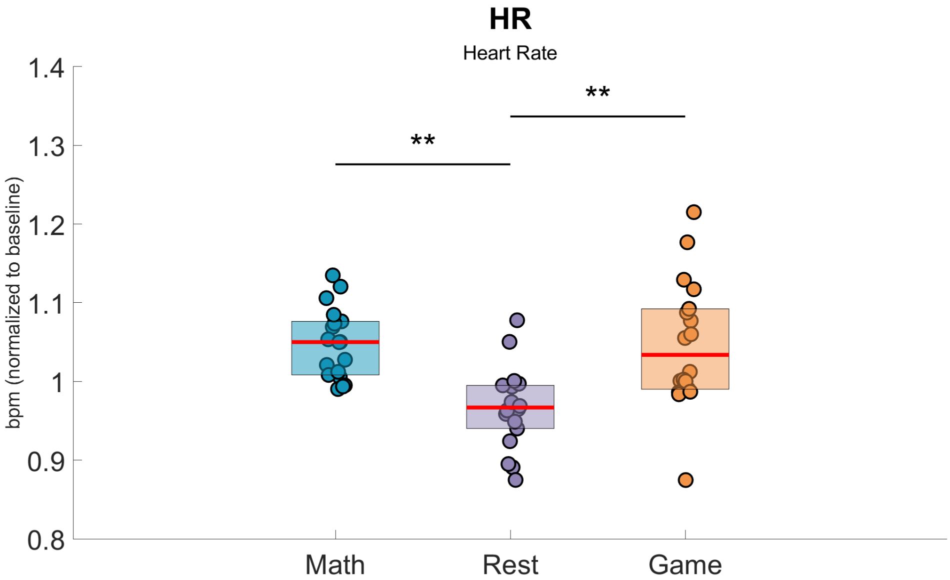

Figure 3.7: Mean HR during Math Task, Rest and Game Task. In the figure, only three distributions are shown because all data were normalized relative to the baseline measured before the Math task. * indicates p between 0.05 and 0.01, ** between 0.01 and 0.001 and *** $<$ 0.001.

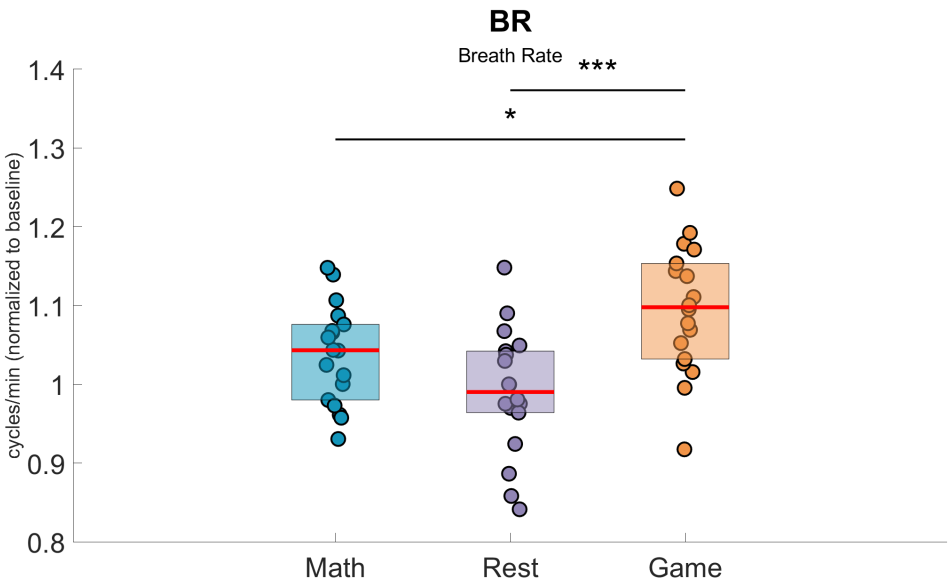

Figure 3.8: Mean BR during Math Task, Rest and Game Task. In the figure, only three distributions are shown because all data were normalized relative to the baseline measured before the Math task. * indicates p between 0.05 and 0.01, ** between 0.01 and 0.001 and *** $<$ 0.001.

Figure 3.8 refers to the breath rate, obtained from EDR. Regarding the comparisons between the stressful moments and the relaxation period, only the comparison with the Game task showed significant differences; however, the comparison with the Math task produced a p-value slightly above the significance threshold. A significant difference was also found between the Game and Math tasks, with the VR game showing higher values, which is consistent with the expected trend.

The heart rate variability metrics, shown in Figures 3.9, 3.10 and 3.11, present less meaningful results: signals display lower variability during stressful tasks compared to relaxation, which aligns with expectations, but the results are not statistically significant.

As noted in Section 2.6.4, the statistical analysis consisted of three steps: the SRH test to investigate global differences, the Wilcoxon signed-rank test to analyze paired differences and the Bonferroni correction to avoid Type I errors. The only metric that passed the SRH test, and subsequently evaluated using the Wilcoxon test, was RMSSD. However, its significance was very marginal and after applying the correction factor it exceeded the significance threshold.

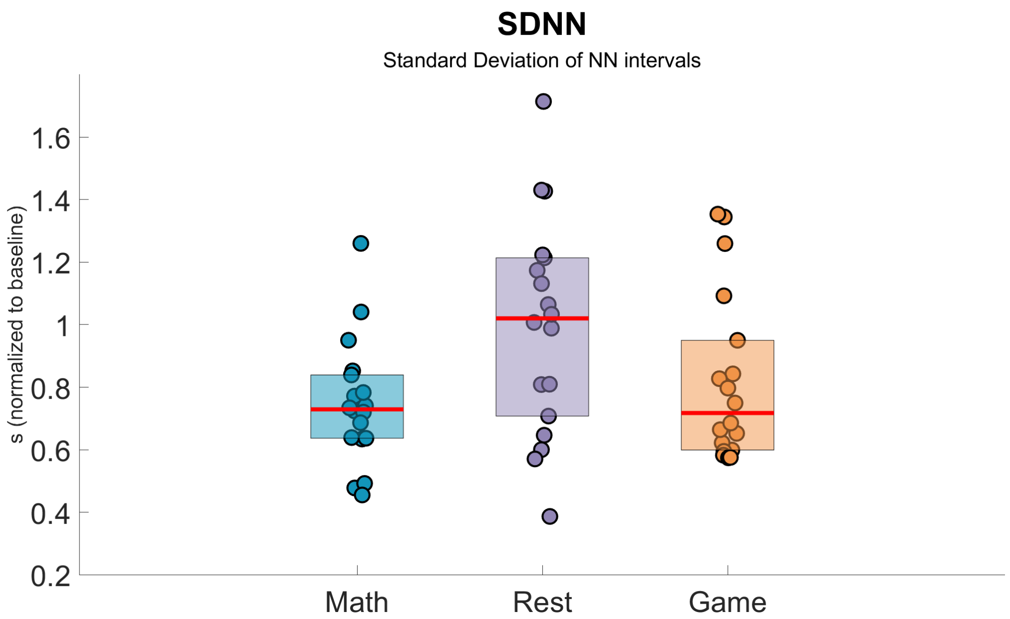

Figure 3.9: SDNN during Math Task, Rest and Game Task. * indicates p between 0.05 and 0.01, ** between 0.01 and 0.001 and *** < 0.001.

