POLITECNICO DI TORINO

MASTER's Degree in
Biomedical Engineering

MASTER's Degree Thesis

Multisignal approach for stress and workload analysis

Supervisors

Prof. DANILO DEMARCHI

Ph.D. IRENE BURAIOLI

Candidate

LAURA BAJARDI

December 2022

If you can dream it
you can do it.
- Walter Disney

Abstract

Every day, high stress and mental workload negatively impact people and, depending on the circumstances, could put them in danger or make them hazardous. When perceived, stress sets the mind and the body in a fight or flight situation, i.e., blood is redirected from the minor vessels and retained in the major vessels so that the muscular system can react and allow quick movements. Body temperature lowering, faster heartbeats, or increased sweating, especially in the palm of the hands, are some of the main consequences of this phenomenon. On the other side, mental workload refers to the amount of mental effort necessary to accomplish a task: people required to do more than their abilities or resources allow may perform worse or even fail. Health issues like chronic stress and depression can emerge in case of long-term exposure, possibly escalating into physical sickness; therefore, it is crucial to understand how these factors affect people. This thesis aims to explore the field of stress and workload assessment by analyzing features extracted from meaningful and easy-to-get biological signals, both in terms of intrusiveness and cost, through feasible methods. The Electrodermal Activity (EDA), the Photoplethysmography (PPG), and the body temperature were recorded from the subjects while taking the BiLoad Test. In particular, EDA is the footprint of skin electrical activity, proportional to skin moisture level and highly dependent on the emotional state. On the other hand, PPG tracks the Blood Volume Pulse in time, including respiration and heartbeat components. It is possible to extract from this signal the Heart Rate and Heart Rate Variability, which contains information about the sympathovagal balance, an indicator of sympathetic and parasympathetic nervous system activation. To generate the mental states this study wants to investigate, the BiLoad Test was developed, via the Matlab GUI. It is composed of two subsequent well-known tests adopted from literature: the first one is the Stroop Color and Word Test, mainly used for psychological assessment, chosen to generate stress in people; the second, the N-back Test, challenges people to endure an increasing mental workload through different levels. The g.HIamp 144 channels (a biosignals amplifier) and sensors from g.tec allowed the biological signals recording from about thirty subjects aged between 24 and 41 years while carrying out the test. Signals and test performances - reaction times and answers of each set - were elaborated and post-processed using Matlab. Each biological signal revealed a strong dependency on the participants' emotional arousal induced by the test, allowing the choice of the optimal features to use as indicators. A possible development of this study could be implementing a real-time, automatic detection system for moments of stress and high workload, especially for safety and health applications.

Table of Contents

List of Tables IV

List of Figures V

Acronyms IX

1 Introduction 1

2 Background 3
2.1 Stress . . . 3
2.1.1 Stressors . . . 5
2.2 Mental Workload . . . 6
2.3 Software . . . 8

3 Biosignals 10
3.1 Signals . . . 10
3.2 Selected signals . . . 11
3.2.1 Electrodermal Activity . . . 12
3.2.2 Photoplethysmography . . . 20
3.2.3 Temperature . . . 25

4 Experimental setup 28
4.1 g.tec Bioamplifier and sensors . . . 28
4.1.1 g.HIamp and g.Sensors settings . . . 31
4.2 The tests . . . 31
4.2.1 Stroop Color and Word Test . . . 32
4.2.2 N-Back Test . . . 33
4.3 Experiment . . . 35
4.3.1 Sample . . . 35
4.3.2 Test procedure . . . 35

5 Signal analysis 43
5.1 Data conversion . . . 43
5.2 Test results . . . 43
5.3 Biological signals elaboration and features extraction . . . 45
5.3.1 EDA . . . 46
5.3.2 PPG . . . 48
5.3.3 Temperature . . . 54
5.4 Statistical Analysis . . . 56
5.4.1 Correlation between individuals . . . 57
5.4.2 Correlation within individuals . . . 68
5.5 Test validation . . . 70

6 Conclusion 74
6.1 Considerations . . . 74
6.1.1 About the features . . . 74
6.1.2 About the test . . . 75
6.1.3 About the questionnaire . . . 75
6.2 Future development . . . 76

Bibliography 77

List of Tables

3.1 [3] Methods of electrodermal recording, features, and abbreviations. 15

5.1 Best performing features obtained by comparison between subjects. 65

List of Figures

2.1 A cartoony representation of a stressed person. 4
2.2 Different factors that affect people's Workload perception. 6
2.3 g.RECORDER interface: in the picture there are respectively EDA, Temperature and PPG plotted. 9

3.1 [25] Skin structure 13
3.2 [26] EDA analyzed recording locations: 1) fingers, 2) distal wrist, 3) central wrist, 4) vertical wrist, 5) chest, 6) foot (instep), 7) calf, 8) forehead, 9) neck, 10) shoulders, 11) back, 12) buttock, 13) abdomen, 14) armpit, 15) upper arm, and 16) thighbone. 14
3.3 [25] Graphical representation of principal SCR components in a typical wave (Type 1[3]). 18
3.4 [3] Four examples of overlapping SCR. It is possible to distinguish two other different types of waves, distinguishable (2) and non-distinguishable(3) evaluated with three criteria: (A) measures the amplitude of the second peak from the interpolated recovery line of the first; (B) is the *standard* method which evaluates the amplitude with respect to the second wave start; (C) in this case the first peak is not considered. 19
3.5 Graphical representation a typical PPG wave from one of the subjects' signals. 21
3.6 PPG wave features from one of the subjects' signals. 22
3.7 PSD of one of the test phases. 23
3.8 [21] Dorsal skin temperature of a subject while watching Alfred Hitchcock's movie "Psycho", used as a stressor. This subject showed a variation in temperature of more than 2°C. 26

4.1 g.HIamp bioamplifier 29
4.2 g.SENSOR temperature 29
4.3 g.GSRsensor2 30
4.4 g.SENSOR Oxygen Saturation 30

4.5 [36] Example of one of the SCWT versions cards. 33
4.6 [37] Example of 2-Back test structure. 34
4.7 Experimental set-up, with a subject carrying out the test. 36
4.8 Sensor positioning on the hand. 37
4.9 BiLoad Test initial window. 38
4.10 BiLoad Test demos and tests window. 39
4.11 BiLoad Test structure. 40

5.1 Signals new starting point selection from a few second window of the beginning of PPG. 46
5.2 [41] Decomposition of EDA signal in SCL, SCR and error, obtained through the CVX method proposed by Greco et al. 47
5.3 Example of a single SCR wave. 48
5.4 10s window of PPG signal. 49
5.5 An example of a PPG wave. 50
5.6 BPM signal extracted from PPG 51
5.7 Time-frequency PSD of the BPM signal; the white lines indicate the tests markers. 52
5.8 The simpathovagal balance over time; the red lines indicate the test markers. 53
5.9 Whole temperature signal over time. 54
5.10 Temperature during the second SCWT level. 55
5.11 Representation of the submatrix of interest contained in the Heat Map. 57
5.12 Heatmap of SCR std correlation among subjects. 58
5.13 Boxplot of SCR std correlation among subjects. 59
5.14 Heatmap of BPM std correlation among subjects. 60
5.15 Boxplot of BPM std correlation among subjects. 60
5.16 Heatmap of LF/HF correlation among subjects. 61
5.17 Boxplot of LF/HF correlation among subjects. 61
5.18 Heatmap of Final Temperature correlation among subjects. 62
5.19 Boxplot of Final Temperature correlation among subjects. 63
5.20 Correlation between features and test difficulty in the SCWT. . . . 66
5.21 Correlation between features and test difficulty in the NBACK. . . 66
5.22 An example of a subject's features correlation over the phases of the test. 68
5.23 An example of some subject's normalized features over the phases of the test, including rest. 69
5.24 Reaction time vs. tests levels. 70
5.25 Reaction time vs. tests levels, with score over-plotted. 71
5.26 Test results vs. tests levels. 72

5.27 Test results vs. tests levels, with score over-plotted. 72

Acronyms

ANS

Autonomous Nervous System

BPM

Beats Per Minute

BVP

Blood Volume Pressure

CNS

Central Nervous System

CVX

Convex Optimization Approach

ECG

Electrocardiography

EDA

Electrodermal Activity

EEG

Electroencephalography

EMG

Electromyography

GUI

Graphical User Interface

GSR
    Galvanic Skin Response

HF
    High Frequency

HR
    Heart Rate

HRV
    Heart Rate Variability

IBI
    Inter-Beat Interval

LED
    Light-Emitting Diode

LF
    Low Frequency

NS
    Nervous System

NS.SCR
    Non-specific Skin Conductance Response

PNS
    Parasympathetic Nervous System

PPG
    Photoplethysmography

RMS
    Root Mean Square

RT
    Reaction Time

S.SCR

Specific Skin Conductance Response

SC

Skin Conductance

SCL

Skin Conductance Level

SCR

Skin Conductance Response

SCWT

Stroop Color and Word Test

SNS

Sympathetic Nervous System

SAQ

Self assessed questionnaire

STD

Standard Deviation

TS

Test Score

VLF

Very Low Frequency

WL

Workload

WLS

Workload Score

Chapter 1

Introduction

In active societies like the one we live in today, people constantly have to face challenging situations, not only in their job field but in daily life too. Emotional activation could arise both from doing or just thinking of doing a task, for example, taking an exam, meeting the boss, or driving in traffic.

For this reason, lately, the effects of **Stress** and **Mental Workload** on people are more investigated: not only do these conditions have consequences in the short term concerning safety and performances[1] in general, but, when experienced for long periods, they could lead to physical and psychological health problems[2]. Many symptoms are addressed to these two causes: from the physical point of view, some examples are headache, insomnia, upset stomach, fatigue, and more; on psychological terms, we could mention emotional instability, frustration, anxiety, and even depression up to serious psychopathologies[3][4].

All these consequences of excessive burden have heavy social and economical consequences. It's easy to think of a worker failing their task or missing days at work due to illness: this brings a loss both for the workers and their companies, lacking results and losing money.[5]

This thesis aims to identify the most revealing biosignals and their best features for stress and mental workload identification and assessment taking into count the subjective perception each subject declared to have experienced at the end of an appositely created test. This way, it is possible to check and compare the perceived arousal with the experienced one. A variety of biosignals, suitable for this kind of research, were identified in the literature before selecting the more significant for this investigation, with particular attention put on the ones directly influenced by emotional state, then expected to exhibit relevant variations.

The Autonomous Nervous System activation, due to high levels of stress and cognitive workload, was induced through the BiLoad test, an apposite test entirely developed via Matlab©, based on the Stroop Color and Word Test and the N-back Test, selected to induce a particular emotional activation. They are acknowledged, in fact, for their aptitude to trigger the desired conditions: the first tends to increase the perceived stress through tasks, following the difficulty of the test; the second, on the other hand, challenges the subjects on the cognitive level, demanding mnemonic performance both for visual and auditory echelon. As previously anticipated, at the end of the test, each subject filled out a questionnaire in which they evaluated their perceived difficulty in performing each level of the test.

For this study, **Electrodermal Activity**, **Blood Volume Pulse** and **Body Temperature** are extracted during the administration of the BiLoad test and examined to find indicators, or features, of high levels of arousal. This choice is justified by two main reasons: first, the fact that the measurements are accomplished in a non-intrusive way, because sensors are fixed through surface electrodes; second, because the wirings don't obstruct the free movement of the subject, although undesirable for a better quality of the recorded signals.

Chapter 2

Background

Even though high stress and workload identification is an essential object of interest to take measures for pursuing a good quality of life, there isn't a gold standard method to follow for their evaluation. It is a significant barrier that prevents responsive intervention when a momentary condition prolongs over time, which can have severe and sometimes fatal consequences for the health of individuals[2].

The same can be said for workload: despite the fact that, especially in certain fields like aeronautics[1] or driving rather than others, it is of essential importance to recognize and measure in the case of exhausting conditions, there is no standard method for the recognition of people perceived level[6].

Anyhow, it seems from research in the literature that it is possible to extract the information needed about emotional evaluation, in particular arousal, through biological signals[7]: not only it is possible, but it is becoming of widespread interest, thanks to technological advances that provide more accurate and easy-to-use tools every day. Researchers are trying to find the best way, i.e. the (for now) missing gold standard, to exploit this technology and the interest of making life always better to reach this goal.

This chapter contains a brief description of Stress and Workload: what they are, how they generate, and how they affect the human organism. In the end, a short description of the software employed for this study.

2.1 Stress

The term **stress** describes an emotional state experienced in a situation in which a person has to react to an event or circumstance that requires a change, such as a dangerous or challenging situation. Psychology defines it as a complex reaction due to physiological and psychological factors.

It was first defined by Hans Selye, for this reason called "father of stress", to indicate the nonspecific body response to a demand, noticing that some of his patients manifested different diseases consequently to the same stimulus[8].

The stress response is a subjective reaction[7]: it is different for each person since everybody has a different capability to face a situation. In Figure 2.1, there is a cartoony representation of a man living his daily life with the burden caused by stress.

Figure 2.1: A cartoony representation of a stressed person.

Until a certain level, people can react adequately and overcome the situation: **eustress** is the positive perception of stress, it can have beneficial consequences on the person, such as higher concentration and, as a result, productivity. On the other side, when stress becomes difficult to manage, it acquires a negative connotation: **distress** is the commonly denoted **stress** and it is the one experienced when the arousal exceeds a subjectively critical level, in terms of intensity or duration, or situations perceived as dangerous or challenging[3].

In short-term responses, the Nervous System is activated as a *fight or flight* response, as defined by Cannon[8], causing variations in cardiac, nervous, and glandular activity so that the human body is ready to react, face the situation, and reestablish homeostasis, both on the psychological and physiological level. On the other hand, when experienced for long periods, distress can lead to physical and psychological consequences since the organism doesn't have the time to recover[9]. Diseases like diabetes, anxiety and depression[4], heart disease, and digestive issues have been related to stress. It can have dangerous long-term impacts on health, and can also trigger flare-ups in those with migraines or other stress-related diseases. Health conditions could be heightened on many levels, such as physical, physiological, psychological, behavioral, and social if an early intervention to handle daily stress was easily achievable through immediate identification[2].

It is possible to evaluate perceived stress from biosignals variations (a direct consequence of the Autonomic Nervous System activity)[3], even though quantifying it as a unitary distinct idea is difficult, especially since the arousal level and emotional experience elicited is subjective[7].

2.1.1 Stressors

As mentioned above, various subjective or objective situations can cause increased stress, such as environmental factors or life events. Stress-inducing stimuli or contexts of any kind that cause chronic strain are defined **stressors**[4]. It is possible to distinguish between psychological or physical stressors, even though most have a combined effect on the organism. From literature, many ways are recognized as stress generating, such as:

- **Visual stressors**, such as visual tests, pictures from the International Affective Pictures System (IAPS)[10], video clips, etc...;
- **Acoustic stressors**, like loud or annoying noises, mostly associated with other stressors[11];
- **Cognitive tasks**, such as mathematical operations solving o mnemonic tests performed in a short time[5];
- **Simulation of negative events or negative environments from everyday life**, for example, bad weather during the flight, traffic during driving or crowded areas[12];
- **Fear conditioning**[1]: thanks to ethical standards, when practiced on people this procedure doesn't involve any painful Stimuli, but in the same way it generates fear or phobias associated to normal stimuli[13];
- **Forced breathing**: voluntary hyperventilation is used as a method of inducing physiological activation to achieve a state of anxiety (not always achieved), to study how this is generated[14];

[1] experimentally mostly practiced on mice, this is an associative learning test in which a noise, a light or any *normal* signal is coupled with an *adverse* Unconditional Stimulus (usually a mild electrical shock) so that when the normal signal is sent, the adverse stimulus is expected, and its effect are generated even if it doesn't follow.

- **Preparation to a public speech**: in expectation of an anxiogenic circumstance, such as a public speech, a person projects themselves into the situation itself, thus presenting characteristic physiological variations.

2.2 Mental Workload

Mental workload indicates a person's brain activity level. It represents the relationship between the resources required to perform a task and the capability[15], in terms of both skills and tools, of the individual, where resources and capability usually depend on the context[16]. Performances also depend on subjects' age, experience, motivation, as well as physical and emotional state (Figure 2.2).

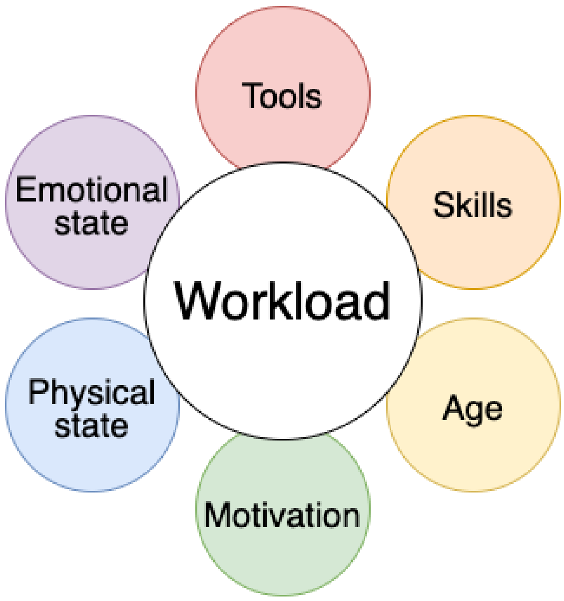

Figure 2.2: Different factors that affect people's Workload perception.

Given this, despite being dependent on objective demands and environments, it is understandable that workload is a highly subjective factor: it reflects not only task specificities but also the operator's abilities and performance. Thus, it is not possible to establish a direct correlation between task demand and mental workload, as this association is always dependent on the strategies and internal factors of each operator.

As a result, it is difficult to find a unique definition for Workload, the same way it is difficult to define a model to make predictions about the performances, whether they are perceived as overwhelming or not, and the attention level required. It is much indagated in the field of driver and pilot safety since there is a direct relationship between accident risk and mental workload: it is statistically demonstrated that the majority of accidents happen for lack or loss of attention. Kantowitz, in his studies expressed how Workload can be considered a subset of attention. In his “Attention and mental workload” article, he lists some studies of car accidents due to a lack of attention caused by acoustic distractions or to benefit second tasks, like talking on the phone[17].