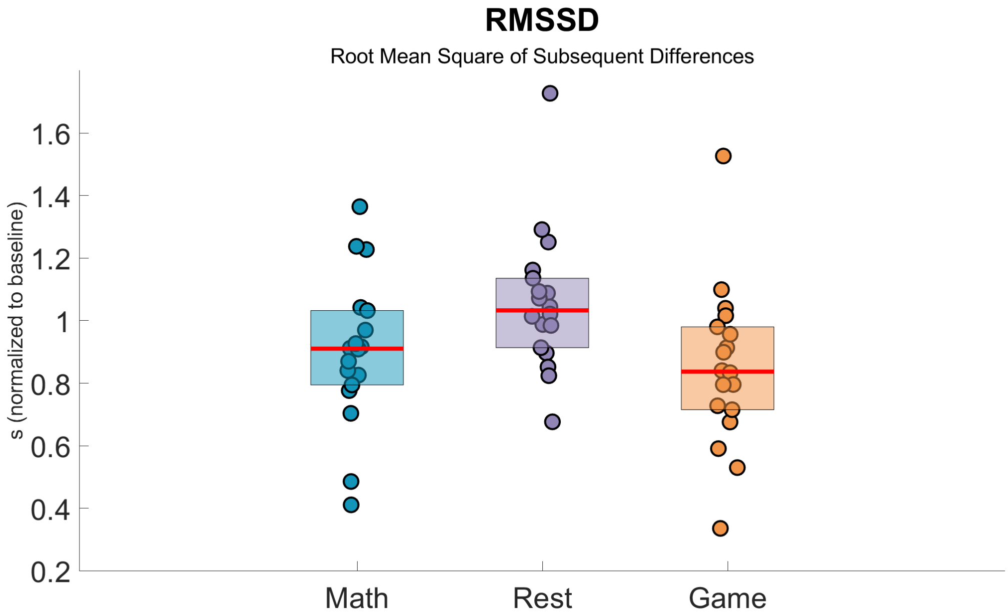

Figure 3.10: RMSSD during Math Task, Rest and Game Task. * indicates p between 0.05 and 0.01, ** between 0.01 and 0.001 and *** < 0.001.

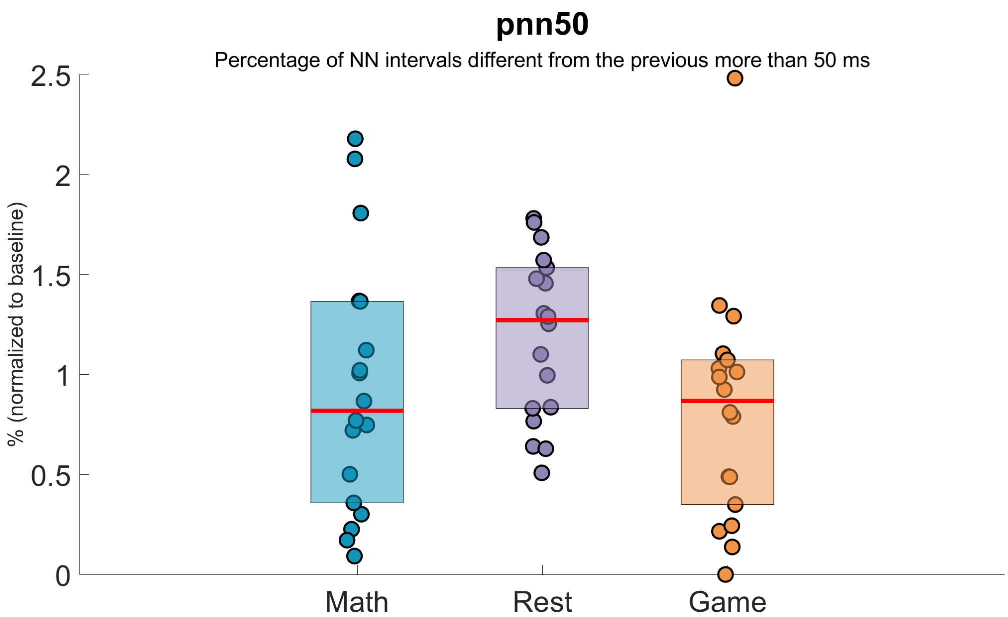

Figure 3.11: pnn50 during Math Task, Rest and Game Task. * indicates p between 0.05 and 0.01, ** between 0.01 and 0.001 and *** < 0.001.

3.3 EDA Analysis

As already explained in Section 1.4.3, the tonic and phasic components of the EDA signal describe different phenomena. The tonic component, characterized by slow variations that occur only over the course of minutes, reflects the overall state of sympathetic activation without capturing specific moments of activation. In contrast, the phasic component is faster and shows peaks within 1–3 seconds of a stressful event, making it more suitable for detecting stress-related peaks. Therefore, we expect that during stressful trials the phasic component will provide the most relevant results, although differences may also be observed in the tonic component.

For the tonic component, only the mean and the median are considered relevant features, as this component does not exhibit peaks. As shown in Figures 3.12 and 3.13, the VR game yields higher values compared to the Math task and the baseline. Nevertheless, none of these differences reach statistical significance.

These two features were computed also for the phasic component. Figures 3.14 and 3.15 show lower data dispersion compared to the tonic component, especially during the Game Task. The most reliable variable is the median, as it is less affected by outliers, and shows significant differences between baseline and both stressful tasks, while no differences emerge between the tasks themselves.

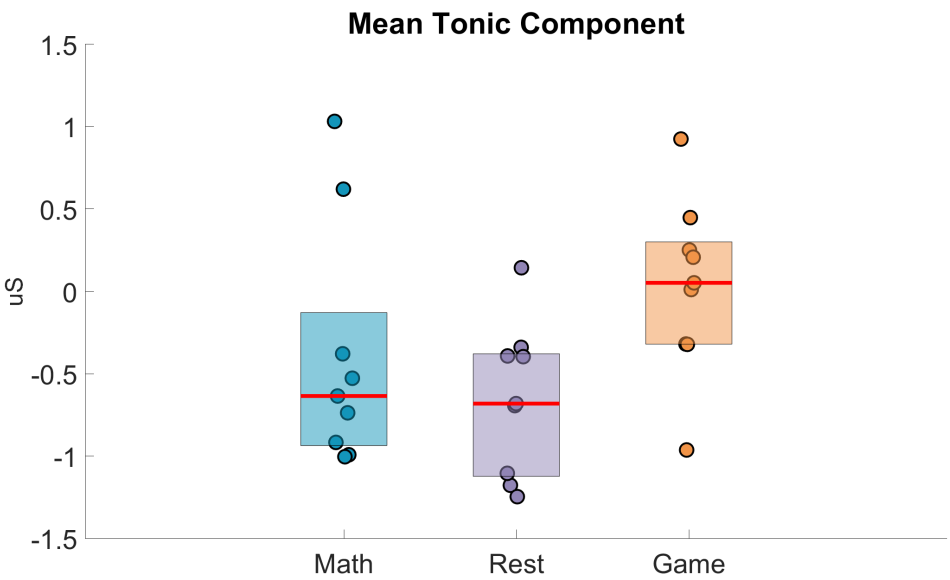

Figure 3.12: Mean of Tonic Component during Math, Rest and Game Tasks. * indicates p between 0.05 and 0.01, ** between 0.01 and 0.001 and *** < 0.001.

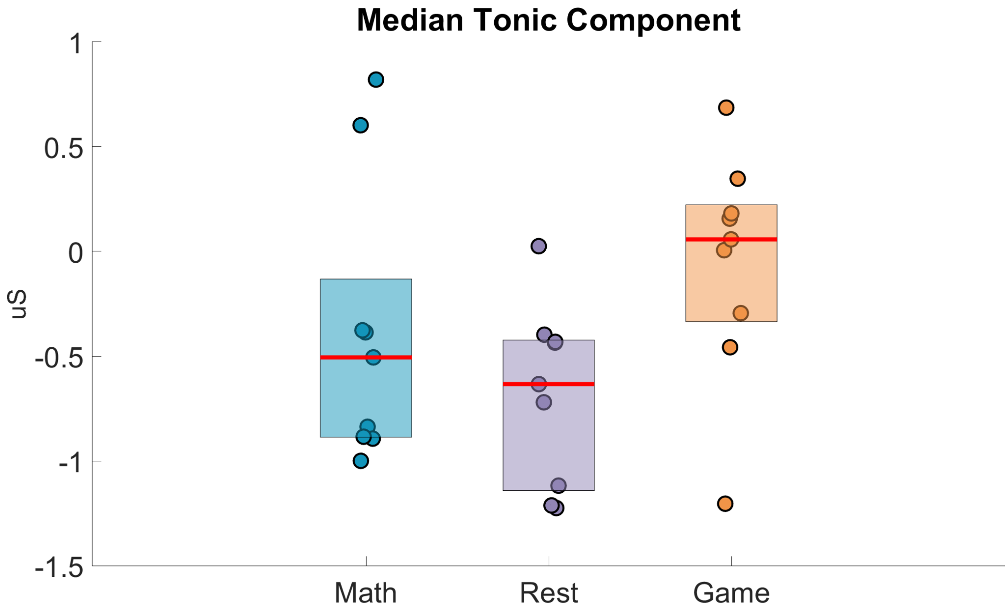

Figure 3.13: Median of Tonic Component during Math, Rest and Game Tasks.

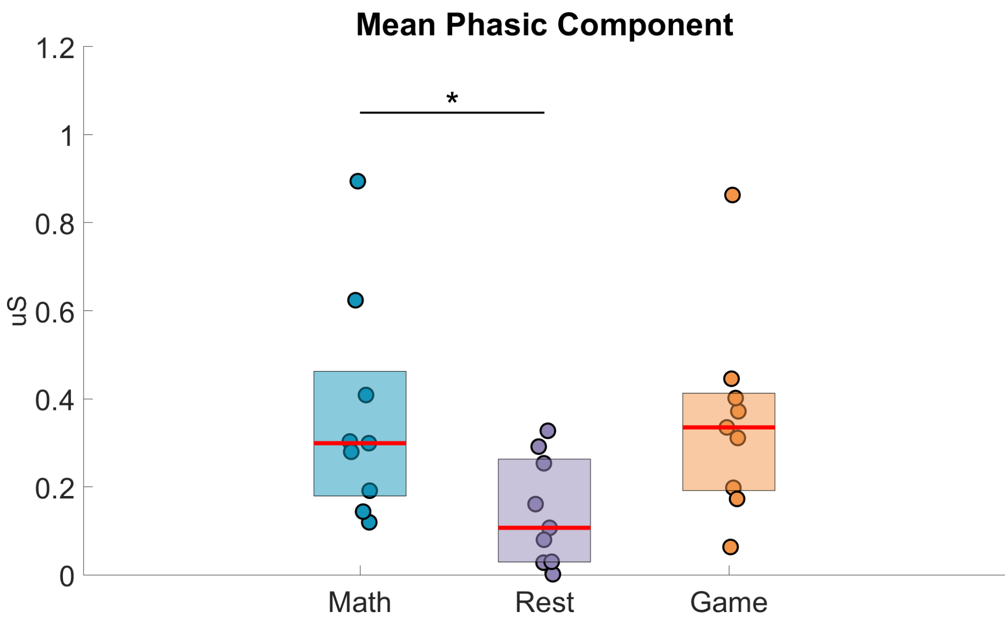

Figure 3.14: Mean of Phasic Component during Math, Rest and Game Tasks. * indicates p between 0.05 and 0.01, ** between 0.01 and 0.001 and *** < 0.001.

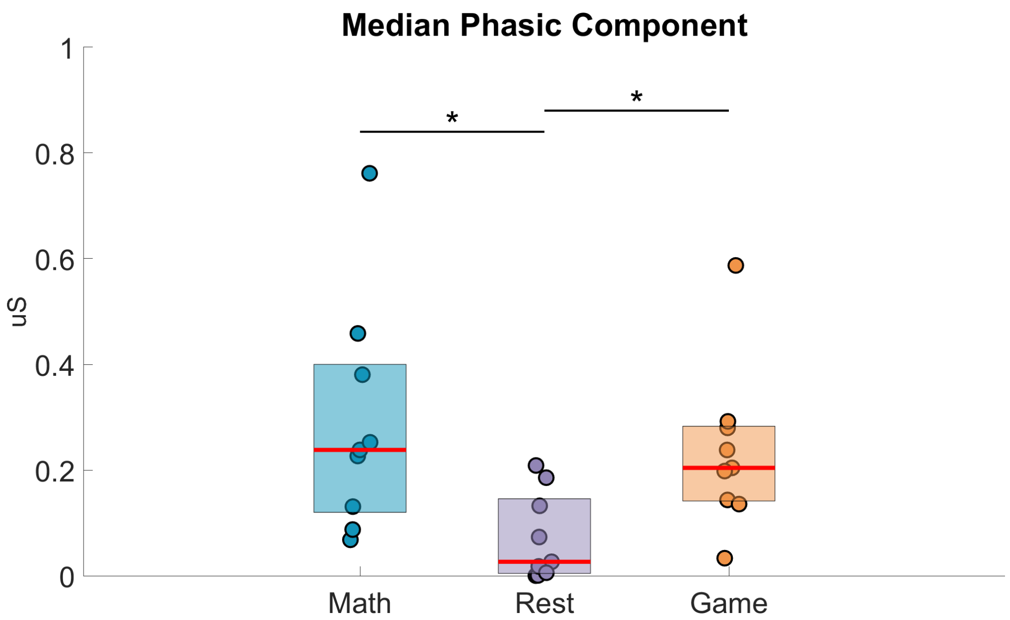

Figure 3.15: Median of Phasic Component during Math, Rest and Game Tasks. * indicates p between 0.05 and 0.01, ** between 0.01 and 0.001 and *** < 0.001.