Mental Workload involves various processes such as neurophysiologic, cognitive, and perceptual. It can be defined as the proportion of information processing capability needed to perform the task[18]. Too high levels of Workload require a lot of concentration and involvement from a person. However, this can cause mental fatigue, decreased performances such as slow information processing, incorrect responses, or even no responses at all, and even health-damaging effects in the long term. On the other hand, too low cognitive workload levels may result in errors or accidents due to boredom; it could be the case with highly automated systems that leave the user little to do[19].

Mulder (1986) distinguishes two types of mental effort: one for information processing in a controlled mode (i.e., *computational effort*) and one required when the operator’s energy state is affected (i.e., *compensatory effort*). While the first is used to maintain an adequate level of task results for variations in task complexity or the introduction of secondary tasks, the second mode plays its role when performance declines to a certain level due to fatigue. This phenomenon is defined as *dissociation*, and it shows when an excessive demand cannot be satisfied; as a result, there is a decrease in task completion. The compensatory effort is only possible for a while because extreme levels of effort increase task difficulty and mental workload. As an alternative to applying high levels of effort, defined by the approach toward the task, the operator may choose adaptive strategies[20].

Therefore, a real-time cognitive workload measuring system may have significant applications for avoiding task error or working under overload conditions. It could find use in various scenarios such as safety, Smart Technology, driver awareness, mental health monitoring, and Brain-Computer Interfacing (BCI). It is then desirable to find an equilibrium between user involvement and fatigue, such as an amount limit of mental effort and tasks that allow maintaining an adequate level of performance.

Many cognitive tests are utilized to generate overload, mostly regarding memory, calculus, or short reaction time tasks.

As is stress, Mental Workload is strongly linked to the activity of the Sympathetic Nervous System too, so it manifests itself in cardiovascular variation, making it recognizable in signals such as the Heart Rate (HR).

2.3 Software

MATLAB® was the main platform involved in this thesis: it is a versatile programming platform which allows a wide variety of operations and algorithm development. It was employed to visualize and analyze the biosignals recorded on the subject, from a first conversion from the bioamplifier extension to a ".mat" file, to the plotting, filtering, extracting feature from the signals and carry data analysis.

This software was also employed for the development of a the BiLoad application through the Matlab Graphic-User Interface, App Designer: this user friendly interface allows the construction of applications both from the code and interface point of view. It is in fact possible to switch any moment from the Design to the Code view, allowing in the first the drag & drop of the components (directly available from the library), and their graphical properties setting, while in the latter it is possible to associate functions and "callbacks" to these elements, i.e. the actual programming part, in the MATLAB editor. Not to be underestimates is the possibility to export the application on desktop or in a web browser using MATLAB online, or the packaging-exportation and installation in the MATLAB Apps tab. The application is also executable outside the MATLAB environment as a standalone desktop and web app using the MATLAB Compiler.

The other software utilized for this thesis is g.RECORDER (Figure 2.3). It is the interface for all g.tec biosignal acquisition devices, offering also tools for simple configuration and setup, data display, storage, and review. Once the utilized device is selected, it is possible to select the channels in which the sensors are connected, apply filters, select the type of signal to analyze, so that it automatically set the measuring units and the scale. During the recording, it is possible to visualize different sized windows on real-time, select the channels to hide and to show, and put different colored markers. It is possible to evaluate signals and settings in the offline/replay mode too, once they have been checked in the display mode and saved to disk.

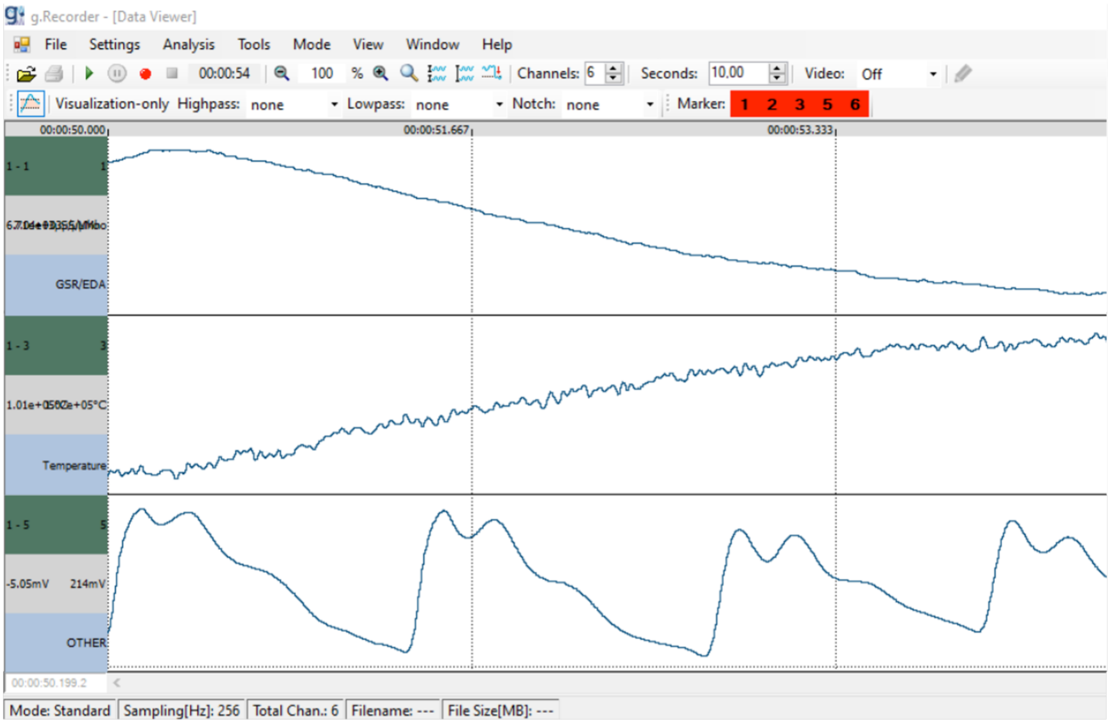

Figure 2.3: g.RECORDER interface: in the picture there are respectively EDA, Temperature and PPG plotted.

Chapter 3

Biosignals

3.1 Signals

In order to select the best signals to use for this study, an overview of the main biosignals that can be acquired non-invasively was made[11].

1. **Breathing**: recorded through a band positioned around the chest to evaluate its expansion and compression, it allows the estimation of the inhaled air volume. It can provide information about the individual emotional state, as negative emotions determine an irregular rhythm. Higher respiratory rate is a consequence in case of stress, in the same way as potential moments of apnea and irregular volume of inhaled air.

2. **Electrodermal Activity** (**EDA**): represents the electrical activity of the skin surface. Its values tend to increase and show higher variability in case of negative emotional states. Just like many other organ activities that can be recorded for medical or generic evaluation, skin is gaining every day more and more attention thanks to its signal information.

3. **Electromyography** (**EMG**): It represents muscle electrical activity. In case of stress and negative emotional states, it shows increased variations. Since the participants of this work experimentation are free to move during the test, this signal could be highly susceptible to several artifacts; therefore, it is excluded from our study.

4. **Electrocardiography** (**ECG**): Indicates the electrical activity of the heart, measured on the body surface. For the measurement of stress values, this signal is not directly used, but from this, the **Heart Rate** (**HR**) and **Heart Rate Variability** (**HRV**) are derived. HRV could also be extracted from PPG.

5. **Photoplethysmography** (**PPG**): measures **Blood Volume Pulse** (**BVP**), it gives information on the blood present in the vessels over time, consequently the oxygenation level, heartbeat, respiration, and vasoregulation, each of which is directly affected by the emotional state.

6. **Electroencephalography EEG**: The study of the neural activity of the brain through the electrical signal allows obtaining information on the central nervous system activity. The signal is analyzed in the frequency domain, and different bands can be distinguished. Contrary to the signals seen so far, this allows us to distinguish different levels of relaxation, therefore also of stress.

7. **Body temperature**: closely linked to BVP, it has lower values than physiological in case of stress, as there is a narrowing of the vessels aiming to re-direct blood to the muscle districts for body movement, a typical reaction in a *fight or fligh* condition. This value is taken at the extremities, then from the fingertips[21].

8. **Facial features**: several studies have analyzed expressive variations under stress conditions, being body language consciously and unconsciously the manifestation of a person's emotional state or intentions.

9. **Eyes signal**: In addition to facial expressions analyzed through the facial landmarks or facial muscles activation, great attention was paid to two other indicators:

   - **eye movement and blink**
   - **dilation of the pupil**.

   The former, however, reports sometimes contradictory results, while the latter demonstrates an increase both in size and in the frequency of diameter variation of the pupil under stress conditions.

3.2 Selected signals

Although EDA is recognized by many studies as the best signal for measuring stress, it is usually combined with other signals[22] to further confirm the results obtained.

For this study, among the biosignals previously listed, **EDA**, **PPG** (and **HRV**), along with **body temperature** were chosen to analyze emotional and cognitive activation.

3.2.1 Electrodermal Activity

Signal description

Electrodermal Activity, also known as Galvanic Skin Response (GSR), since it is very affected by the emotional state of an individual, found most of its application in the last 100 years in psychological assessment, and later in the medical field as well.

It was first associated with vasoregulation as the main regulatory mechanism (*vascular theory of EDA* by Neumann & Blanton[23]); afterward it was linked to sweat glands activity (*secretory theory of EDA*), first by Tarchanoff and later supported by Darrow[24]. This last study, in particular, showed the relation between these two measures, although with some latency.

Since strongly linked to the number of active sweat glands, Electrodermal Activity is the best biomarker of the Sympathetic Nervous System to identify and quantify the emotional activation of the individual.

Skin physiology

Skin is made of different adjacent layers, with different structures and properties[25]. As represented in Figure 3.1 main are:

- the **epidermis** (0,5 - 3 mm),
- the **dermis** (3 - 4 mm),
- the **subdermis** (0,5 - 2 cm).

The behavior of each component determines the skin's electrical characteristics (EDA). The corneum layer of the surface is the most affected: here, the sweat pours from the **sweat ducts**, regulated by the **sweat glands** governed by the SNA.

Glands are distinguished in endocrine and exocrine depending on where their product is secerned: in the case of exocrine, it goes on the surface, such as sweat. They are governed by the Nervous System, in particular, exocrine glands can be distinguished into eccrine and apocrine, where the first depends on the Sympathetic Nervous System while the latter is on Parasympathetic Nervous System. The first is activated in case of stress and a negative emotional state, while the latter has a calming effect on the organism. Sweat production is affected by various factors, but mainly emotional activation. Glands are distributed all over the body, with a high concentration on hands and foot palms[26].

As expected from the physiology of the body, van Dooren et al.[26] evaluated the differences between various recording locations (Figure 3.2 and found out that foot, fingers, and shoulders are most responsive, whereas arm, back, armpit, and thighbone were less responsive. However, being one of the best-performing sites, for

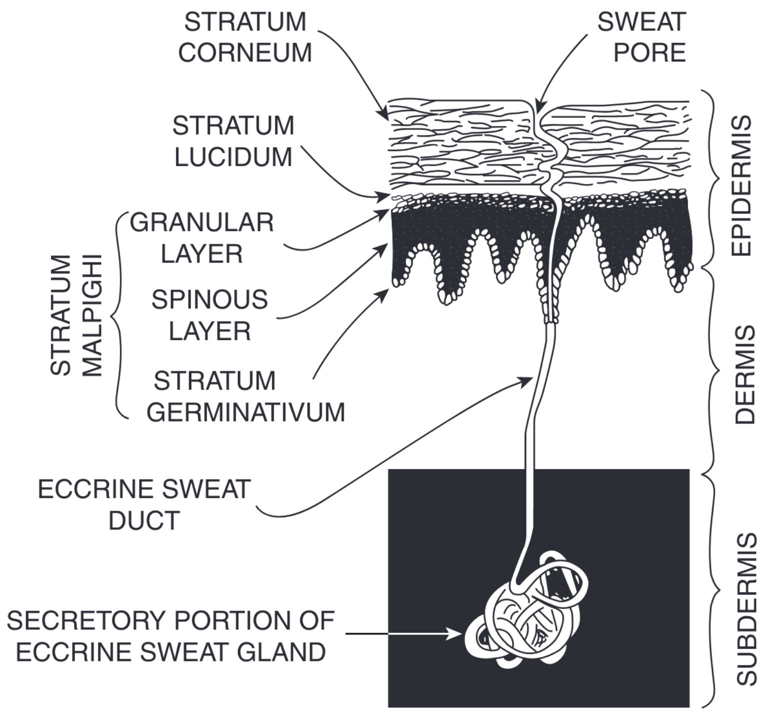

Figure 3.1: [25] Skin structure

convenience in the execution in the experimental laboratory setting, fingers were chosen as the EDA signal recording site, in particular, in the non-dominant hand.

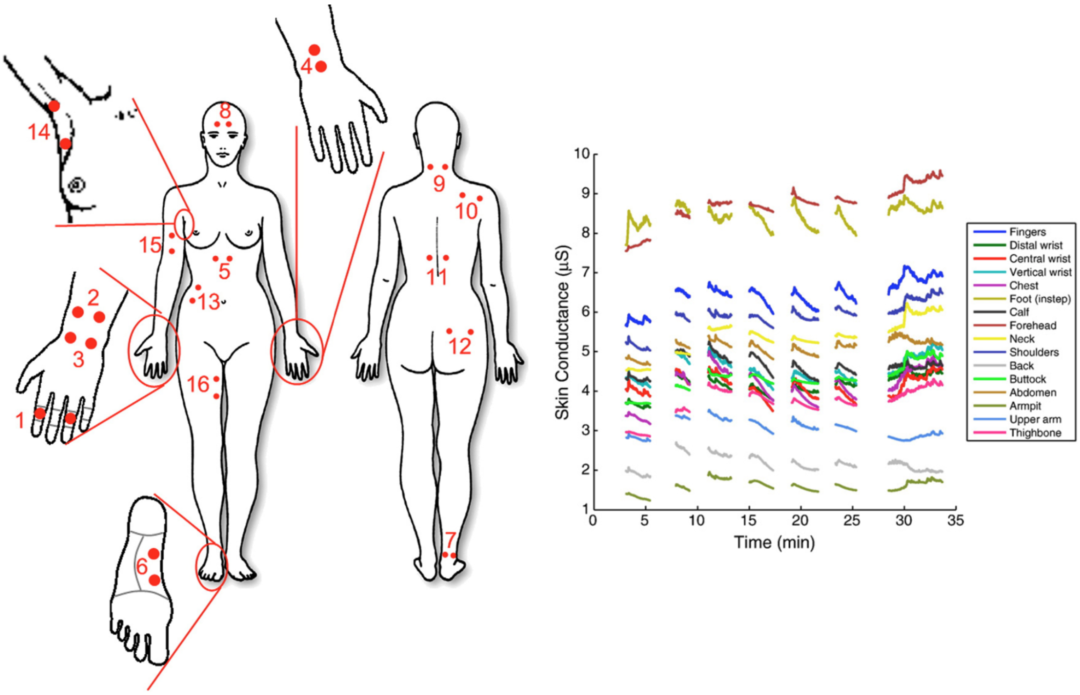

Figure 3.2: [26] EDA analyzed recording locations: 1) fingers, 2) distal wrist, 3) central wrist, 4) vertical wrist, 5) chest, 6) foot (instep), 7) calf, 8) forehead, 9) neck, 10) shoulders, 11) back, 12) buttock, 13) abdomen, 14) armpit, 15) upper arm, and 16) thighbone.

EDA shapes

Skin respects Ohm Law, so it is possible to extract the electrical impedance Z (100 Ω - 4.6 MΩ) by injecting an external current, to obtain a value highly affected by skin hydration.

As anticipated above, skin signal can have different forms (shown in Table 3.1) depending on the recording method.

These are:

- **Endosomatic**: the recorded signal is the potential difference (**Skin Potential** (**SP**)) between two sites of the skin, so it does not need external sources of energy. For this method, it is preferable to use one active and one passive recording site for the electrodes, otherwise, the signal would have a flat profile, with no relevant phasic component content. Moreover, it is necessary to lightly scratch the skin surface and to apply a gain of at least 1 MΩ entrance impedance (suggested 5 MΩ) to obtain a relevant signal. For these reasons, this first methodology does not find much use nowadays.

- **Exosomatic**: on the contrary, skin response to external energy is recorded in this case. Electrodes are positioned in two active locations for this measurement method. It is possible to distinguish the signal nature for continuous or alternate current applications. In DC:
    - When the current is constant, the measured signal is **Skin Resistance** (**SR**);
    - When the potential is constant, the measured signal is **Skin Conductance** (**SC**);

  In AC (5-100Hz):
    - When the effective current is constant, the measured signal is **Skin Admittance** (**SY**);
    - When the effective potential is constant, the measured signal is **Skin Impedance** (**SZ**);

Methods of recording	Endosomatic	Exosomatic			
Applied current		Direct Current		Alternating current	
Units	Skin potential	Skin resistance	Skin conductance	Skin impedence	Skin admittance
In general	SP	SR	SC	SZ	SY
Tonic (level)	SPL	SRL	SCL	SZL	SYL
Phasic (response)	SPR	SRR	SCR	SZR	SYR
Supplementary abbrevations					
nonspecific response	NS.SPR	NS.SRR	NS.SCR	NS.SZR	NS.SYR
frequency	SPR freq.	SRR freq.	SCR freq.	SZR freq.	SYR freq.
amplitude	SPR amp.	SRR amp.	SCR amp.	SZR amp.	SYR amp.
latency	SPR lat.	SRR lat.	SCR lat.	SZR lat.	SYR lat.
rise time	SPR ris.t.	SRR ris.t.	SCR ris.t.	SZR ris.t.	SYR ris.t
63% recovery	SPR rec.tc	SRR rec.tc	SCR rec.tc	SZR rec.tc	SYR rec.tc
50% recovery	SPR rec.t/2	SRR rec.t/2	SCR rec.t/2	SZR rec.t/2	SYR rec.t/2

Table 3.1: [3] Methods of electrodermal recording, features, and abbreviations.

In the biomedical field, **Skin Conductance** (**SC**) finds most of the applications.

EDA in disease studies

EDA finds wide use both for physiological and psychological applications. It finds much use in Dermatology, Neurology, and other medical contexts where the skin itself or the sympathetic branch of the ANS is involved.

Many studies have been carried out on people with various disorders. In the medical field, we can mention thyroid dysfunction, diabetes, rheumatism, asthma, headache, dialysis patients, other psychosomatic issues such as ulcers and hypertension, and other cardiovascular diseases such as chest pain, angina, hypertension, ischemia, and infarction.

Even though much research was made on EDA for clinical dermatological applications, this field prefers qualitative measurements of sweat rather than quantitative measurements of the skin's electrical activity. However, as discussed before, EDA is strictly dependent on skin physiology, so skin pathologies that alter the normal characteristics of the skin determine the variation of EDA features that can be analyzed and evaluated.

Neurology uses EDA and sweats secretion measurements to determine the type and degree of injury to the Peripheral and Central Nervous Systems. Due to minor differences in the sensory and sudomotor dermatomes, EDA is particularly useful in this area. EDA measurements offer an objectivity advantage over patient reports because electrodermal phenomena can produce responses directly corresponding to the severity of the damage that subjectively may not be perceived. In this case, however, endosomatic recordings are preferred over exosomatic due to the instrumentation available in neurologic laboratories (mostly EMG recorders); moreover, the recorded SPR signal is referred to as Sympathetic Skin Response (SSR).

Even though many studies showed inconsistent results, in the Psychopathological field, EDA analysis from studies on patients with Generalized Anxiety, Stress Phobias, Schizophrenia, and Panic Disorders anticipated a possible use of this signal for disease studies. As an example, Lader and Wing carried out many studies on anxious patients, getting as a result a different mean SCL value and a different trend in rest condition between the pool of patients and the control subjects. In fact, when the control pool tended to lower their SCL value due to rest and adaptation, anxious subjects showed a continuous slight increase during the recording, both with and without an acoustic stressor (1 kHz, 100 dB, 1 s duration noise). On the contrary, phobic patients didn't show a considerable variation in the EDA level, but high phobic responses to stimuli. For these studies, EDA recordings are often sided by fMRI recordings, to have an image of the CNS activation. Other studies have been conducted on panic diseases and psychopathic disorders. In particular, the latter was indagated by Hare, who concluded that psychopathic diagnosed patients showed a tendency for electrodermal hypoactivity and hyporeactivity (mostly for SCL than SCR). Follow-up studies also demonstrated that low SCL is already present at a young age in people who develop disruptive behavior when adolescents.

All the mentioned studies and others are described in Electrodermal Activity by Boucsein[3].

EDA and Stress

As anticipated in Chapter 2, stress causes identifiable emotional responses in ANS measures, especially noticeable in EDA.

To explain better how stress influences EDA, here follows a brief description of a study performed by Nomikos et al.[27]. In this study, two groups of subjects were shown a safety film about three wood-mill accidents. One of them had a shorter version without the scene anticipating the accident: this group showed a lower increase in SCL than the other group. Moreover, the higher variations in EDA level were recorded before the accident was shown, while just a little increase was recorded after seeing the accident.

Other studies compared different wait duration before the stressing event, obtaining as a result that, for shorter waits, EDA showed higher variations. These results demonstrate why, for example, the preparation before a public speech is a consistent stressor: the anticipation of a stressful event can elicit EDA changes comparable to or even higher than the event itself. Similar results can be observed in the outcomes of this work experiment results.

Other studies investigated the effects of predictability and controllability of the stressors during the test over the EDA. Geer and Maisel (1972) analyzed the differences between three groups under different conditions, elicited with the same stressor: a presentation with pictures of victims of violent deaths. Only the first group was able to *control* the stressor by deciding whether and when to change the picture; the *predictability* group knew the duration of each slide (the 60s) and were warned 10s before the picture changed through a warning tone; the third group had *no control* it didn't have any information about the duration of the single slide or why they would hear a warning tone. As a result, the first group exhibited lower EDR, while no control e predictability groups showed high amplitude values; however, the second group showed higher levels of adaptation over trials. This study demonstrated that controllability and predictability act independently from the ANS activation. EDL, on the contrary, didn't show a profound difference between groups except for the third, which resulted in lower values. This outcome suggests that this indicator could be the best to represent cognitive activity, while EDR contains information about the stimulus, like its aversiveness or whether an anticipator or not[3].

SC components and features

SCR The **phasic** component, also called **response**, quantifies the reaction of individuals in presence of a stimulus and corresponds to the rapid component of the EDA variations (high reactivity to stimuli). Unfortunately, there is not always a clear relationship between stimuli and EDR: this is due to the fact that both external and internal stimuli affect people, where the first can't be observed from the outside. For this reason, we can distinguish Non-Specific SCR (NS.SCR) and Specific SCR (S.SCR)[28], where the latter is the one generated on purpose. For this thesis, these signals are treated the same way, since it is difficult to discriminate the two cases.

In Figure 3.3, it is possible to see some of the main features of a single SCR wave. These are:

1. **Amplitude**, the difference between the peak value and the baseline, proportional to the number of activated eccrine glands; often difficult to evaluate in case of overlapping waves (Figure 3.4), typical in case of high arousal.

2. **Latency**, the time difference between the application of the stimulus and the response, of about 1 – 5 s, usually around 1.8 s.

3. **Rise time**, needed from the application of the stimulus to reach the peak of the response, evaluated between 10% and 90% of the peak value;

4. **Half recovery time**, which elapses between the peak and 50% of the peak value respecting the baseline value;

5. **Number of responses**;

6. **Area** under the curve;

7. **Slope** of the curve at the response, calculated as the ratio between amplitude and rise time;

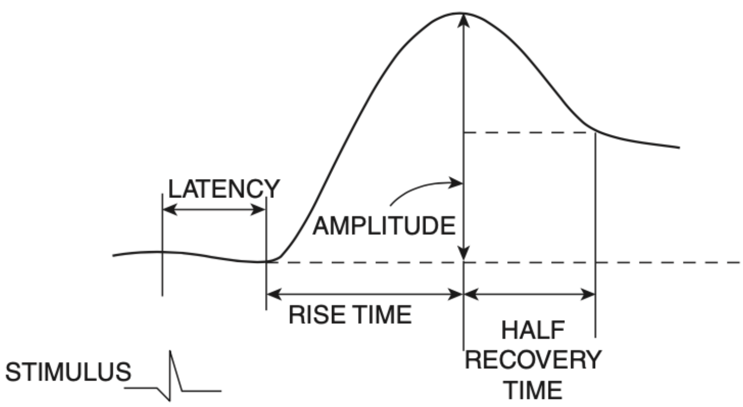

Figure 3.3: [25] Graphical representation of principal SCR components in a typical wave (Type 1[3]).

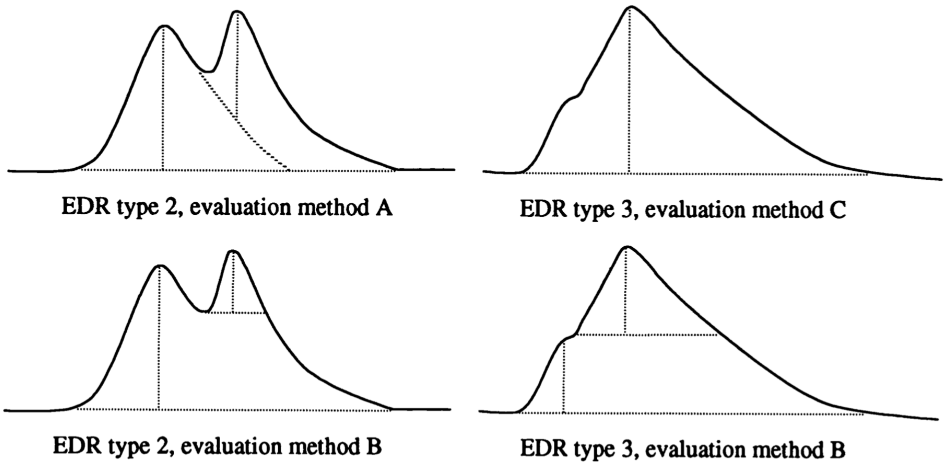

Figure 3.4: [3] Four examples of overlapping SCR. It is possible to distinguish two other different types of waves, distinguishable (2) and non-distinguishable(3) evaluated with three criteria: (A) measures the amplitude of the second peak from the interpolated recovery line of the first; (B) is the *standard* method which evaluates the amplitude with respect to the second wave start; (C) in this case the first peak is not considered.

Unfortunately, it isn't always easy to distinguish a single SCR waveform, since many peaks can overlap (Figure 3.4).

Higher amplitudes, slopes, and times, and lower latencies are recorded in case of greater stimuli.

SCL Tonic Skin Conductance or Skin Conductance Level is the baseline level of the signal; it would be the SC value in absence of stimuli.

It is the slow component of the EDA variation, in fact it can be found in a range between 0 - 0.05Hz[10]. This value is proportional to the number of sweat glands activated, therefore, to the ANS activation.

3.2.2 Photoplethysmography

Photoplethysmography is the signal that quantifies **Blood Volume Pulse** (**BVP**), blood volume changes in a particular vessel. It is measurable in a non-invasive and low-cost optical method that involves a light-emitting diode (LED) at infrared (IR) wavelength to illuminate the skin: the amount of transmitted or reflected light to a photodiode is measured, allowing the evaluation of the light absorbed by the blood, therefore, the quantification of the volume flowing.[29]

This signal contains much information, since it is affected by respiration and heartbeat, and allows the extraction of the heart pulse, respiratory rate, blood oxygen level (saturation), and information on tissue perfusion and some vascular and cardiac disorders. This is what makes PPG signals very appealing to assess vascular diseases, especially the effects of vascular aging, hypertension, and atherosclerosis, providing information on arterial stiffness and elasticity too, but also breathing pathologies such as hypovolemia.[30]

It presents noticeable variations in the case of SNS activation: the body enters in the *fight or flight* condition, therefore heartbeat fastens, and HRV lowers, as PPG waves amplitude and rise time do too.

For this reason, this signal is analyzed both in time and frequency domains.

Signal description

PPG signal is intrinsically constituted of two components[29].

- A pulsatile physiological waveform due to the cardiac activity: blood volume synchronously varies upon each heartbeat. In Figure 3.5, it is possible to distinguish the signal peak reached during heart systole, followed by a dicrotic notch registered right before heart diastole.

- A slow baseline with low-frequency content of respiration, thermoregulation, skin characteristics, and vasoregulation (see HRV section).

Even though the wave shape is well-defined, each subject has a different form, with a more or less evident dicrotic notch. Moreover, it is strongly affected by the sensor position.

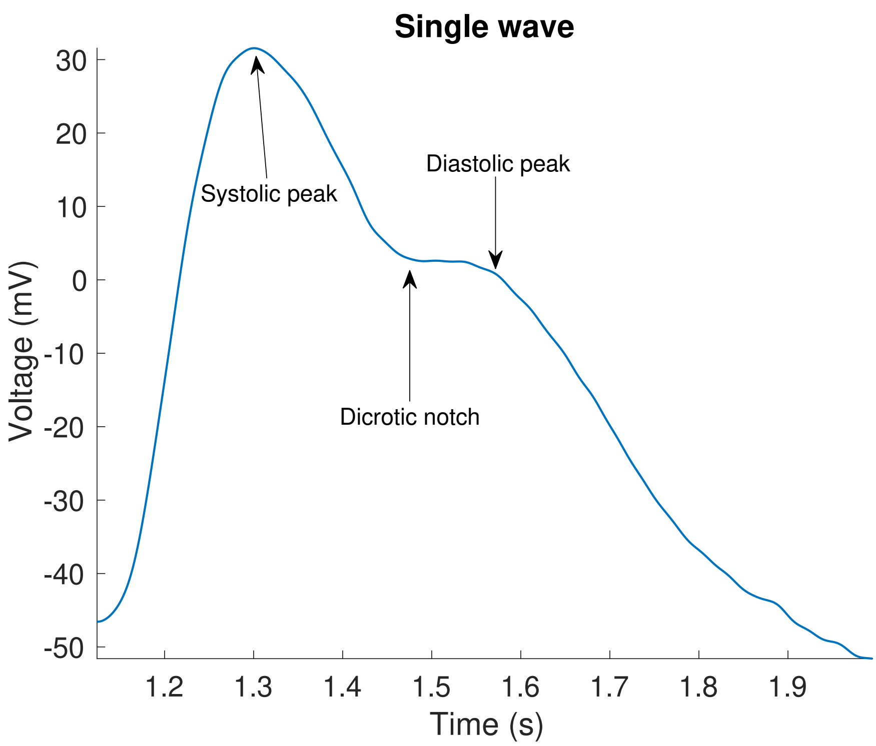

Figure 3.5: Graphical representation a typical PPG wave from one of the subjects' signals.

PPG features

As mentioned above, from this signal temporal and frequencies features are extracted, both from the single waves and the whole signal.

From every single wave (Figure 3.6), in time domain:

1. **peak position**;

2. **peak height**, calculated as the distance between the peak and the line crossing the wave's initial and final points;

3. **duration (width)**, converted in seconds as samples divided the sampling frequency;

4. **rise time**, elapsed from the first point to the peak.

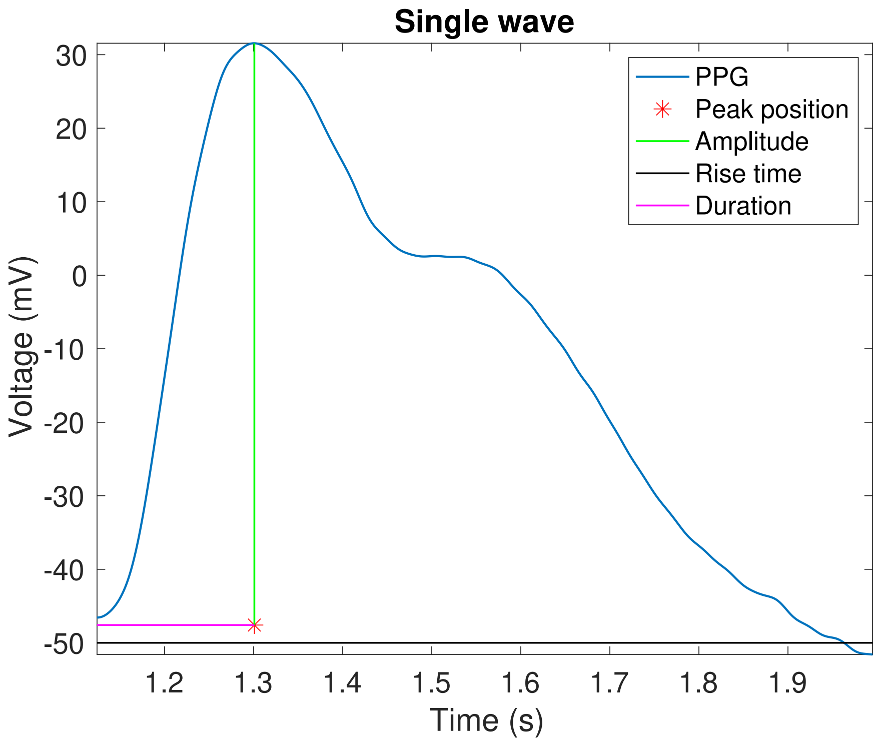

Figure 3.6: PPG wave features from one of the subjects' signals.

From the whole signal:

1. **mean peak distances**, which is the **Inter Beat Interval** (**IBI**) in samples, from which the **Beats Per Minute** (**BPM**) signal was calculated as:

$$BPM = \frac{fs[Hz]}{IBI[samples]} \times 60[s]. \tag{3.1}$$

   where fs = sampling rate;

2. **pNN50**, percentage of pairs of adjacent peaks intervals differing by more than 50 ms.

From the BPM signal, the frequency domain is indagated with the Power Spectral Density (PSD) (Figure 3.7), and contains information about the **Heart Rate Variability** (**HRV**), described in the following paragraph:

1. **Power content**, the signal energy found within the entire frequency band[31];

2. **LF norm**, the normalized power in low frequency band;

3. **HF norm**, the normalized power in low frequency band;
4. **LF/HF ratio**.

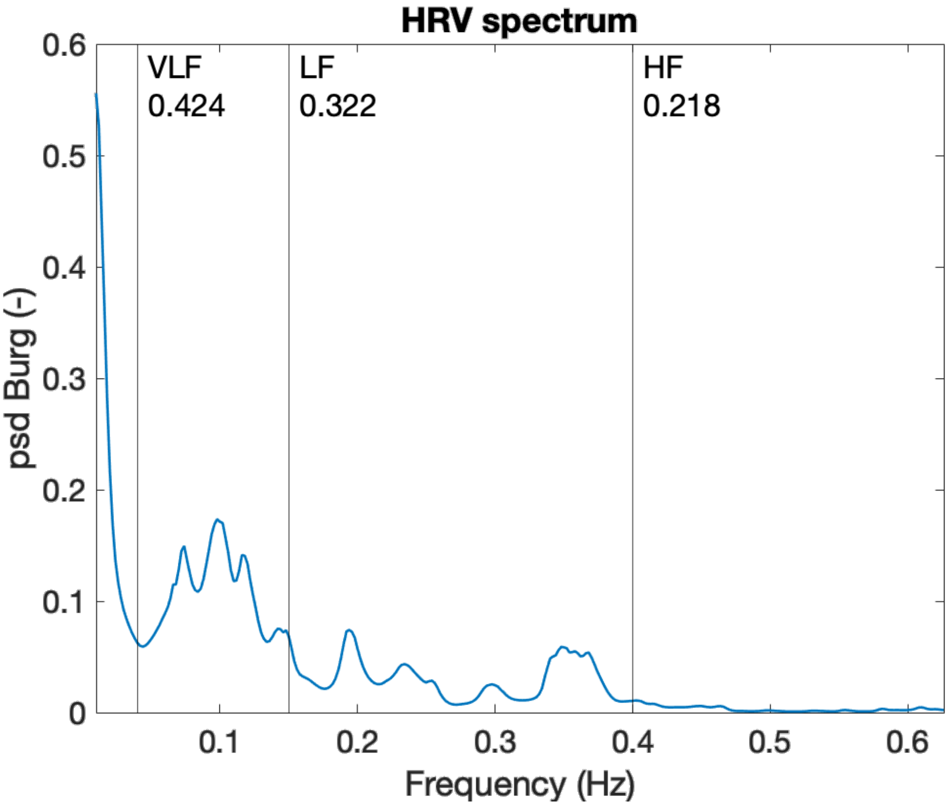

Figure 3.7: PSD of one of the test phases.

HRV **Heart Rate** (**HR**) is an indicator extracted from the ECG, as the difference in samples between two R peaks. Likewise, it can be derived from the difference between two BVP peaks (Equation 5.6): in this case, this signal is referred to as **Pulse Rate** (**PR**).

$$HeartRate = \frac{1}{IBI(s)} \times 60 \tag{3.2}$$

The variability of these signals is expressed as **Heart Rate Variability** (**HR**) and **Pulse Rate Variability** (**PRV**) respectively. Even if it isn't the same signal, PRV shows a level of correlation to the HRV of more than 99%[32], so it can be considered analogous. Therefore in this work, it will be called indifferently by either name.