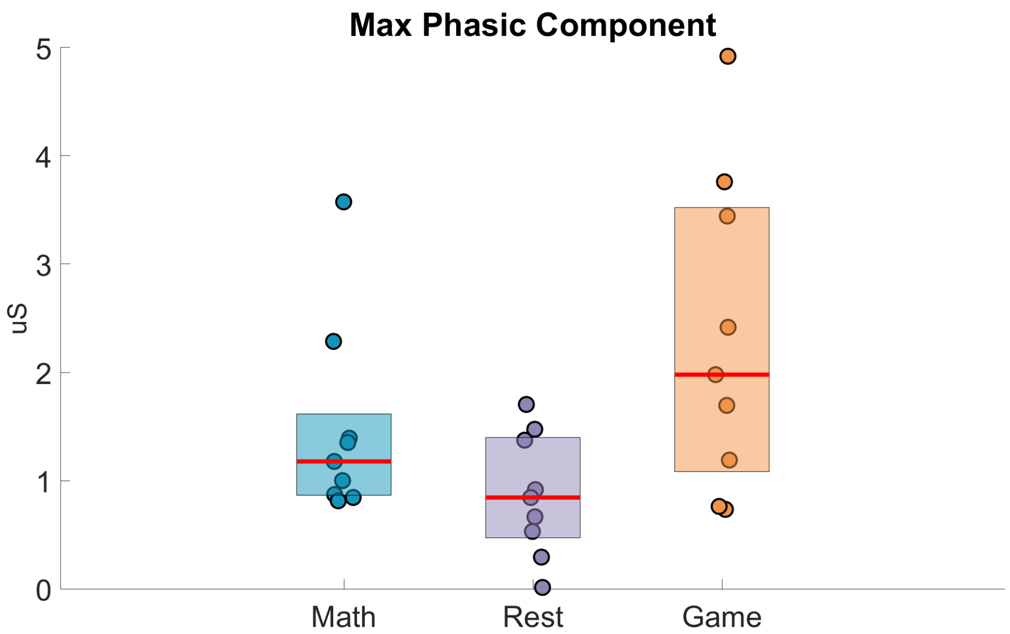

Figure 3.16: Maximum value of Phasic Component during Math, Rest and Game Tasks. * indicates p between 0.05 and 0.01, ** between 0.01 and 0.001 and *** < 0.001.

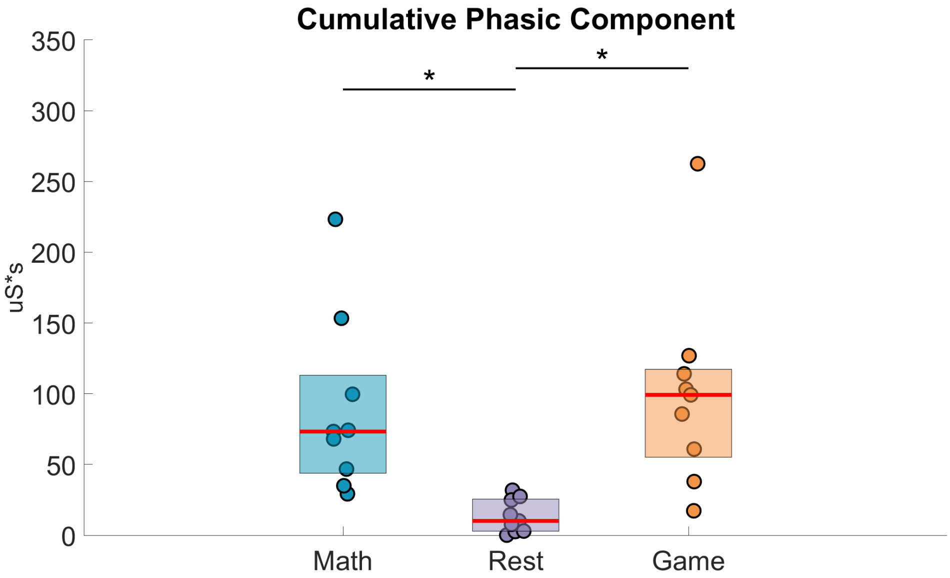

Figure 3.17: Cumulative of Phasic Component during Math, Rest and Game Tasks. * indicates p between 0.05 and 0.01, ** between 0.01 and 0.001 and *** < 0.001.

The maximum value and the cumulative area, describing peaks amplitude and frequency and showed in Figures 3.17 and 3.16, were computed only for the phasic component. Only the second one shows differences between relaxation and stressful tasks: this indicates that peaks are more frequent during the stressful tasks, but not necessarily higher in magnitude

3.4 Regressor

As already mentioned in Section 2.6.3, one of the distinguishing characteristics of the Random Forest model is its ability to produce an accurate and robust predictor without the need for careful manual feature selection. This is because the model autonomously evaluates feature importance and identifies the most informative variables to construct a reliable and accurate representation.

In Figure 3.18, the frequency with which the various features were selected among the four to train the final model is shown, with a comparison between training on the VR Game and on the Math Task. It can be observed that the RR interval was by far the most frequently selected in both trainings, followed by the respiratory cycle width. This is consistent with expectations: heart rate and breath rate are the physiological characteristics that are most rapidly and significantly affected by stress, and are therefore considered good indicators of it. Most of the derived features, on the other hand, were found to be of little relevance.

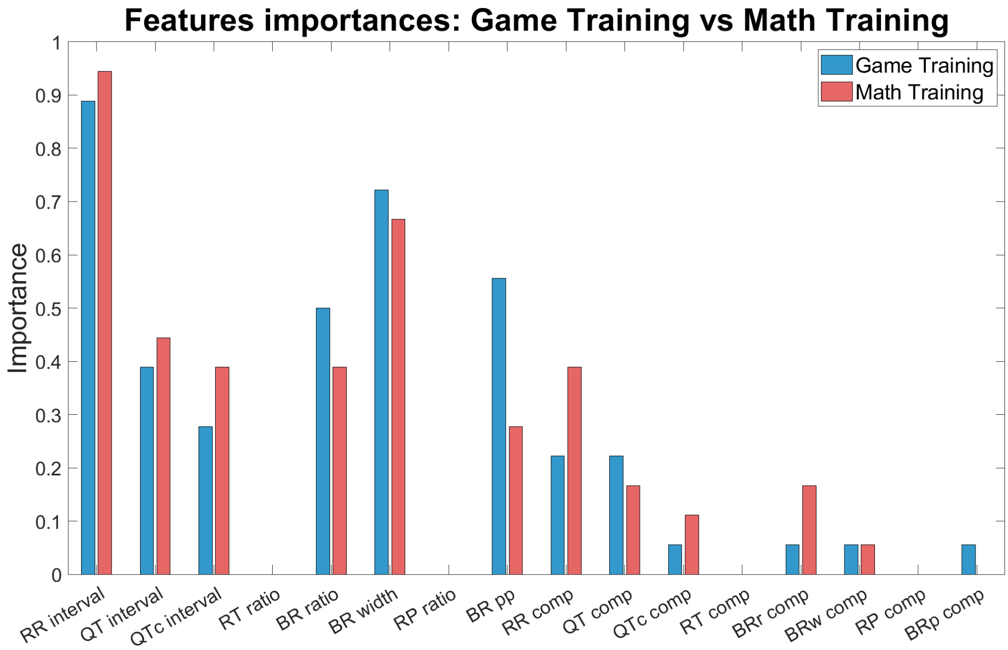

Figure 3.18: Comparison of features selected by the regressor trained on the Game Task vs those selected by the regressor trained on the Math Task.

Subsequently, a comparison is presented between the output of the regressor trained on the Math Task and that of the regressor trained on the Game Task.

Figure 3.19 shows this comparison for a single subject: the output of the regressor during the relaxing task is shown in purple, with its mean highlighted in red, while the output during the stressful task is displayed in light blue and its mean in green.

It can be observed that, both with the regressor trained on the Math Task (top) and with the one trained on the Game Task (bottom), the difference between the mean outputs in the relaxing and stressful conditions is highly pronounced. It can also be noticed that both the means of the relaxing condition and the Math Task obtained with the Game Training are slightly higher than those obtained with the Math Training, however without marked differences.

Figure 3.20 extends this observation to all subjects: in every case, each regressor clearly distinguishes its corresponding stressful task from the relaxing condition with very high statistical significance. By contrast, no evident differences emerge between the mean outputs of the two stressful tasks, thereby confirming the possibility of applying a regressor trained on the Math Task to the Game Task.

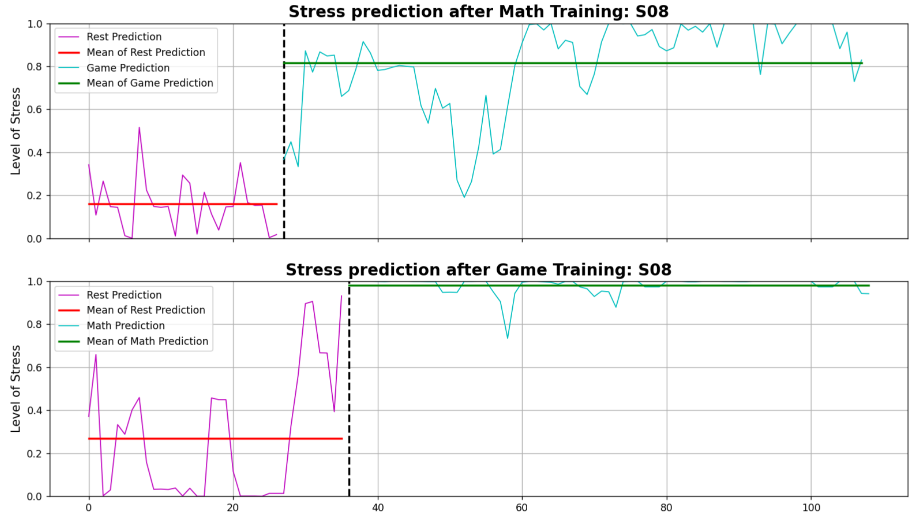

Figure 3.19: Example of a regressor trained on the Math Task to estimate stress levels during the Game Task compared to the preceding relaxation period (top), and a regressor trained on the Game Task to estimate stress levels during the Math Task compared to the preceding relaxation period (bottom).

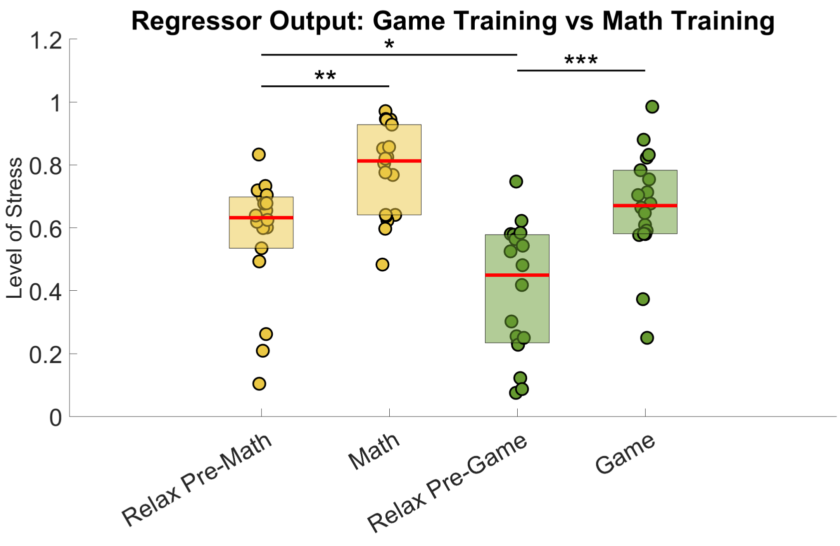

Figure 3.20: Average output of the regressor trained on the Game Task and tested on the Math Task with its corresponding rest period (left), and average output of the regressor trained on the Math Task and tested on the Game Task with its corresponding rest period (right). * indicates p between 0.05 and 0.01, ** between 0.01 and 0.001 and *** < 0.001.

Chapter 4

Discussion and Conclusion

The pre-task questionnaires revealed that some participants showed high levels of both trait and state anxiety: the former indicating a general predisposition to experience this emotion, the latter reflecting elevated values specifically at the time of the protocol execution. These tests were administered to identify individuals who might exhibit smaller differences between relaxation and task phases, because they will probably show higher baseline stress levels. However, these participants were not excluded a priori, since the sample size was already limited and further reduction would have compromised the stability of significance estimates. Instead, a different approach was adopted: the features extracted from these participants' data were closely monitored, and they would only have been excluded if all their features consistently fell outside the expected range. As none of them turned out to be outliers, they were retained in the sample.