Heart Rate Variability (HRV) frequency spectrum content is found as:

- **Very Low Frequency (VLF)**: f < 0.04 Hz
- **Low Frequency (LF)**: 0.04 Hz < f < 0.15 Hz
- **High Frequency (HF)**: f > 0.15 Hz

From which:

$$EnergyRatio = \frac{LF_{Energy}}{HF_{Energy}} \tag{3.3}$$

This ratio is also called **Simpathovagal balance** and represent the sympathetic nervous system activation over the parasympathetic. Higher value are found in a condition of stress or emotional arousal, since there is a more distinguished activation of the HF bands. Finally, VLFs are usually considered unreliable, so they are not considered in this study.

HRV gives information about the heart rate variations from the mean value. This index is strictly correlated to cardio-respiratory activity, which, in turn, could be regulated by the sympathetic and parasympathetic autonomous nervous system activation[31].

The HRV frequency content allows the estimation of the **sympathovagal balance**, calculated as the ratio between LF/HF. This parameter is very useful in the clinical field to monitor, for example, post-infarct and diabetic patients since it also quantifies the risk of Sudden Cardiac Death (SCD).

HRV globally depends on many factors such as health conditions, circadian rhythm, posture, and age, yet in particular is affected by factors such as breathing, posture, and emotional state.

From the anatomic point of view, the Autonomic Nervous System (ANS) innervates the Cardiac System (CS) through twelve different ramifications, which form the cardiac plexus. Cardiac ramifications originating from the sympathetic system have an excitatory function on the CS, tending to increase cardiovascular functions, heart rate (HR), and blood pressure (BP); those originating from the vagus or parasympathetic have inhibitory functions, tending to lower these cardiovascular activities. The counterposed actions of the vagus and sympathetic are combined such that under physiological conditions at rest, responding to continuous external and internal stimuli, the organism maintains a heart rate of about 70 bpm. This is due to the tendency of living beings to always reach homeostasis, i.e., an equilibrium condition of certain parameters, searched every time environmental conditions change. These changes in the sympathovagal balance determine fluctuations in the HR, thus generating HRV. Three main mechanisms regulate this activity[32].

1. **Respiration rhythm**: when inhaling, intrathoracic pressure arises, reducing the pulsatile volume, therefore, the cardiac output and the blood pressure. Baroceptors – blood pressure receptors sensitive to pressure variations in

vessels - detect these changes and alert brain centers, from which an inhibitory signal to the vagal system starts. Consequently, the heart rate increases again, thereby heart output and blood pressure. The intervention of the brain centers causes respiratory sinus arrhythmia, i.e., the heart rate exhibits fluctuations synchronous to the respiratory rate. The frequency content of this component is found between 0.15 - 0.4 Hz, and is the High-Frequency band, representing the vagal activity.

2. **Baroceptor reflexes**: when baroceptors activate the previously mentioned compensation activity from the cerebral centers, the frequency fluctuation originated by the loop of baroceptive reflexes, the *10-sec rhythm*, is accompanied by synchronous blood pressure fluctuations called Mayer waves. This contribution is found between 0.04-0.15 Hz, in the Low Frequency component.

3. **Thermoregulation** and **slow regulation mechanisms**: thermoregulation system uses four main mechanisms, two to raise the temperature and the other two to lower it. For internal combustion (thermogenesis) decreasing or irradiation increasing, in case of a higher temperature than the optimal temperature threshold, the circuit closes. In contrast, when the temperature is below the threshold the circuit shuts to minimize evaporation and irradiation. Changes in peripheral vascular resistance brought on by thermoregulation lead to variations in blood pressure, which set off the previously discussed baroceptive reflex mechanism. Various hormone regulation mechanisms regulate peripheral vascular resistance too; however, because it has a much longer latency than nervous system control mechanisms, it will result in very slow fluctuations in blood pressure. This last factor results in a Very-Low Frequency (VLF) lobe found before 0.04 Hz and represents the sympathetic system activation.

Higher HF power is correlated with stress, panic, anxiety, or worry[31]. As previously mentioned, the ratio of LF to HF power (LF/HF ratio) may estimate the ratio between the sympathetic nervous system (SNS) and parasympathetic nervous system (PNS) activity under controlled conditions[31]. Therefore, for this study, this sympathovagal balance (LF/HF) is essential to quantify the level of distress and cognitive load perceived by the subject during the test.

3.2.3 Temperature

Signal description

Body temperature is one of the most commonly known biosignals since it gives relevant information about a person's health condition: it is mostly used to check if a person has a fever and its severity. Its ease of measurement comes from the fact that thermometers are, by now, almost all superficial, so it's only necessary to hold it under the armpit for a few minutes to get the result.

In this case, the importance of this signal comes from the fact that body temperature is strongly dependent on the blood flowing in superficial vessels, which are mostly part of the microcirculation. So, when a person shows vasoconstriction and reduced blood flow, as it happens, for example, under stress, skin temperature is expected to vary as a direct consequence.

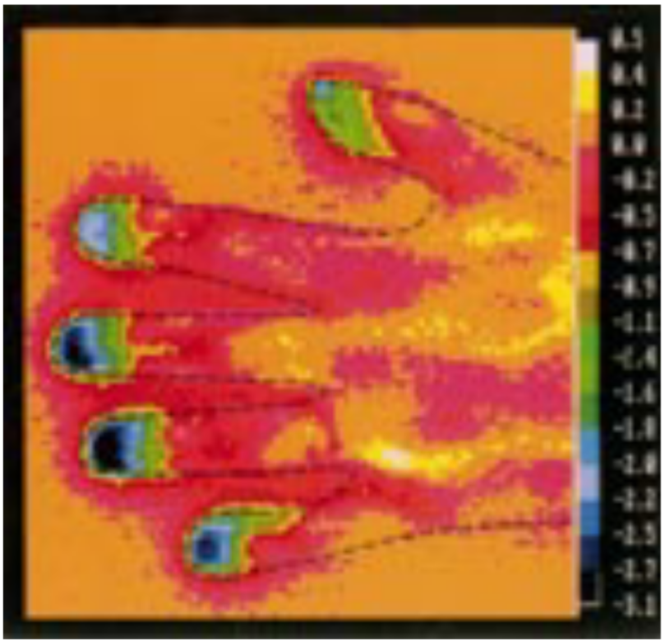

Figure 3.8: [21] Dorsal skin temperature of a subject while watching Alfred Hitchcock's movie "Psycho", used as a stressor. This subject showed a variation in temperature of more than 2°C.

Since fingertip skin arterioles, in particular, exclusively have sympathetic, adrenergic constrictor nerves, cutaneous microcirculation is primarily regulated by sympathetic activity. Therefore, they are frequently used experimentally as a measure of ANS activation, independent of the cause, which could be: contact with cold, forced breathing variations, physical pain, or mental or emotional stressors.

With their study, Kistler et al.[21] showed, through infrared thermography, that the hand regions most reacting to sympathetically mediated vasoconstriction reflex were indeed the fingertips (Figure 3.8). In this study, the stressor selected to arouse participants was a video from a horror movie, the shower murder from Alfred Hitchcock's "Psycho".

Temperature features

The features selected from this signal are:

- the **maximum** value,
- the **minimum** value,
- the **mean**,
- the **standard deviation**,
- the **slope**.

The same parameters have been indagated on the first derivative, but they were mostly eliminated after statistical analysis.

Chapter 4

Experimental setup

In this chapter, the experimentation is described: first, the instrumentation employed (bioamplifier, sensors, and software); second a section with the most common methods for emotional arousal and a theoretical description of the chosen tests; finally the experiment itself.

This last part presents a description of the sample of people, the procedure adopted for all of them, and a section with the BiLoad Test description, a custom-made application developed for the execution of the tests, fundamental in this study.

4.1 g.tec Bioamplifier and sensors

The g.HIamp-Research 144 channels bioamplifier (Figure 4.1) is a multi-channel high-end biosignal amplifier with USB technology. The device can record up to 144 biosignal channels with 24-bit resolution, such as EEG (Electroencephalogram), with a very high signal-to-noise ratio. The sampling rate can be set up to 38400 Hz. It can be connected directly to a computer via USB, eliminating the need for an additional data acquisition device.

Figure 4.1: g.HIamp bioamplifier

The sensors produced by g.tec employed in this study are:

- g.SENSOR temperature (Figure 4.2), to measure temperature values through a thermosensor, in the working range of 20-45°, 0,2° accuracy;

Figure 4.2: g.SENSOR temperature

- g.GSRsensor2 (Figure 4.3), to measure the EDA: the sensor measures the current by applying a small constant DC voltage to a couple of special finger electrodes; the output signal is given in mV (6.77 mV corresponds to 1 μS) and is proportional to skin conductivity in the 0-30 μS/μMho (micro Siemens/micro Mho) range. The sensor has an internal isolation barrier to prevent interference from GSR electrodes and other electrophysiological recordings.

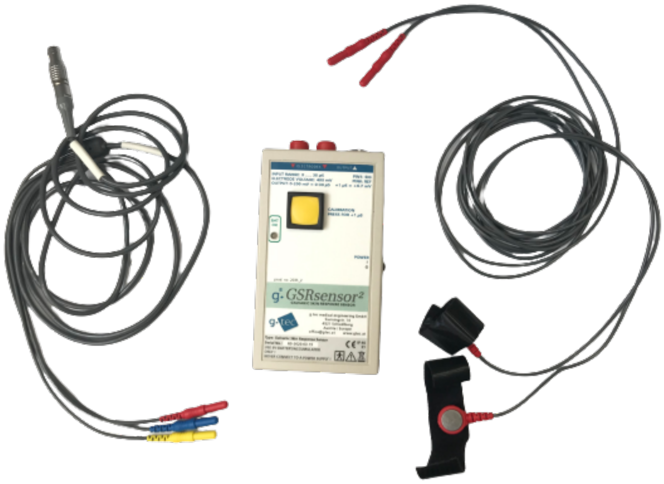

Figure 4.3: g.GSRsensor2

- g.SENSOR Oxygen Saturation (Figure 4.4) to measure the PPG. It actually allows three measures: SpO2-signal, pulse frequency, and plethysmogram. This last one is the only one recorded.

Figure 4.4: g.SENSOR Oxygen Saturation

4.1.1 g.HIamp and g.Sensors settings

The g.HIamp, interacts directly with the g.Recorder software (described in Chapter 2) that allows the visualization with a real-time pre-filtering application, as well as the recording and exporting signals in ".hdf5" format.

To use the software, first, run it in the *administrator mode*, rather than the user mode, to be able to apply changes to the settings for signals recording. After connecting the hardware (g.Hiamp, through the settings) and sensors and turning them on, it is possible to configure the device: after selecting the Channels utilized by the sensors, a sampling rate of 1200 Hz was set, and a notch filter at 50 Hz for all of them. To each channel (sensor) were applied other filters:

- for the temperature sensor, a low pass filter of 30 Hz, no high pass filter;
- for the GRS sensor, no high pass filter, a low pass filter at 30 Hz; it is possible to record only quick responses applying an additional high pass filter (0.05 - 0.1 Hz), but this wasn't the case;
- for the PPG sensor, a passband filter between 0.1 – 30 Hz.

It is possible to set markers to use during signal acquisition: I only used one color marker to indicate the moment when the participants started each of the two tests. When the **Start** or the **Recording** button is clicked, a Recording dialog opens, where it is possible to select the path to save the recording, set the Filename, add the subject information and any comments, and session data. It is possible to pause and resume at any moment the recording, whereas to terminate it just click the stop button and it is automatically saved.

4.2 The tests

After comparing various tests from the literature for the emotional activation of subjects, two tests were selected: respectively, the Stroop Color and Word Test (SCWT) for stress generation and the N-back Test for cognitive overload. Both act visually and auditory on the individual; moreover, all of them, the SCWT and the three tests composing the N-Back, are structured with increasing levels of difficulty.

There are numerous tests for psychological and cognitive assessment in the literature. Among them, we can distinguish different categories:

- Attention and concentration test: visual and auditory reaction time evaluation, with visive or auditory stimulus and logic games;
- Cognitive tests, such the arithmetical Montreal Imaging Stress Task;

- Cognitive tests, such arithmetical (for example the Montreal imaging stress task (MIST))[5];
- Language;
- Perceptual organization: visuospatial and tactile;
- Verbal and visual learning and memory[33];
- Motor functions;
- Concept formation and reasoning.

4.2.1 Stroop Color and Word Test

The Stroop Test assesses the relative speed of reading color names, naming colors, and naming colors in incongruous color words (e.g., the color red used to print the word "blue"), overriding the reading response, i.e., inhibiting an overlearned response in favor of an unusual one. Poor Stroop Test performance is attributed to brain dysfunction, psychopathologies, or injuries[34].

"**Stroop Effect**" refers to this conflict/interference situation.

The Stroop Test paradigm is one of the most recognized in experimental psychology[35]. The first published study about the relative speeds of color naming and color-word reading dates back to 1886, before Stroop began his studies. Around 1935, he developed and used the test that now bears his name. His original studies involved three cards with 100 elements each:

1. A black color-word reading card;
2. A chromatic color-word naming card with incongruent word color;
3. A pure color card filled with different solid color squares.

Five colors and/or color words were employed (red, blue, green, purple, and brown), arranged in a 10 x 10 grid of evenly spaced rows and columns, afterward reduced to 5 columns.

Both the colors and the words were arranged so that each would appear twice in every column and row, with no color word immediately succeeding itself. Moreover, for the chromatic color-word naming cards, no word was printed in the corresponding color, but in each of the other four colors.

The Stroop test has been translated into several languages, and in a "language-neutral" version that uses numbers rather than words. The main problem is the presence of many different versions of the test.

Figure 4.5: [36] Example of one of the SCWT versions cards.

Moreover, numerous studies have been carried out to investigate various factors that may influence performance, such as age, gender, nationality, language, brain damage, IQ, and so on[34]).

4.2.2 N-Back Test

The N-back test is a robust cognitive test that manipulates working memory load, one of the major components and a good approximation of workload. Thus, by analyzing biosignal variations during each task, this test aims to modulate cognitive workload and to understand how it affects people on a psychological and physiological level[37].

The N-back test presents the participants with a stream of stimuli, and the task is to decide whether each stimulus matches the one presented N items before (Figure 4.6).

It was first introduced by Kirchner in 1958 (even though there was an unpublished dissertation by Kay, 1953, which presented this test for the first time) as a visuospatial task with four load factors (0 to 3) and by Mackworth (1959) as a visual letter task with up to six load factors. Gevins et al. (1990) introduced it to the field of neuroscience as a "visuomotor memory task" with one load factor (3-back)[38].

Several processes are involved, including the encoding of incoming stimuli,

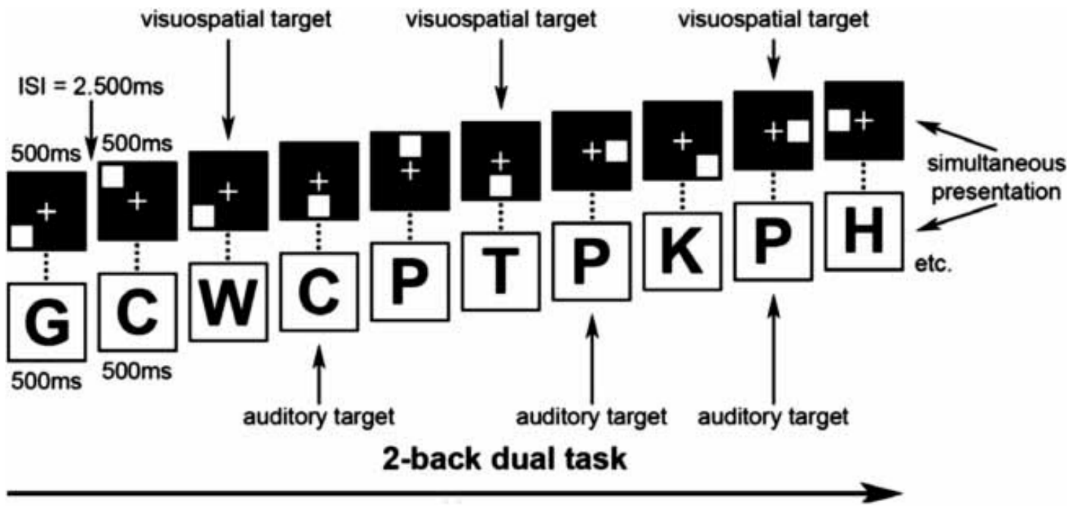

Figure 4.6: [37] Example of 2-Back test structure.

monitoring, maintenance, and updating of the material, and matching the current task to the previous Nth, along with decision, selection, inhibition, and interference resolution processes. These last two, in particular, step in when the current task matches a preceding stimulus but not the previous Nth to resolve this conflict[37]. Because of the task's sequential nature, all those processes must be carried out simultaneously, particularly the storage and processing of the material, which presumably let the N-back test become a Working Memory measurement method (Jonides et al., 1997; Kane & Engle, 2002).

As a result, the N-back task is a complex measure that appears to involve multiple stimulus and material-independent processes. As N increases, the processing load increases monotonically, and so do the number of errors and RTs. On a neural level, it appears that, as the processing load increases, activation grows alongside. Although behavioral and imaging results appear to be largely consistent, the N-back task's psychometric properties remain largely unexplored.

Only a few studies have addressed the psychometric properties of the N-back task in visual and verbal 0 to 3-back tests, but only some explicitly reported measures, with controversial reliability too. Other studies also indagated the correlation of N-back test performances with age, intelligence and IQ, attentional control processes, and stimuli nature (introducing dual-task versions)[37].

4.3 Experiment

4.3.1 Sample

The subjects who participated in the signal collecting were recruited from the university by volunteering; they were a total of 33 people, aged around 26 ± 3.5 years, but unfortunately, only 28 of them were considered. However, as the SNS reaction is subjective and not every person has the same amount of responsiveness or signal, it is preferable to have a pool as large as possible.

4.3.2 Test procedure

For the signals acquisition, the same procedure was followed for all the participants. The test took around an hour to complete, beginning with a period of adaptation to the atmosphere in the room, followed by the first test, a pause to restore the signal to its resting state, a second test, and finally, another rest period. Before the beginning, they were asked to remove watches, rings, and bracelets from the opposite hand than the one taking the test, to set the mobile phone to silent mode without vibration, and not to wash their hands before the test to avoid superficial skin conditions far from the "natural".

As shown in Figure 4.7, the subject was seated in front of the laptop on which they had to perform the test, in a chair placed sideways to the same table where the supervisor was sitting. This was the optimal solution since, during the entire test duration, they could monitor both the correct signals acquisition from the computer equipped with the g.Recorder interface, and the correct functioning and execution of the test from the laptop. Once seated, as long as their hands were free, the participants were asked to enter their data in the BiLoad application interface, firstly fundamental to integrate the test results with the recorded signals and the self-assessed questionnaire, afterward anonymized.

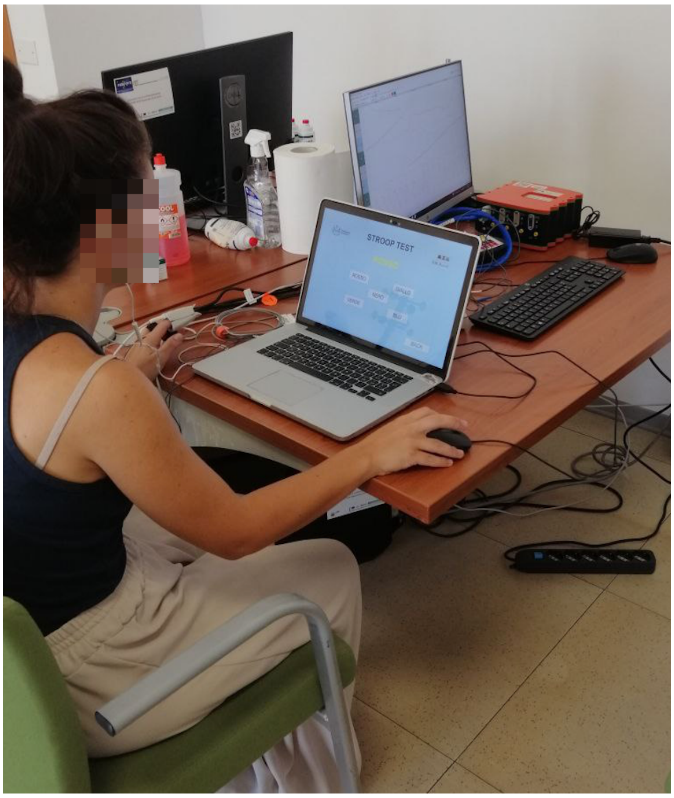

Figure 4.7: Experimental set-up, with a subject carrying out the test.

An interval of time was chosen, depending on the subject, to acclimatize to the room, both in terms of environment and temperature.

After this, the three sensors were connected to the hand of the subject (in this case, for all 28 participants on the left hand). As decided following the analysis of the previously mentioned studies related to the best recording sites, the sensors were placed as shown in Figure 4.8:

- for the EDA, on the second phalanx of fingers 2 and 3,

- for the PPG, on the tip of the index finger with the LED positioned above,

- for the temperature, the thermo-sensor on the fingertip of the fifth finger, chosen as a less irrorated area, therefore in anticipation of greater temperature variations than other positions.

Figure 4.8: Sensor positioning on the hand.

The subjects were asked to keep this hand relaxed for the whole BiLoad test, with the arm lying on the chair armrest. Moreover, they were given necessary and mandatory earphones for performing the test. All sensors and earphones were sanitized with alcohol before the session began. The signals recording began immediately, to have longer recordings possible.

During a first time interval, the motivation and objective of the test procedure were explained to the subject. They received information about the rules and the number of levels of the tests whereas the single tests' precise duration, the number of questions and the difference between the Stroop Test levels were omitted to avoid "spoiling the surprise effect", hence any emotional reactions.

Once the explanation was over, the subject was invited to practice as long as they wished (up to 15 min for timetable issues), executing the specifically created demos of the two tests to clarify the rules without compromising the preparation for the emotional test (explained in the following section).

Inside the room, there was a silent environment to allow the subjects to concentrate on the test; during the three rests, on the opposite, they were distracted, as much as possible, by a conversation from the activity in progress to allow them to relax and bring the signals to the state of rest.

After recording at least 10 minutes of this initial phase, called *rest* as the first resting phase, during which the subjects became familiar with the test, we moved on to the first test, the Stroop Test. Once the first test was over, the subjects were invited to relax and distract themselves in order to bring the signals back to resting conditions. After at least a 5 minutes signal was recorded during the *middle rest*, the subject could start the N-Back test. At the end of the test, the sensors were further kept on to record the *final rest* signals of at least 5 minutes.

At the end of the signals recording, the subjects completed a self-assessed questionnaire about the test.

BiLoad application development

The tests were administered through the BiLoad Test, a custom-made application developed through App Design (Matlab®), entirely written from the start. It takes its name from the dual load, emotive and cognitive, applied to the subjects through the Stroop Color and Word Test and the N-Back Test.

The interface is designed to be as simple as possible, and the background is in light blue, chosen because considered a calming color[39]. The initial screen requires the keyboard entry of personal information: name, surname, and age; instead, there is a drop-down menu for gender, with three alternatives: M, F, and other.

Figure 4.9: BiLoad Test initial window.

As previously said, all information was subsequently anonymized. At the end of the data entry, the START button below activates, and, when pressed, leads to the presentation screen of the two tests, with the two demos and the Stroop test accessible.

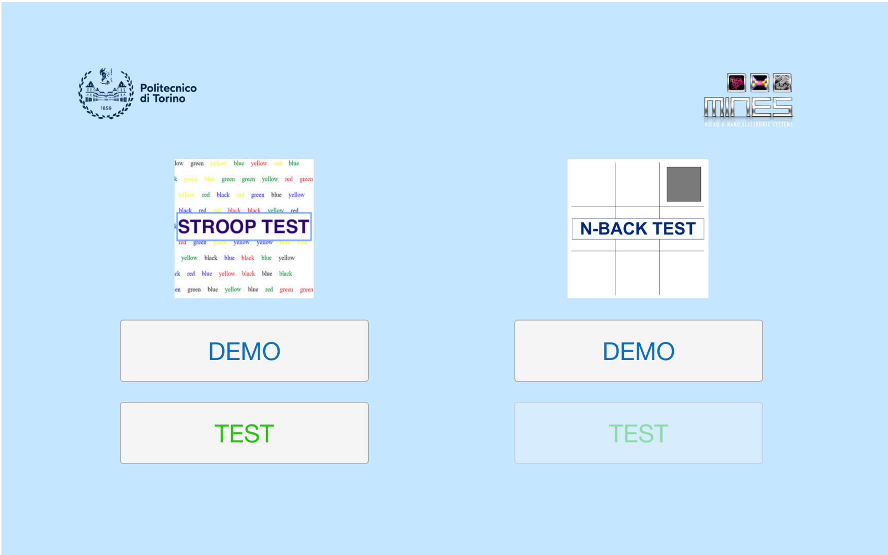

Figure 4.10: BiLoad Test demos and tests window.

The demo of the Stroop test is very similar to the test, differing in no time limit to answer, congruent and incongruent questions proposed indifferently, and answers fixed in the same buttons for all the questions. For the N-BACK test demo, the dual 2-back level is proposed to allow a clear explanation of what both the visual and auditory and dual test consist of, with a medium difficulty level between those proposed (from 1 to 3). From both the demos windows, it is possible at any moment to go back to the former one through the BACK button; on the opposite, once started the test it isn't.

Here is a detailed description of the two tests' structure.

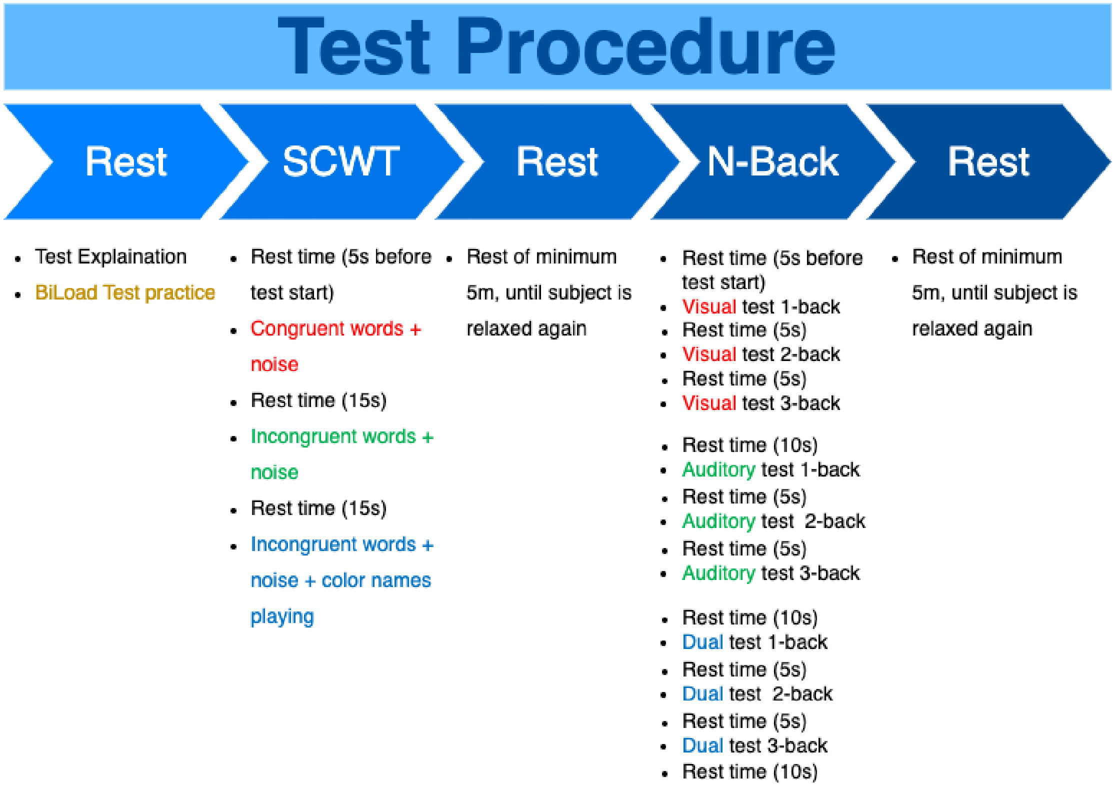

Figure 4.11: BiLoad Test structure.

SCWT Structure The Stroop Test exploits the discrepancy between the name of the color and its color font to create a conflict in the subject, who needs more time to say the correct answer (see Stroop Color and Word Test subsection). Since there is often little time for each question, this test is an efficient stressor. An additional acoustic stressor is introduced to make this test more stressful: a rapid ticking clock at varying volumes aims to annoy, and preferably stress, the subjects. The choice of the varying volume is to try and avoid people's adaptation[40] to noise.

For this application, the test is composed of three increasing difficulty levels, each of 90 questions, with a maximum of 2.5 s to answer.

Moreover, all the wrong answers activate an acoustic buzz, aiming to generate stress in an inversely proportional way respecting the test performances.

1. The first level is designed to be the easiest: all the colors appearing in the question are congruent, i.e., their color font is the same as the color word; moreover, all the answer buttons keep the same position during the whole test (as in the demo).

2. The second level presents two important differences from the first:
    - the colors are incongruent, so the color font is always different from the color word, requiring the subject more time to process the information, not always available since the time allowed to respond is short ;
    - the answer button change for each question: this is the main change the subjects have to face since they indeed tend to remember the button position and try to answer relying on their memory. This could lead to wrong answers, increasing the performance stress level.

   These two differences are expected to generate a medium-higher level of difficulty than the first one.

3. The last level introduces a distractor, so it isn't expected to be much harder than the second: in addition to the inconsistency between word and color, an audio clip that randomly repeats color names is introduced, creating further conflict between sight, reading, and hearing. People used to noisy environments are expected to isolate themselves from the audio and show performances similar to the second level, while the others are not.

To generate the questions for the test, two vectors of five elements are generated:

- one with the color words {ROSSO, GIALLO, VERDE, BLU, NERO} (i.e., {RED, YELLOW, GREEN, BLUE, BLACK} in Italian);
- one with the correspondent color codes {[1 0 0], [1.00,1,0], [0.39,0.83,0.07], [0 0 1], [0 0 0]} for the color font.

A random entire number comprised between 1 and 5 is generated for the question.

- In the congruent level, the same number is used to select the color word and color font from the arrays.
- For the incongruent level, another random number comprised between 1 and 5, different from the former, is generated, so that one of them selects the color word from the array, while the other the incongruent color font.

N-back Test Structure The N-back test, as previously said, is a cognitive evaluation tool to assess the difficulty level perceived by a subject to complete tasks. In my BiLoad application, this test is composed of three different tests, an auditory, a visual, and a dual test, each proposed in 3 increasing levels (1-back, 2-back, and 3-back) of 30 questions each with a maximum of 2.5 s per question to answer.

This is profoundly different from the Stroop Test: for instance, not all questions require an answer. It is, in fact, structured so that all the questions are randomly generated, and the correspondence of a task with the previous Nth is guaranteed for only 20% of the questions. When there is no correspondence, the subject only has to wait for the timer to run out.

Another difference is that the subject doesn't have to find the correct answer button: there is only one button for the first two tests (respectively "POSITION" and "AUDIO"), while the two of them are activated for the dual test.

In this last test, in case of both audio and position coincidence, it doesn't matter which of the two buttons is clicked first, it is sufficient to click both to proceed to the following question. As for the previous test, the N-Back notifies each answer with a sound, but, in this case, correct answers too.

To generate the questions for the audio, a vector re-calling correspondent letter audio with the English letters [c, h, j, k, l, q,r, s, t] was compiled. To generate the question, in both acoustic and visual and dual tests, just a random entire number comprised between 1 and 9 is generated. In the case of the visual task, it just corresponds to the position of the square on the grid; for the audio task, it is employed to call the position of the letter in the vector, which corresponds to an audio clip.

Self assessed questionnaire

As anticipated, at the end of the signal recording for the *final rest phase*, subjects were given a **self-assessment questionnaire** (**SAQ**) to fill out.

The questions first involved entering the personal information put in the BiLoad application starting interface. Thereafter they were asked to assign a score between 1 and 5 for each level of the two tests, based on the perceived cognitive WL, where 1 meant *none* and 5 was *too high*.

Finally, they were invited to insert any comment they had about the test. More than one subject pinpointed the difficulty perceived at increasing SCWT levels, particularly the fact that, from the second, responses' positions began to shift and that the third level's distractor, the voice randomly repeating color names, created high conflict. On the contrary, unfortunately, some pointed out that, in the last two sections of the N-Back Test, the dual 2-back and 3-back, particularly the latter, the perceived difficulty was so high that they stopped making an effort to answer the test correctly and felt like giving up.

Chapter 5

Signal analysis

All the recorded signals, together with the test results, both the scores and the reaction times of each response, were processed through MATLAB.

5.1 Data conversion

As mentioned in the g.HIamp and g.Sensors settings section, the output file from the g.HIamp is in the ".hdf5" format. An apposite code was developed to extract the signals, where, through the *h5read* function, for each subject, they were converted in a ".mat" matrix of three rows (1-GSR, 2-Temperature, 3-PPG) and n (= time duration x sample rate) columns, called *data*. In addition, the vector of the positions of the markers is extracted: it is a 1x2 matrix since I only marked the SCWT and the N-Back tests start.

Since the output signal of the bioamplifier is in mV, the three signals were converted following the datasheet indication: the ranges are expressed both in mV and in the corresponding measuring unity (μS, °C, mV), so the formulas are empirically found as:

$$GSR\,[\mu\text{S}] = data(1,:) \times \frac{10^3}{6.67} \tag{5.1}$$

$$Temperature\,[°\text{C}] = data(2,:) \times \frac{10^3}{8} + 20 \tag{5.2}$$

$$PPG\,[\text{mV}] = data(3,:) \times 10^3 \tag{5.3}$$

5.2 Test results

The outputs from the BiLoad application are five different files:

1. *surname_name_Data_dd-mm-yyyy_hh-mm-ss.mat*:
   a matrix with the subject data inserted in the initial interface, the test performances (reaction times and percentual results), and the final score.

2. *surname_name_SCWT_Time_dd-mm-yyyy_hh-mm-ss.mat*:
   with the answering reaction time matrix of the Stroop Color and Word Test, with 90 rows (questions) and 3 columns (levels).

3. *surname_name_SCWT_Raw_dd-mm-yyyy_hh-mm-ss.mat*:
   the test results in a 90x5 matrix, where the rows are the questions and 5 columns with:

   1 - [M,N], where M = color font random generated number and N = color-word random generated number

   2 - Answer, expressed color font code of the button

   3 - 1/0 correctness

   4 - 1/0 wrongness

   5 - 1/0 missing answer

4. *surname_name_NBACK_Time_dd-mm-yyyy_hh-mm-ss.mat*:
   the answering reaction time matrix of the N-Back Test with 90 rows (30 questions per three levels) and 4 columns ("Visual", "Acoustic", "Dual - Visual", "Dual - Acoustic"). There are entire rows of zeros in the first N rows of each N-back set since tasks are referred to the Nth before, and at the beginning, there are no former tasks to refer to.

5. *surname_name_NBACK_Raw_dd-mm-yyyy_hh-mm-ss.mat*:
   a 2x3 cell, with the names of the levels in the first row and the results matrices in the second. The matrices are structured as follows:

   - *Visual* and *Acoustic* are 90x5 matrices:

     1 - Question, Position/Audio random number
     2 - 1/0 answered
     3 - 1/0 correctness
     4 - 1/0 wrongness
     5 - 1/0 missing answer

   - *Dual* is a 90x10 matrix:

     1 - Question, Position random number
     2 - Question, Audio random number
     3 - 1/0 Position answered

4 - 1/0 Audio answered
5 - 1/0 Position correctness
6 - 1/0 Audio correctness
7 - 1/0 Position wrongness
8 - 1/0 Audio wrongness
9 - 1/0 Position not given
10 - 1/0 Audio not given

As previously specified, personal information was only needed to recollect all the test outputs and signals of each person in a folder, later it was anonymized.

From these matrices, the reaction times were extracted and analyzed by comparison with the perceived difficulty, as described next.

5.3 Biological signals elaboration and features extraction

The following subsections explain in detail the specific MATLAB scripts and functions developed to analyze each biological signal.