The subjective perception of stress levels, assessed through VAS and NASA-TLX, revealed no significant differences between the MAT and the VR game. This suggests that the VR game, despite not requiring specific skills (unlike the mathematical task), was still perceived as stressful. This is likely due to the fact that stress can arise from multiple factors, such as the desire to perform well, the need to manage unexpected stimuli or the drive to excel in a competition, which depend primarily on the participant's own approach to the task.

The feedback questionnaire on the VR game structure was particularly interesting: 95% of participants reported feeling immersed in the environment, which was the main strength of using VR in this study. Thus, it can be assumed that the stress experienced by participants derived exclusively from the game elements, ensuring greater experimental control over the environment, which is an essential aspect in experimental research.

The analysis of the metrics extracted from the ECG signal revealed clear and significant differences between the relaxation phase and both stress-inducing tasks for the basic metrics, namely heart rate and breath rate. In contrast, differences in heart rate variability (HRV) metrics were minimal and not statistically significant. This can be easily explained: heart rate and breath rate are highly sensitive and change rapidly, so they readily reflect differences even over short recordings. HRV metrics, on the other hand, require medium-to-long observation intervals to detect meaningful changes and are truly meaningful only when compared across signals of equal length. Considering that the relaxation phase used as a baseline lasted only 1.5 minutes while the stress-inducing tasks lasted 4–5 minutes, HRV metrics in this context are of limited value. Moreover, participants had already completed a previous task and knew another was coming, so the recorded values relative to the relaxation phase may not represent a genuine baseline.

The analysis of the metrics extracted from the EDA signal revealed clear and significant differences in those related to the phasic component, whereas differences in the tonic component were not statistically significant. This lack of significance, despite the visibly noticeable differences in the medians of the two variables between the relaxation phase and the VR game, may be explained by a combination of factors, like a limited number of participants and highly variable data. It is possible that a larger sample would be sufficient to provide more robust results. However, the tonic component is more susceptible to accumulation effect, which z-score normalization can only partially mitigate, so it is also possible that the variables were higher in VR game due to this effect.

In conclusion, the VR environment designed appears suitable for the controlled investigation of stress, as it induces measurable psycho-physiological changes and is perceived as immersive. However, there are several modifications that should be considered for **future developments**.

HRV-related metrics may be exluded or at least studied on longer segments. It would be better to measure baseline periods of similar duration to the stress tasks, at least 4–5 minutes, providing sufficient time to observe these parameters. Additionally, recording baseline data before explaining the protocol could help ensure that measurements are not influenced by any prior stress.

Concerning EDA, in order to obtain a reliable measure of the tonic component that is not affected by summation effects, it would be advisable to randomize the order of task administration across participants to cancel out the average accumulation effect. However, the intention to employ a real-time regressor in future studies precludes random task sequencing, since a stress-inducing task must always be administered first to allow regressor training. Therefore, the tonic component might simply need to be discarded.

The second part of the questionnaire about VR environment was devoted to collecting suggestions for making the environment more stressful for future developments. Participants' feedback highlighted several potential improvements:

- introducing a dual-race mode, where participants compete simultaneously with a friend, or displaying the continuous progress of the best performer (as a semi-transparent car running ahead), to enhance competitiveness;
- occasionally inverting the colors of game elements associated with penalties and bonuses or adding random sounds unrelated to any specific trigger, to increase the demand for concentration and thus cognitive load;
- implementing progressively harder challenges as the game advances, similar to a multi-level game, in order to reduce the learning effect;
- making both progress and regressions more visible, since on-screen text displayed during the task was largely ignored and therefore did not effectively induce stress.

This approach is valuable because, even if stress is a highly subjective psychophysiological response, by introducing additional stress-inducing elements the likelihood that all participants perceive the environment as stressful increases. Creating a more consistently stressful environment facilitates the identification of the factors that actually trigger stress and subsequently enables the exploration of strategies for its mitigation. In particular, the regressor developed in this work was intended as a preliminary step toward modulation.

A potential future direction could be the creation of a game element that modulates stress, such as a vibration or visual stimulus delivered at a specific frequency, aimed at periodically minimizing the output generated by the regressor. This idea was inspired by studies on binaural beats for stress reduction [50][51][52][53], which use auditory waves to generate a beat at a specific frequency. Through a phenomenon known as **brainwave entrainment** [54], certain populations of neurons can temporarily synchronize with that frequency, thereby modifying EEG oscillations. This effect can be targeted within relaxation-related frequency bands, such as the theta band (4–8 Hz), to induce stress reduction [55][56][57]. In our case, the environment already includes multiple auditory stimulations, which may limit the effectiveness of this approach. However, it is hypothesized that replacing sounds with flashing lights or tactile stimuli could produce a similar effect.

Appendix A

Unity Environment Feedback Questionnaire

This section presents the set of questions that were given to participants regarding the VR environment. It is not a standard questionnaire. The questions are written in Italian.

Sezione 1: Istruzioni e Obiettivi

1. Hai letto attentamente le istruzioni prima di iniziare a giocare? SÌ / NO

2. Ti è stato chiaro come pilotare l'auto? SÌ / NO

3. Ti è stato chiaro che era una gara di velocità? SÌ / NO

4. Ti è stato chiaro che bisognava evitare le pietre? SÌ / NO

5. Ti è stato chiaro che bisognava raccogliere le scatole? SÌ / NO

6. Ti è stato chiaro che serviva un numero minimo di bonus e un numero massimo di penalità per far contare il giro? SÌ / NO

7. Ti è stato chiaro il limite di tempo per completare il task? SÌ / NO

8. Hai trovato utile poter fare una prova prima di iniziare? SÌ / NO

9. Considerazioni: ____

   ____

Sezione 2: Feedback Visivi e Sonori

1. Ti sei sentito “immerso” nell’ambiente? SÌ / NO
2. Hai notato il contatore delle penalità aumentare quando colpivi una pietra? SÌ / NO
3. Hai notato il contatore dei bonus aumentare quando raccoglievi una scatola? SÌ / NO
4. Il frastuono provocato dall’errore ti ha trasmesso una sensazione di stress? SÌ / NO
5. Il fatto che, finendo nel fango, progredissi più lentamente ti ha trasmesso una sensazione di stress? SÌ / NO
6. Il fatto che, se non rispettavi il numero minimo di bonus e il numero massimo di penalità, il giro non contasse ti ha trasmesso una sensazione di stress? SÌ / NO
7. Cosa ti ha fatto stressare maggiormente? ______________________________

   ____
8. Considerazioni: ____

   ____

Sezione 3: Competizione

1. Hai mai usato un visore prima? SÌ / NO
2. Sei abituato ad utilizzare joysticks? SÌ / NO
3. Hai visto comparire nella schermata di gioco i messaggi relativi alla tua posizione in classifica? SÌ / NO

   Eri interessato a migliorare la tua posizione in classifica? SÌ / NO

   Ti sei stressato cercando di vincere? SÌ / NO
4. Ti saresti stressato di più se avessi gareggiato contro un amico? SÌ / NO
5. Ti sei divertito? SÌ / NO
6. Considerazioni: ____

   ____

Sezione 4: Altro

1. Hai qualche consiglio da dare su come aumentare il livello di stress durante il gioco? ______________________________

   ____

2. Considerazioni: ______________________________

   ____

Bibliography

[1] Matteo Raggi, Lee J. Moore, and Luca Mesin. «Inducing and Assessing Acute Mental Stress in Controlled Conditions: Topical Review and Guidelines for Effective Experimental Protocols». In: *IEEE Access* (2025).

[2] Frederic H. Martini, Judi L. Nath, and Edwin F. Bartholomew. *Fondamenti di Anatomia e Fisiologia.* 2nd. Edises, 2020.

[3] M. Le Fevre, J. Matheny, and G. S. Kolt. «Eustress, distress, and interpretation in occupational stress». In: *Journal of Managerial Psychology* 18.7 (2003).

[4] Mohamad Awada, Burcin Becerik Gerber, Gale M. Lucas, and Shawn C. Roll. «Stress appraisal in the workplace and its associations with productivity and mood: Insights from a multimodal machine learning analysis». In: *PLoS One* 19.1 (2024).

[5] Julie Bienertova-Vasku, Peter Lenart, and Martin Scheringer. «Eustress and Distress: Neither Good Nor Bad, but Rather the Same?» In: *BioEssays* 42.4 (2020).

[6] Mohamad Awada, Burcin Becerik-Gerber, Gale M. Lucas, Shawn C. Roll, and Ruying Liu. «A New Perspective on Stress Detection: An Automated Approach for Detecting Eustress and Distress». In: *IEEE Transactions on Affective Computing* 15 (2023).

[7] James A. Russell and Albert Mehrabian. «Evidence for a three-factor theory of emotions». In: *Journal of Research in Personality* 11.3 (1977).

[8] George M. Slavich. «More than a feeling: A unified view of stress measurement for translational neuroscience». In: *Psychoneuroendocrinology* 98 (2018).

[9] Maria Gardani, Daniel R. R. Bradford, Kirsten Russell, Stephanie Allan, Louise Beattie, Jason G. Ellis, and Umair Akram. «A systematic review and meta-analysis of poor sleep, insomnia symptoms and stress in undergraduate students». In: *Sleep Medicine Reviews* 61 (2022).

[10] Archana S. Nagaraja, Nouara C. Sadaoui, Piotr L. Dorniak, Susan K. Lutgendorf, and Anil K. Sood. «Snapshot: Stress and disease». In: *Cell Metabolism* 23.2 (2016).

[11] Hannes Ernst, Matthieu Scherpf, Sebastian Pannasch, Jens R. Helmert, Hagen Malberg, and Martin Schmidt. «Assessment of the human response to acute mental stress–An overview and a multimodal study». In: *PLoS ONE* 18.11 (2023).

[12] Jonathan D. Quick, Rebecca S. Horn, and James Campbell Quick. «Health Consequences of Stress». In: *Job Stress*. Routledge, 2014.

[13] Madhu Kalia. «Assessing the economic impact of stress—the modern day hidden epidemic». In: *Metabolism: Clinical and Experimental* 51.6 (2002).

[14] B. McCarthy, S. Datta, G. Sesa-Ashton, et al. «Differential control of sympathetic outflow to muscle and skin during physical and cognitive stressors». In: *Clinical Autonomic Research* 34 (2024).

[15] Sean M. Smith and Wylie W. Vale. «The role of the hypothalamic-pituitary-adrenal axis in neuroendocrine responses to stress». In: *Dialogues in Clinical Neuroscience* 8.4 (2006).

[16] Ibraheem Rehman and Afzal Rehman. *Anatomy, Thorax, Heart*. Treasure Island (FL): StatPearls Publishing, 2025.

[17] Mario Merone, Paolo Soda, Mario Sansone, and Carlo Sansone. «ECG databases for biometric systems: A systematic review». In: *Expert Systems with Applications* 67 (2017).

[18] S. Schliemann and P. Elsner. «Anatomy and Physiology of the Skin». In: *Skin: Aging and Disorders*. Berlin, Heidelberg: Springer, 2014.

[19] Y. Sawada and K. Hashimoto. «Distribution and function of eccrine and apocrine sweat glands in human skin». In: *Journal of Dermatological Science* 58.3 (2010).

[20] Wolfgang Boucsein. *Electrodermal Activity*. New York, NY: Springer Science & Business Media, 2012.

[21] Risheka Walls et al. «Virtual reality as an engaging and enjoyable method for delivering emergency clinical simulation training: a prospective, interventional study of medical undergraduates». In: *BMC Medicine* 22 (2024).

[22] Marzieh Alsadat Minouei, Athar Omid, Alireza Mirzaie, Hamid Mahdavifard, and Alireza Rahimi. «Effectiveness of virtual reality on medical students' academic achievement in anatomy: systematic review». In: *BMC Medical Education* 24 (2024).

[23] Belén Rubio Ballester, Jens Nirme, Irene Camacho, Esther Duarte, Susana Rodríguez, Ampar Cuxart, Armin Duff, and Paul F.M.J. Verschure. «Domiciliary VR-Based Therapy for Functional Recovery and Cortical Reorganization: Randomized Controlled Trial in Participants at the Chronic Stage Post Stroke». In: *JMIR Serious Games* 5.3 (2017).

[24] Mohsen Soleimani, Marjan Ghazisaeedi, and Soroush Heydari. «The efficacy of virtual reality for upper limb rehabilitation in stroke patients: a systematic review and meta-analysis». In: *BMC Medical Informatics and Decision Making* 24 (2024).

[25] Moritz Queisner and Karl Eisenträger. «Surgical planning in virtual reality: a systematic review». In: *Journal of Medical Imaging* 11.6 (2024).

[26] Florentine Huettl, Patrick Saalfeld, Christian Hansen, Bernhard Preim, Alicia Poplawski, Werner Kneist, Hauke Lang, and Tobias Huber. «Virtual reality and 3D printing improve preoperative visualization of 3D liver reconstructions—results from a preclinical comparison of presentation modalities and user's preference». In: *Quantitative Imaging in Medicine and Surgery* 9.13 (2021).

[27] Maria V. Sanchez-Vives and Mel Slater. «From presence to consciousness through virtual reality». In: *Nature Reviews Neuroscience* 6 (2005).

[28] Giuseppe Riva, Brenda K. Wiederhold, and Fabrizia Mantovani. «Neuroscience of virtual reality: From virtual exposure to embodied medicine». In: *Cyberpsychology, Behavior, and Social Networking* 22.1 (2019).