All of them were divided into five signals (rest, SCWT, middle rest, NBack, final rest) and analyzed as different but sequential so that, every time the signals were windowed, no window overlapped two different phases of the test.

Since the sensors take a few moments to couple with the subjects, the beginning of each signal had some noise, therefore, an apposite function was written to set a new starting point from the PPG signal, which resulted in being the most affected. This function works as follows. First, a few seconds of the signal are selected and highly filtered with a passband filter between 0.5 and 1.5 Hz, to only keep the trend, which is a sinusoidal wave (Figure 5.1). Considering the distance between the first two valleys as a neighborhood centered in the first of them, if the starting point is found before the first sample of the signal, the neighborhood shifts in the following valley. The valley is the signal's new starting point. Since there is a lag between the filtered and the original signal due to the filtering, the same neighborhood is centered in the same position of the sinusoidal signal to find the same valley position in the original signal, and the minimum value of PPG found that neighborhood is the new starting point.

A number of samples equal to the samples excluded were removed from all the recorded signals' beginning so that all the signals had the same length, and the markers were consequently shifted.

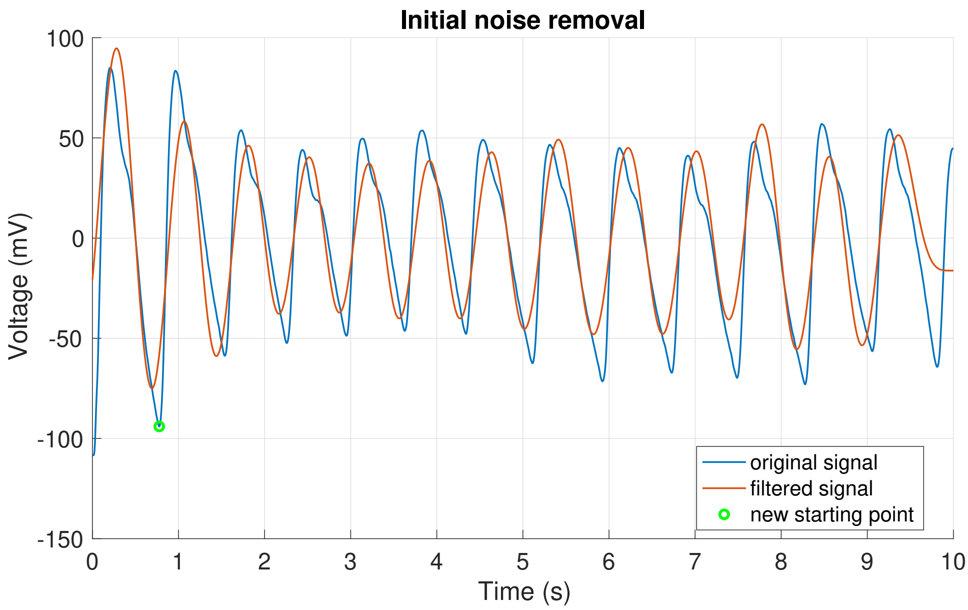

Figure 5.1: Signals new starting point selection from a few second window of the beginning of PPG.

5.3.1 EDA

For the EDA signal decomposition in its SCR and SCL components, the Convex Optimization Approach model offered by Greco et al.[41] was adopted (Figure 5.2). This method allows the extraction of the tonic (slow) component through the interpolation of 10s distant points (since the SCL has a low-frequency content), considering it an optimization problem. The phasic component is derived from every single wave, modeling them through a bicompartmental pharmacokinetic model representing the diffusion of the sweat through the gland ducts. Finally, an error component is found as the difference between the original signal and the two components:

$$EDA = SCL + SCR + Error \tag{5.4}$$

Since it is expected to have a very low frequency content (between 0.05 and 1.5 [10]), the SCR component was filtered again with a low pass filter of 5 Hz, to avoid as much noise as possible[42].

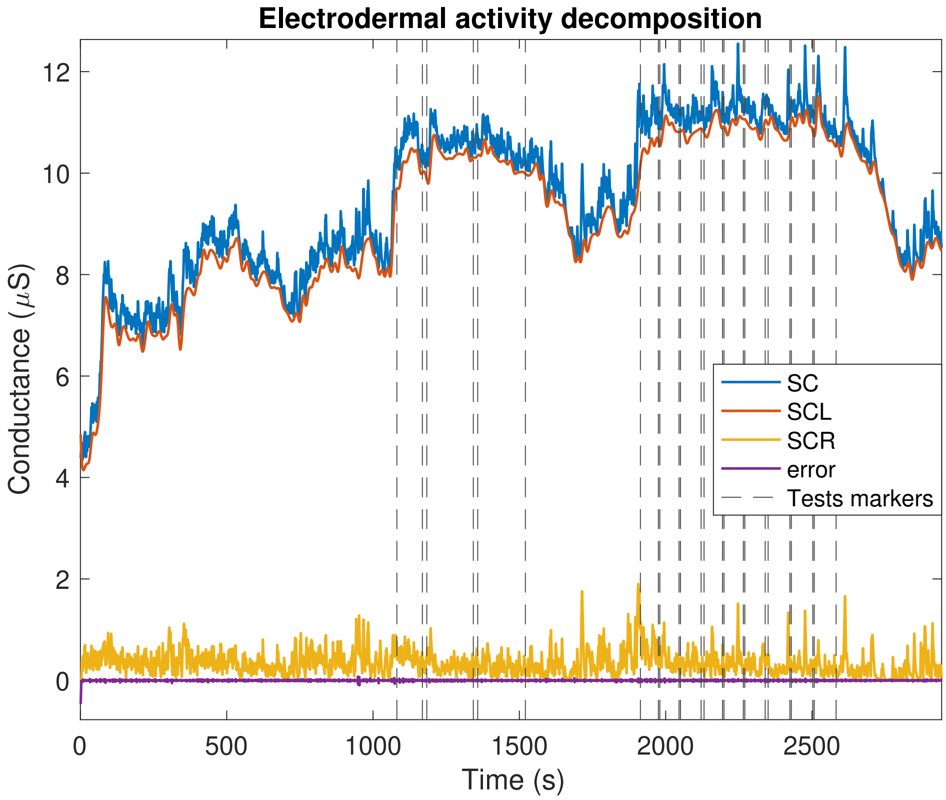

Figure 5.2: [41] Decomposition of EDA signal in SCL, SCR and error, obtained through the CVX method proposed by Greco et al.

Both the two components were analyzed in 90s windows: for each window, different features were evaluated. In particular, from the SCL:

- the **mean value**;
- the **standard deviation**;
- the **interpolated slope**;

and for the SCR (Figure 5.3):

- the **mean peak amplitude**;
- the **standard deviation**;
- the **slope**, calculated as amplitude over rising time;
- the **number of peaks per minute**;

- the **mean rise time** for all the bursts.

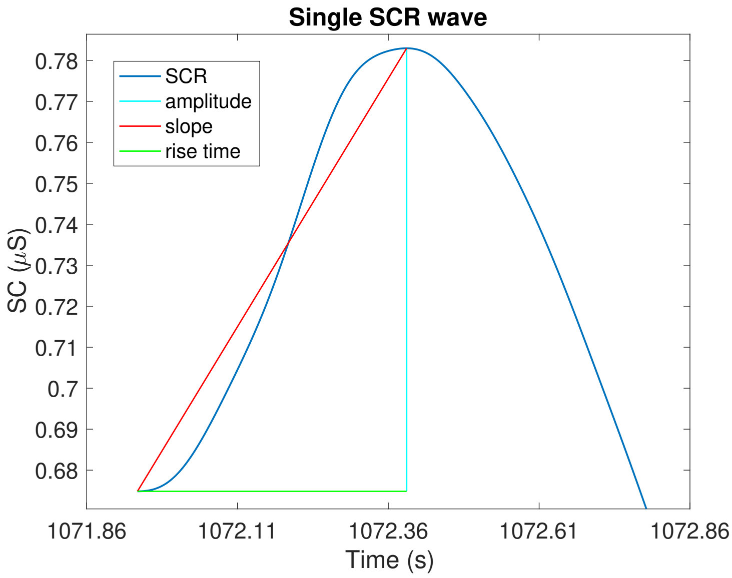

Figure 5.3: Example of a single SCR wave.

Since not all SCR peaks can be linked to an external cause[28], control was implemented to only select peaks distanced by at least 0.4 s. This value was chosen by observing the signal: in fact, from the literature, there isn't a well-defined minimum rise time, but it ranges between 0.8 and 3 s[43] or even less[3]. In this case, there were multiple external stimuli, and by analyzing the signal it seemed that 0.8s led to too many peaks exclusion.

5.3.2 PPG

The PPG signal was the most elaborated to process. I expected it to be a very high content signal since it has dual information: the first about the PPG itself, the second about the HRV, extracted from the positions of the peaks.

The signal was first filtered between 0.5-10 Hz to exclude noise[30].

Just for the PPG analysis, a new starting point was selected for all the five signal phases: this choice was due to the fact that to extract the features from the waves, all the considered waves had to be whole, i.e., from minimum to minimum, with the peak comprised between them. Consequently, a short signal segment was lost both at the beginning and the end of each of the five signals, but it wasn't a damaging loss since it was just less than a second for each of the five phases over a total signal length of about forty to fifty minutes.

The following procedure was applied for each of the five parts of the signal. Each signal was divided into 10s windows (Figure 5.4), but continuity between windows was guaranteed since, in case of an interrupted wave at the end of the window, the following would start from the final point of the former.

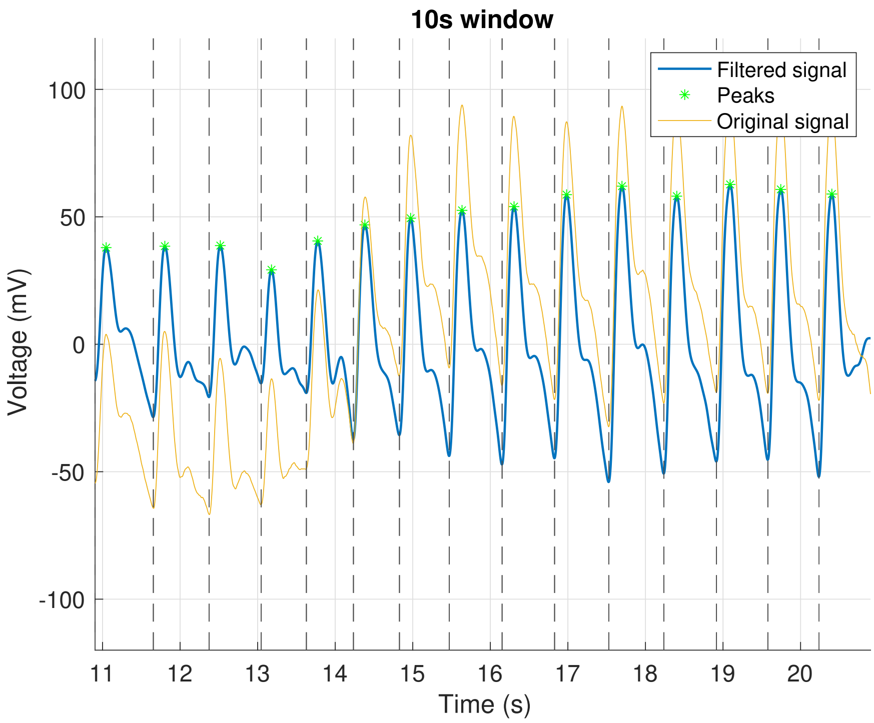

Figure 5.4: 10s window of PPG signal.

For each window, with the same method as the starting point selection, the valleys and the peaks were found, paying attention to exclude the dicrotic notches as peaks: to do so, it was imposed a minimum distance between the peaks of 0.35 s. This value was chosen since 0.35 s distant peaks would translate into a heartbeat of 171 bpm, a non-physiological rate considering this experiment activity: even though the aim was to generate stress, all subjects were in a rest condition, seated and in a quiet and mostly empty environment. Moreover, since the dicrotic notch could present a local minimum, the end of the wave was set as the last minimum found between the peak and the following one.

After finding them, all the single waves were analyzed, and the following features were extracted from each (Figure 5.5):

- the **amplitude**, calculated as the distance of the peak from the *base*, i.e. the line passing through the initial and final point of the wave,

- the **time duration**, from the wave formation to its last point,
- the **rise time**, elapsed between the wave start and peak.

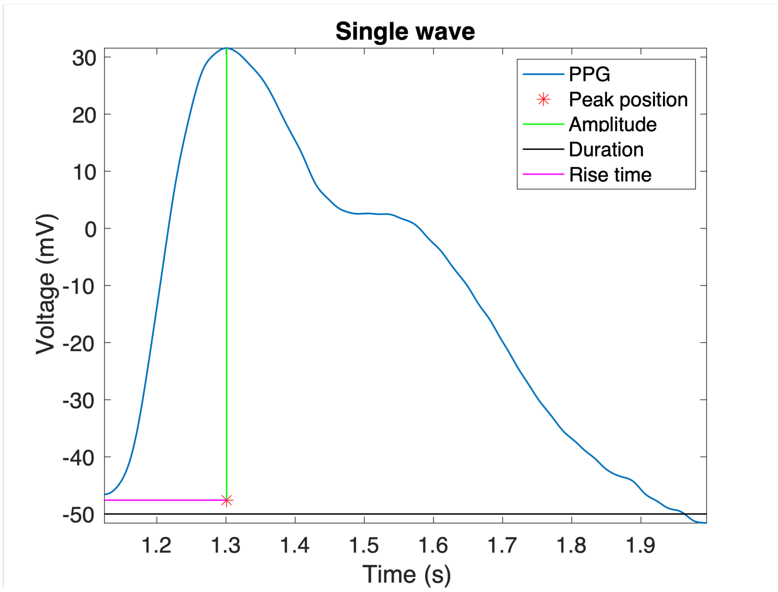

Figure 5.5: An example of a PPG wave.

For each wave, different controls were applied to check the conformity of the wave, whether it was affected by noise or movement artifacts.

Since the shape of the wave is well-defined, the adopted controls were:

1. the interval didn't have to be too short or too long with respect to a range between 0.35 and 1.2 s (i.e., 50 and 171 bpm);
2. the peak had to be located in the first half of the wave period;
3. the peak width at 80% of the height had to be larger than 1/4 of the wave;
4. the rise time had to be more than 0.1 s;
5. the height had to be positive;
6. there must be only one peak;
7. the wave had to be whole and not interrupted.

From the peak distances, the **Inter Beat Interval** (**IBI**) was calculated and converted in **Beats per minute** (**BPM**) through the formula 3.1.

When doing so, another control was applied to exclude the outliers. For this study, the allowed considered values were included in the range:

$$[mean(BPM) \pm 4\ STD(BPM)]. \tag{5.5}$$

The obtained result[1] is shown in Figure 5.6.

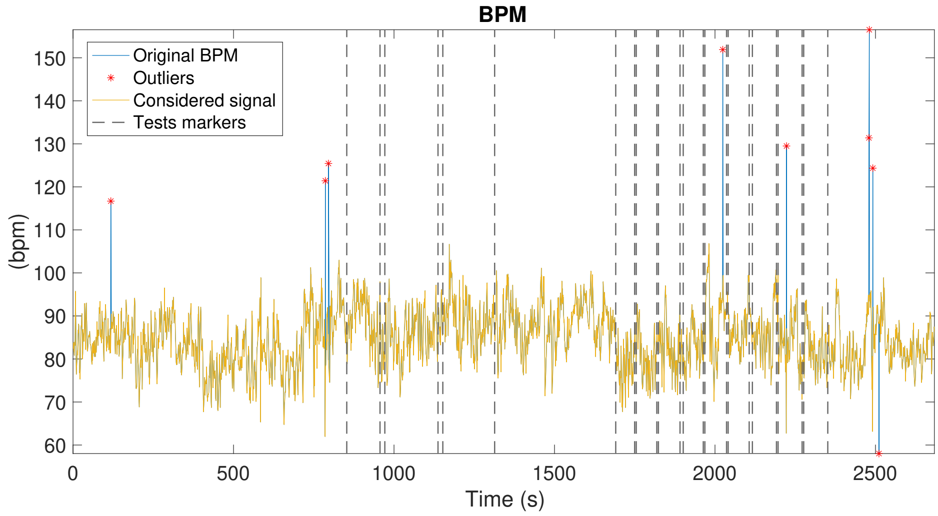

Figure 5.6: BPM signal extracted from PPG

.

HRV

The BMP signal was exploited to extract information about the **Heart Rate** (**HR**) and **Heart Rate Variability** (**HRV**). To do so, this signal was divided into about two minutes ($mean(BPM) \times 2\ [samples]$) windows[44] and analyzed. From this signal, the frequency domain analysis was executed, obtaining the Power Spectral Density of each window.

From this:

- the **total power** (Ptot);
- the **normalized Very Low-Frequency** (VLF_norm);
- the **normalized Low-Frequency** component (LF_norm);

[1] It is possible to notice from this subject a relevant variation in heartbeat already few seconds before the test started: this could be due to the fact that the anticipation of a stressful event, i.e. performing a test, can be itself a stressor, as mentioned in the Background section about stressors.

- the **normalized High-Frequency** component (HF_norm);
- the **LF and HF ratio**, the **Simpathovagal balance** ($\frac{LF}{HF}$).

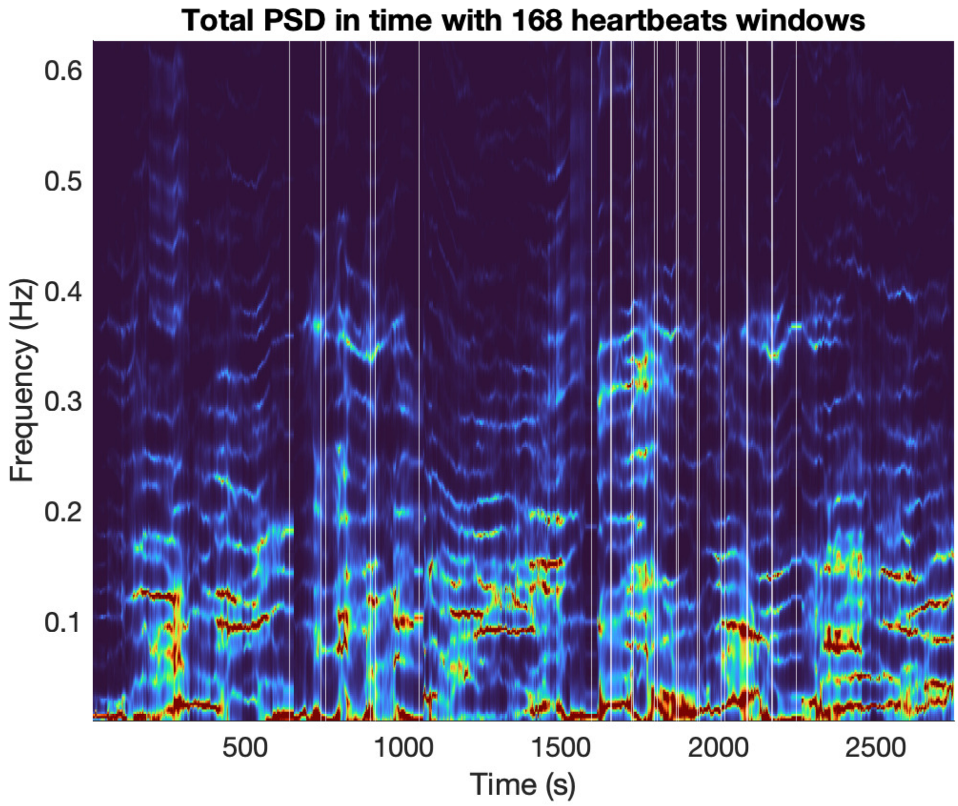

Figure 5.7: Time-frequency PSD of the BPM signal; the white lines indicate the tests markers.

From Figure 5.7, it is possible to distinctively notice the high-frequency bands activation during the test phases, marked by the white lines: on the left, six lines mark the start and finish of each of the three SCWT levels, on the right, three sets of six markers very close to each other mark the three levels for each of the three tests constituting the N-Back test. These bands' activation is exactly what we would expect to find in the situation of emotional activation[45].

Due to the increasing HF content, the LF norm showed a decrease, so in consequence, the LF/HF ratio did too. As mentioned, the sympathovagal balance is the best HRV feature for emotional arousal identification, therefore, from these results, it appears that the test generated the desired results (Figure 5.8).

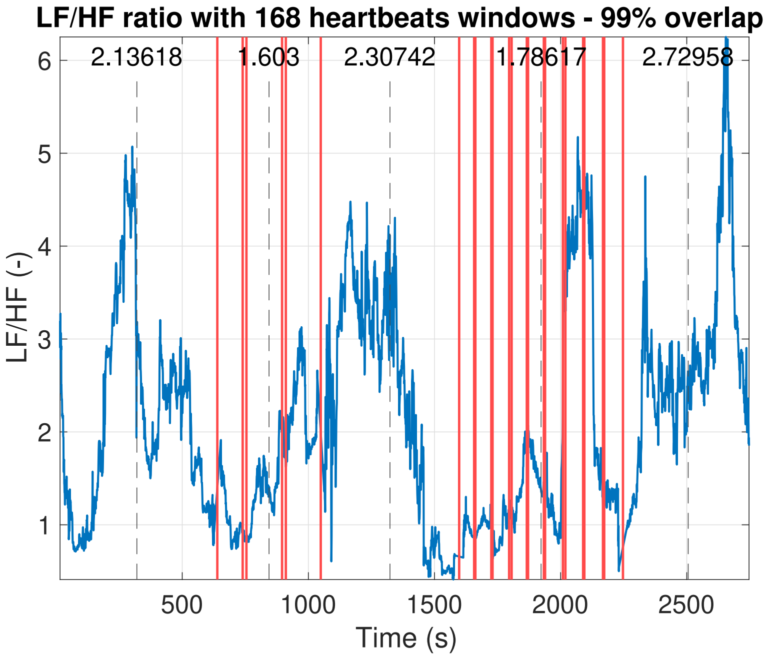

Figure 5.8: The simpathovagal balance over time; the red lines indicate the test markers.

5.3.3 Temperature

The temperature signal (Figure 5.9) was first filtered with a 10 Hz low pass filter, to exclude noise.

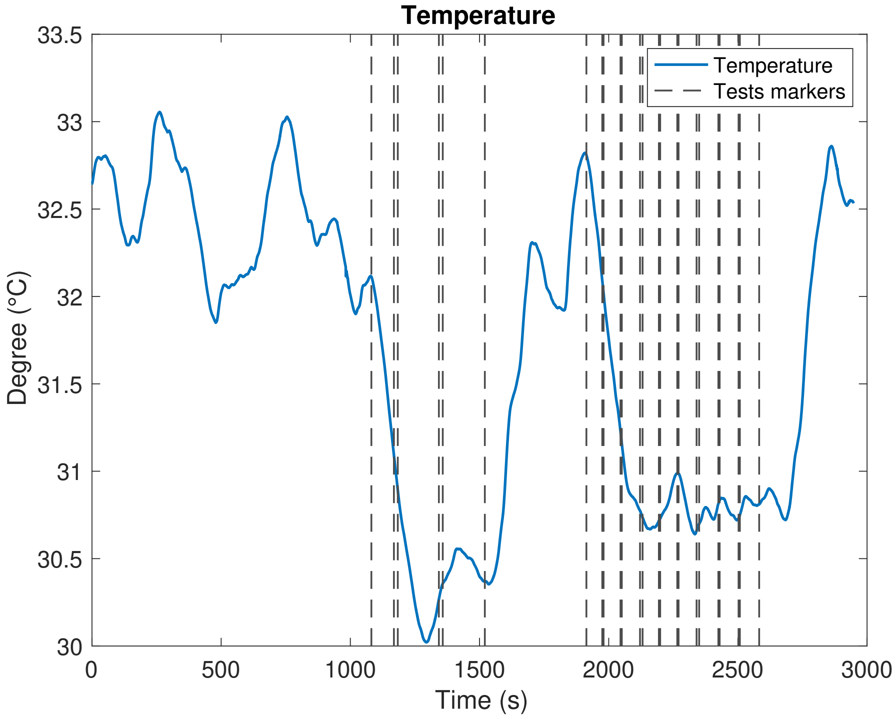

Figure 5.9: Whole temperature signal over time.

From each phase of the test (each rest and each test level) the following features (Figure 5.10) were extracted:

- the **initial value**,
- the **final value**,
- the **temperature variation**,
- the **mean value**,
- the **standard deviation**,
- the **slope**.

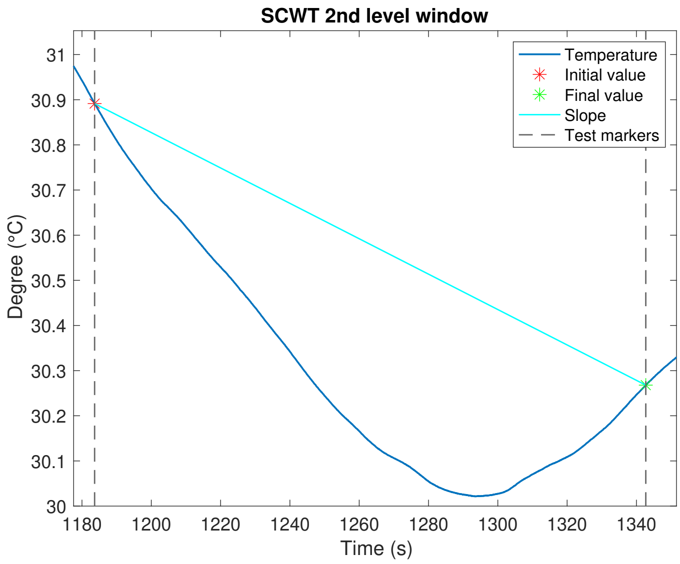

Figure 5.10: Temperature during the second SCWT level.

The same features were extracted from the first derivative, to check whether it contained useful information too.

5.4 Statistical Analysis

The thesis aims to individuate the best-performing features for evaluating stress and workload, which are those showing significant variation through the test phases since these were structured to arouse defined responses in each. To do so, particular attention was given to the comparison between the subjects' features' values in each BiLoad Test level, and their TS indicated in the SAQ.

The mean values of each feature (specified in table 5.1) were extracted from each signal feature extraction script in each window the signal was divided into. Then, a features matrix was created by calculating the mean value between the windows belonging to the same phase and collecting the resulting data. Except for the standard deviations, which were normalized over the mean value of the associated feature, each feature was first normalized over the maximum value for each subject.

This procedure repeated for 28 individuals resulted in the extraction of a three-dimensional matrix, rearranged so that in the first two dimensions were the subjects and the test phases, and in the third features.

The test was divided in 15 phases:

1. Rest 1
2. SCWT 1
3. SCWT 2
4. SCWT 3
5. Rest 2
6. Visual N-Back 1
7. Visual N-Back 2
8. Visual N-Back 3
9. Acoustic N-Back 1
10. Acoustic N-Back 2
11. Acoustic N-Back 3
12. Dual N-Back 1
13. Dual N-Back 2
14. Dual N-Back 3

15. Rest 3

Afterwards, the phases *Rest 2* and *Rest 3* were excluded, leaving 13 phases.

The investigation was performed for all features extracted from the subjects' signals, first over the whole pool of subjects, then within each subject, among their features. The best features were selected and correlated with the **test score** (**TS**), calculated on the answers given by the subjects to the self-assessed questionnaire. To better visualize their overall trend, signals were also plotted in a boxplot and compared to the TS.

5.4.1 Correlation between individuals

From the three-dimensional matrix obtained from the features extraction, a correlation analysis was performed for each feature, between all subjects' features trends during the test phases. The result of each correlation was represented in a heatmap and a boxplot for each feature. Some example are shown in the following figures. For the Heat Map, there is a distinction to be made between the tests, since this study doesn't compare the SCWT with the N-Back result. Therefore, the submatrices to observe in each Heat Map are indicated in the figure.

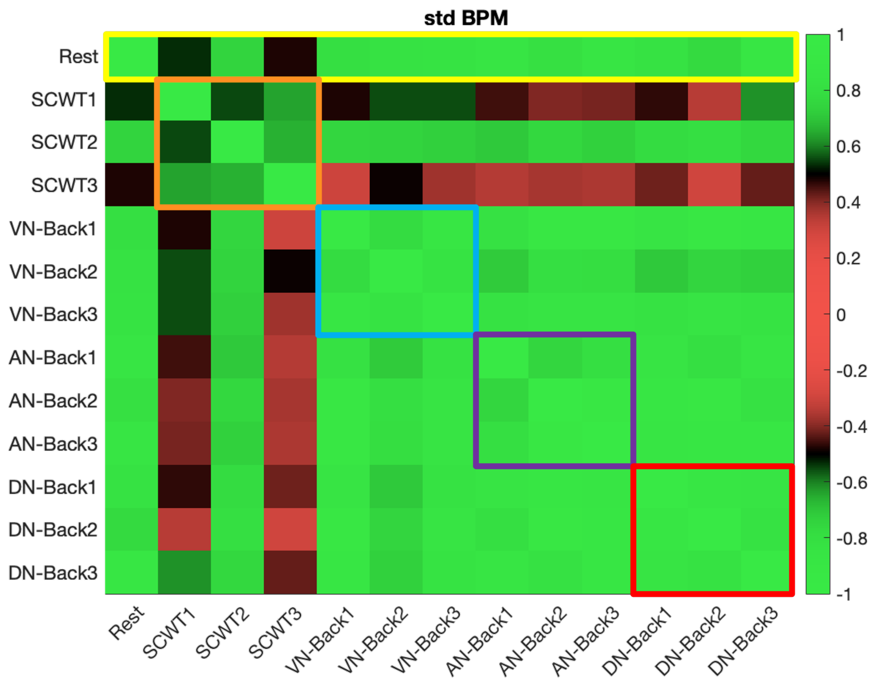

Figure 5.11: Representation of the submatrix of interest contained in the Heat Map.

In particular, there are:

- the Rest matrix, in yellow;
- the SCWT matrix, in orange;
- the Visual N-Back matrix, in blue;
- the Auditory N-Back matrix, in purple;
- the Dual N-Back matrix, in red;

SCR standard deviation This signal shows a good pattern of correlation among the SCWT phases, reaching a value of 0.91 outside the main diagonal, a little lower in the visual and acoustic N-Back, up to 0.69 and 0.65 respectively, and between 0.37 and 0.59 in the dual N-Back. These values are represented in a Heat Map (Figure 5.12), to better visualize the difference between phases.

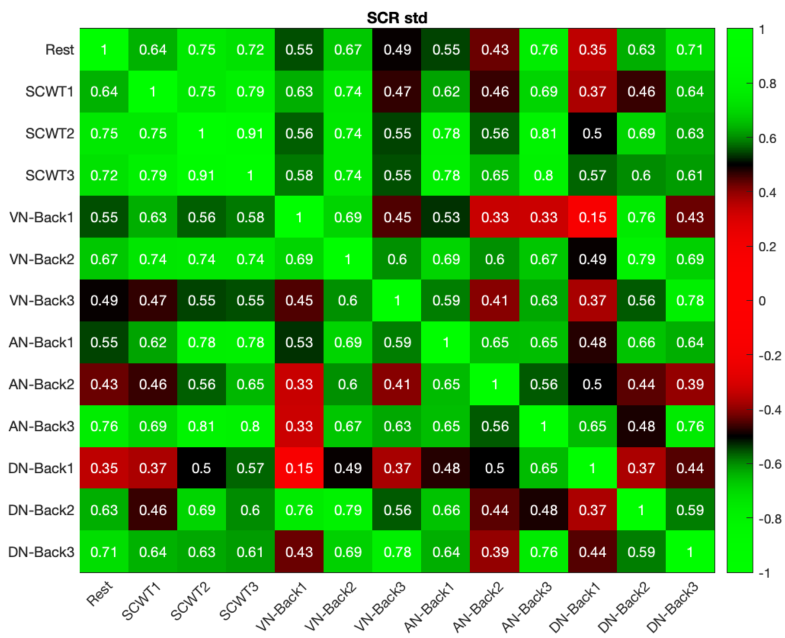

Figure 5.12: Heatmap of SCR std correlation among subjects.

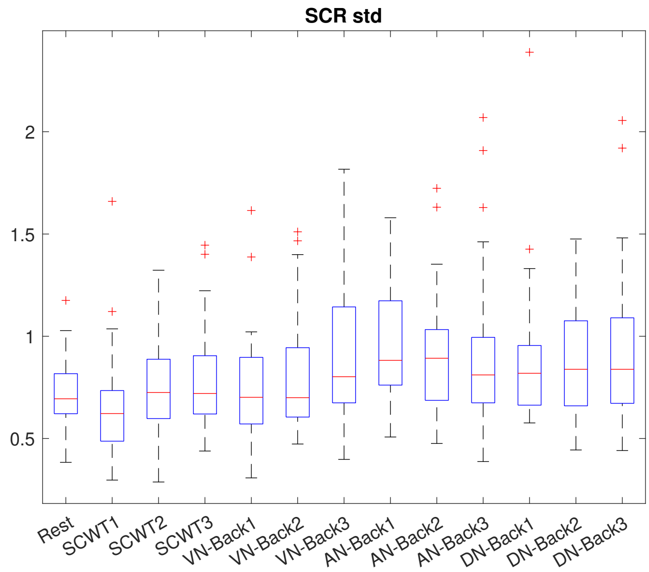

Figure 5.13: Boxplot of SCR std correlation among subjects.

From the boxplot (Figure 5.13), it is possible to visualize the global trend of the feature followed by the population, with a rapid increase from the first to the second SCWT level, and a small decrease between the second and the third. During the N-Back it shows a global increase from the start to the end, with lower values in the first two levels of the Visual N-Back, and a rapid increase in the third. In the Acoustic Test, the value reaches its global maximum, then it lowers before the Dual Test start. This last test shows a pretty constant SCR standard deviation value.

BPM standard deviation It is possible to notice high levels of correlation (Figure 5.14) for each of the four tests, with values around 0.64 for the SCWT and between 0.78 and 0.92 for the N-Back. From the boxplot (Figure 5.15), a more defined trend emerges: during the whole SCWT there is a decrease of the value, reflecting what was expected since with a distressed condition HRV lowers. Regarding the N-Back, the mean values oscillate among the tests levels, showing a decrease in the last level of the visual and dual test, and in the second level of the acoustic. Le lowest value reached at the end of the Dual Test may suggest that in those conditions the subject was perceiving high values of stress due to the difficulty of the test and its duration.

Figure 5.14: Heatmap of BPM std correlation among subjects.

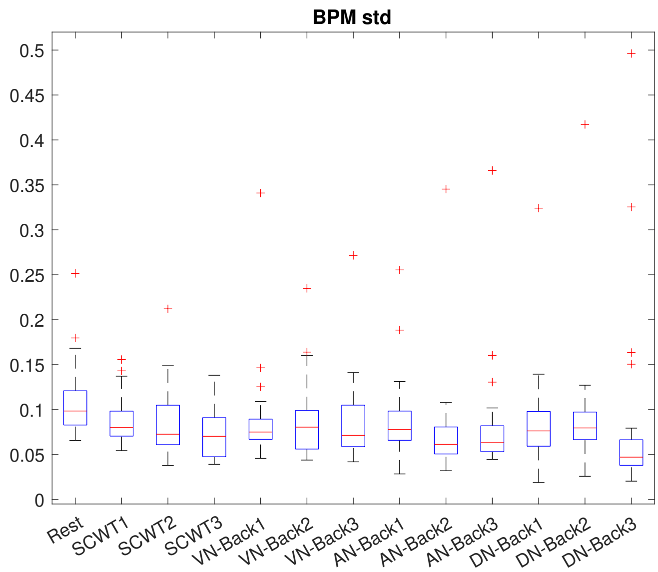

Figure 5.15: Boxplot of BPM std correlation among subjects.

Sympathovagal balance LF/HF shows a strong correlation among the population within each test. (Figure 5.16).