[29] Nan Bu. «Mental Stress Tests with Autonomic Responses Induced in Virtual Reality Environment». In: *Proceedings of the IEEE 2nd Global Conference on Life Sciences and Technologies (LifeTech)*. National Institute of Technology (NIT), Kumamoto College. Kumamoto, Japan: IEEE, 2020.

[30] R. Vaitheeshwari, Shih-Ching Yeh, Eric Hsiao-Kuang Wu, Jyun-Yu Chen, and Chia-Ru Chung. «Recognition Based on Multiphysiological Data in High-Pressure Driving VR Scene». In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28.5 (2020).

[31] Richard S. Lazarus and Susan Folkman. *Stress, appraisal, and coping*. New York: Springer Publishing Company, 1984.

[32] Dan Archer and Anthony Steed. «Optimizing Performance through Stress and Induction Levels in Virtual Reality Using Autonomic Responses». In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. *Dan Archer, †Anthony Steed. IEEE. Atlanta, GA, USA: IEEE, 2020.

[33] Justin Asbee, Kimberly Kelly, Timothy McMahan, and Thomas D. Parsons. «Factor analysis of the virtual reality Stroop task». In: *Virtual Reality* 24.4 (2020).

[34] Celia Kessassi. «Modeling the Induction of Psychosocial Stress in Virtual Reality Simulations». In: *Proceedings of the Conference (specific conference name if available)*. Nantes, France, 2020.

[35] M. A. Fallon, M. M. E. Riem, L. E. Kunst, W. J. Kop, and N. Kupper. «Multi-modal responses to the Virtual Reality Trier Social Stress Test: A comparison with standard interpersonal and control conditions». In: *Psychoneuroendocrinology* 113 (2020). Business School, University of Mannheim; Tilburg University.

[36] Emily C. Helminen, Melissa L. Morton, Qiu Wang, and Joshua C. Felver. «A meta-analysis of cortisol reactivity to the Trier Social Stress Test in virtual environments». In: *Psychoneuroendocrinology* 115 (2020).

[37] Emily C. Helminen, Melissa L. Morton, Qiu Wang, and Joshua C. Felver. «Stress Reactivity to the Trier Social Stress Test in Traditional and Virtual Environments: A Meta-Analytic Comparison». In: *Psychoneuroendocrinology* 115 (2020).

[38] Max Planck Institute for Human Cognitive and Brain Sciences. *Induction of Social Stress through the Trier Social Stress Test (IMAGE)*.

[39] William Standard, Daniel Zielasko, Bernadette von Dawans, Gregor Domes, and Benjamin Weyers. «Open TSST VR: Trier Social Stress Test in Virtual Reality». In: *GI VR/AR Workshop* 2020 (2020).

[40] Kjell N van Paridon, Matthew A Timmis, Charlotte M Nevison, and Matt Bristow. «The anticipatory stress response to sport competition; a systematic review with meta-analysis of cortisol reactivity». In: *BMJ Open Sport & Exercise Medicine* 3.1 (2017).

[41] Renato A Souza, Oscar A B Beltran, Diana M Zapata, Elisângela Silva, Wagner Z Freitas, Rubens V Junior, Fabiano F da Silva, and Wonder P Higino. «Heart rate variability, salivary cortisol and competitive state anxiety responses during pre-competition and pre-training moments». In: *Biology of Sport* 36.1 (2018).

[42] Songfa Zhong, Idan Shalev, David Koh, Richard P. Ebstein, and Soo Hong Chew. «Competitiveness and Stress». In: *International Economic Review* 59 (2018).

[43] NF Narvaez Linares, Valérie Charron, Allison J. Ouimet, Patrick R. Labelle, and Helene Plamondon. «A systematic review of the Trier Social Stress Test methodology: Issues in promoting study comparison and replicable research». In: *Neurobiology of Stress* 13 (2020).

[44] Karla Gambetta-Tessini, Rodrigo Mariño, Mike Morgan, Wendell Evans, and Vivienne Anderson. «Stress and health-promoting attributes in Australian, New Zealand, and Chilean dental students». In: *Journal of Dental Education* 77.6 (2013).

[45] Peter J. Bieling, Martin M. Antony, and Richard P. Swinson. «The State–Trait Anxiety Inventory, Trait version: Structure and content re-examined». In: *Behaviour Research and Therapy* 36.7-8 (1998).

[46] F. X. Lesage, S. Berjot, and F. Deschamps. «Clinical stress assessment using a visual analogue scale». In: *Occupational Medicine* 62.8 (2012).

[47] Atyanti Dyah Prabaswari, Chancard Basumerda, and Bagus Wahyu Utomo. «The mental workload analysis of staff in study program of private educational organization». In: *IOP Conference Series: Materials Science and Engineering*. Vol. 528. IOP Publishing, 2019.

[48] Hugo F. Posada-Quintero and Ki H. Chon. «Innovations in Electrodermal Activity Data Collection and Signal Processing: A Systematic Review». In: *Sensors (Basel)* 20.2 (2020).

[49] Pekka Siirtola and Juha Röning. «Comparison of regression and classification models for user-independent and personal stress detection». In: *Sensors* 20.16 (2020).

[50] Hector D. Orozco Perez, Guillaume Dumas, and Alexandre Lehmann. «Binaural beats through the auditory pathway: from brainstem to connectivity patterns». In: *Eneuro* 7.2 (2020).

[51] Ruth Maria Ingendoh, Ella S. Posny, and Angela Heine. «Binaural beats to entrain the brain? A systematic review of the effects of binaural beat stimulation on brain oscillatory activity, and the implications for psychological research and intervention». In: *PLOS ONE* 18.5 (2023).

[52] Michael J. Hautus, Daniel Shepherd, Edmund Giang, and Jason Landon. «Can binaural beats facilitate autonomic recovery following exposure to an acute stressor?» In: *Complementary Therapies in Clinical Practice* 45 (2021).

[53] Matteo Raggi, Stefania Chiri, Silvestro Roatta, Rosita Rabbito, and Luca Mesin. «A Novel Approach for Acute Mental Stress Mitigation Through Adapted Binaural Beats: A Pilot Study». In: *Applied Sciences* 15.10 (2025).

[54] Tina L. Huang and Christine Charyton. «A comprehensive review of the psychological effects of brainwave entrainment». In: *Database of Abstracts of Reviews of Effects (DARE): Quality-assessed Reviews* (2008).

[55] N. Jirakittayakorn and Y. Wongsawat. «Brain Responses to a 6-Hz Binaural Beat». In: *PMC* (2017).

[56] I. Scala et al. «Analysis of the effect of 6-Hz binaural beats on cortical activity and autonomic nervous system parameters». In: *PMC* (2025).

[57] Leila Chaieb, Elke Caroline Wilpert, Thomas P. Reber, and Juergen Fell. «Auditory beat stimulation and its effects on cognition and mood states». In: *Frontiers in Psychiatry* 6 (2015).