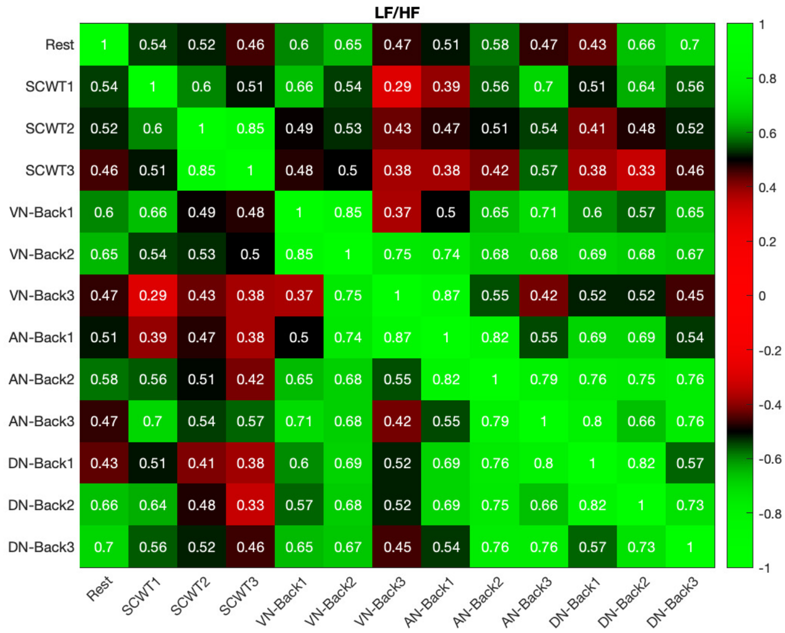

Figure 5.16: Heatmap of LF/HF correlation among subjects.

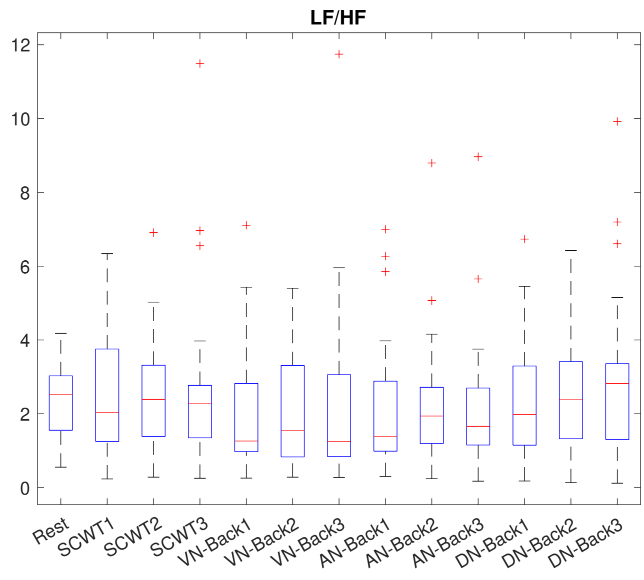

Figure 5.17: Boxplot of LF/HF correlation among subjects.

From the boxplot (Figure 5.17), in particular, it can be seen from the whiskers' width how the subjects reacted differently, but the overall trend shows a decrease with respect to the rest phase, except for the lasts levels of the Dual N-Back. This can be justified by the fact that, as stated in the SAQ, many participants were too tired or the test was perceived as too difficult to stay concentrated.

Final Temperature The correlation between subjects' trend during the test for the final Temperature of each phase, shows a very high level of correlation for all four tests too (Figure 5.18).

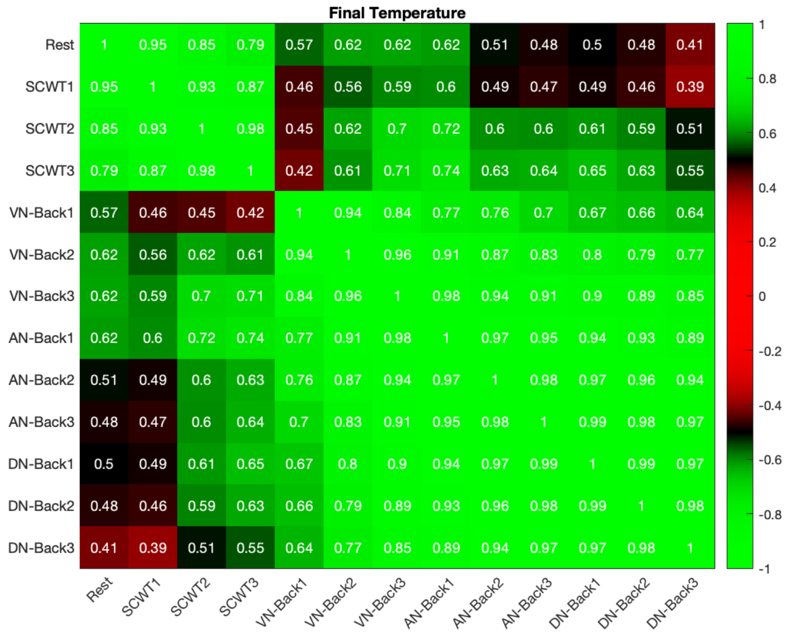

Figure 5.18: Heatmap of Final Temperature correlation among subjects.

From the boxplot (Figure 5.19) it appears that the global value of temperature decreases a the beginning of the SCWT, increases until the half of the acoustic N-Back, to eventually decrease up to the end of the test. It has to be noticed too that these values are distributed almost over the whole range, indicating a very different subjective trend.

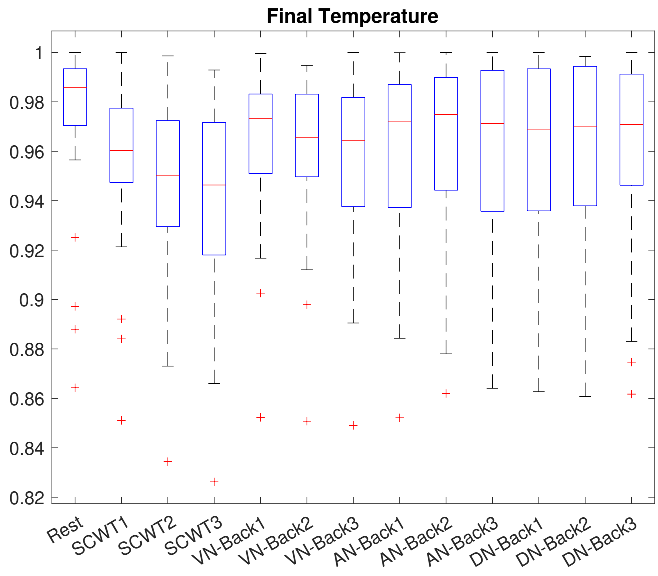

Figure 5.19: Boxplot of Final Temperature correlation among subjects.

Features selection

After the observation of each feature's trend over the test, the following step was to establish which features performed the best. To do so, a threshold of 0.4 was set to select the features whose mean absolute value of correlation was high enough: it was calculated among the levels of each of the four tests and between rest and all the test levels (rest, SCWT, Visual N-Back, Acoustic N-Back, and Dual N-Back) (submatrices shown in Figure 5.11. As the mean value of the N-Back test, it was considered the mean of the means of the three tests. For each of the four correlation matrices, the mean value, excluding the main diagonal because filled with ones, was calculated as follows:

$$Mean(phase) = \frac{\sum |corr(i,j)|}{N}. \tag{5.6}$$

Where:

- phase: Rest, SCWT or N-Back Test (Visual, Auditory or Dual);
- i, j: phases correlation coefficient, with $i \neq j$ to exclude the diagonal;
- N: number of considered elements, excluded those on the diagonal.

After this, the features that performed over the imposed threshold were evaluated in Table 5.1.

The features performing over the threshold in at least two out of three phases were considered as the *best*.

Afterward, the correlation between the best-performing features and the subjective TS of all the participants was evaluated. In particular, in the Heat Maps in Figure 5.20 and Figure 5.21, all the diagonals from the performed correlation are shown, since I wanted to evaluate the correlation between the perceived difficulty on each determinate phase, therefore the other values in output are not significant for this study. For a better understanding of the results, the values of the SCWT and N-Back are arranged in separate figures.

		Rest	SCWT	NBack
SCL	mean		x	x
	std	x	x	x
	slope		x	x
SCR	mean		x	x
	std	x	x	x
	slope		x	x
	# peaks/min		x	
	rise time			
PPG	mean amplitude			
	std amplitude	x	x	x
	rise time		x	x
	mean IBI		x	x
	std IBI	x	x	x
	rms IBI		x	x
	pNN50	x	x	x
	mean BPM			x
	std BPM	x	x	x
HRV	Ptot		x	x
	LF norm	x	x	x
	HF norm	x	x	x
	LF/HF	x	x	x
Temperature	initial value		x	x
	final value	x	x	x
	difference		x	
	mean		x	x
	std			
	slope		x	
Temperature First Derivative	initial value			
	final value			
	difference		x	
	mean			
	std			

Table 5.1: Best performing features obtained by comparison between subjects.

Figure 5.20: Correlation between features and test difficulty in the SCWT.

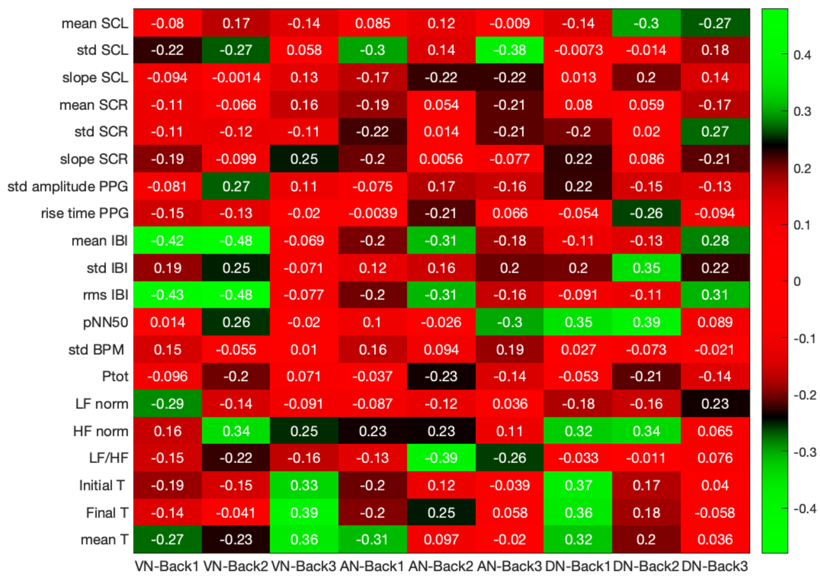

Figure 5.21: Correlation between features and test difficulty in the NBACK.

It is possible to observe that, in both figures 5.20 and Figure 5.21, the correlation hardly reaches the 0.5 value. Even if it is partly true, unfortunately, this is due to the fact that each subject reacted pretty differently, as it is possible to notice, for example, from the final temperature boxplot (Figure 5.19) standard deviation. Moreover, the difficulty assigned to the test is only expressed through an entire number ranging between 1 and 5, with a general evaluation distributed on three different scores for each level. Therefore, subjects showing the same variations in their signal but different evaluation of the test level, are classified with a low value of correlation. Anyhow, an absolute value comprised between 0.3 and 0.4 can be considered significant enough, taking into count the small pool of people, the few points considered (score values), and the high variability. Further comments can be found in the Conclusion chapter.

5.4.2 Correlation within individuals

Afterward, the correlation of each subject's feature response over the test was evaluated. The same matrix was employed, but in this case, it was transposed to have in the rows the features and in the columns test phases, while the third dimension was occupied by subjects. The obtained result is shown in Figure 5.23 for one of the subjects.

Figure 5.22: An example of a subject's features correlation over the phases of the test.

This result, again, reflects the fact that the correlation itself is not sufficient to analyze the trend of the variables because insignificant or inexistent changes between phases are also recognized as highly correlated. To better visualize this concept, an example of subject's normalized features is shown in Figure 5.23.

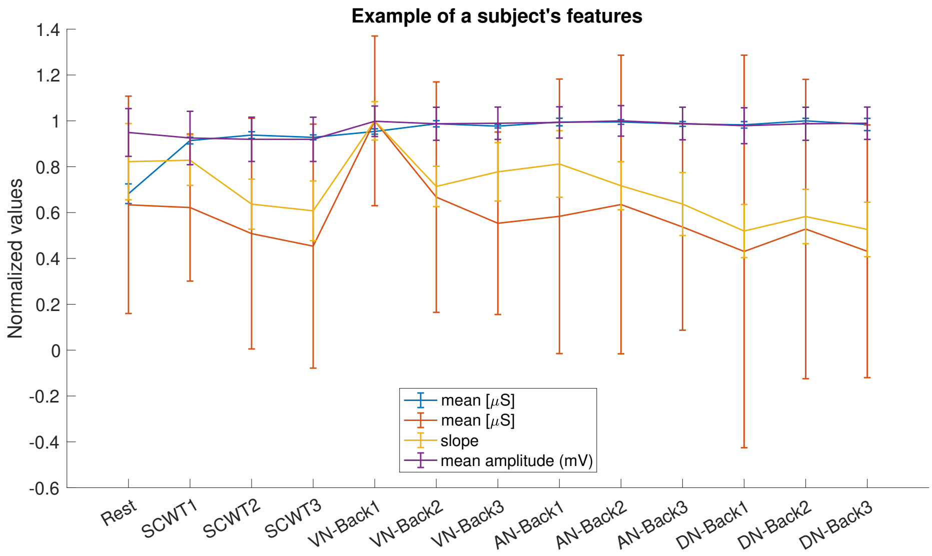

Figure 5.23: An example of some subject's normalized features over the phases of the test, including rest.

Here, it is possible to notice how features considered performing in the evaluation over the whole population actually show little variations among phases.

5.5 Test validation

To verify that the subjects really perceived the level of workload declared in the SAQ, I chose to evaluate, in addition to recorded signals, test performances both in terms of correct answers and reaction time. They are showed in figures 5.24 and 5.26.

It is possible to notice a little decrease in the population's reaction time, also reflected in a small increase in the general test performances, in the SCWT that, as stated in the SAQ, demonstrate how the subject went through adaptation to the audio stressor, while the test difficulty was the same as the second level. So, if someone was distracted or stressed by the voice, the majority of people could isolate themselves from the noise and perform almost equally in the last two levels of the SCWT. On the contrary, the N-Back test shows correspondence with the evaluation in the reaction time but discrepancies with the performances.

Reaction time against test score Since performance also contributes to the generation of stress, especially in the case of poor performance, in order to give more weight to incorrect responses, the maximum time available to respond was set as the reaction time for wrong answers. The result of this operation is represented in a boxplot format (Figure 5.24) to assess the overall participants' performance.

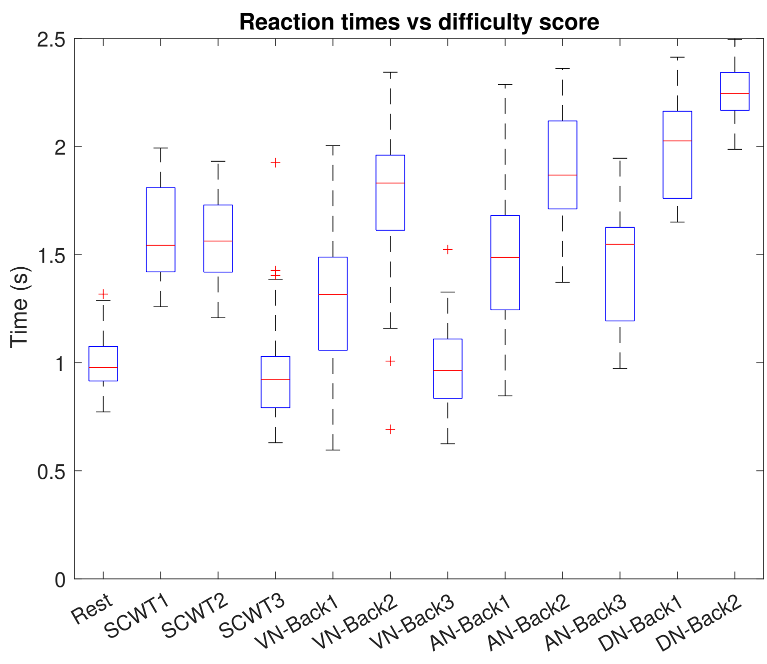

Figure 5.24: Reaction time vs. tests levels.

The same plot is reported in Figure 5.25 with over-plotted the mean values of the TS given by the subjects in the same range of the reaction time (0 - 2.5 rather than 1 - 5), to show the similarity between the reaction times and the perceived TS. From this, it is clear the closeness between the real and the perceived difficulty.

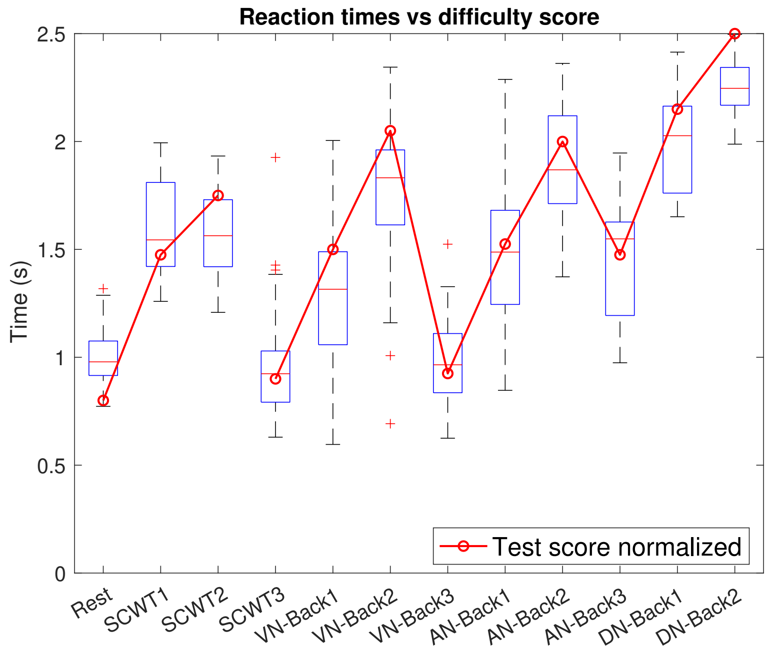

Figure 5.25: Reaction time vs. tests levels, with score over-plotted.

To further confirm the correspondence between perceived WL and test reaction time, the correlation between the mean results across the entire population was evaluated, resulting in a value of r = 0.9779.

Test performance against test score Test performance, on the contrary, doesn't show such a strong correlation with the overall reaction time, resulting in r = -0.6622. If we consider separately the SCWT and NBack correlation with the TS, instead, the correlation values appear to be much more significant for the first test rather than the second, with results equal to r = -0.9251 and r = -0.6586 respectively.

These values, accompanied by the boxplot of the test score for each level, could be interpreted by the fact that, when perceived for a long time, workload leads to decreasing performances and errors, especially when sided by increasing stress due to the wrong answer buzz noise.

Figure 5.26: Test results vs. tests levels.

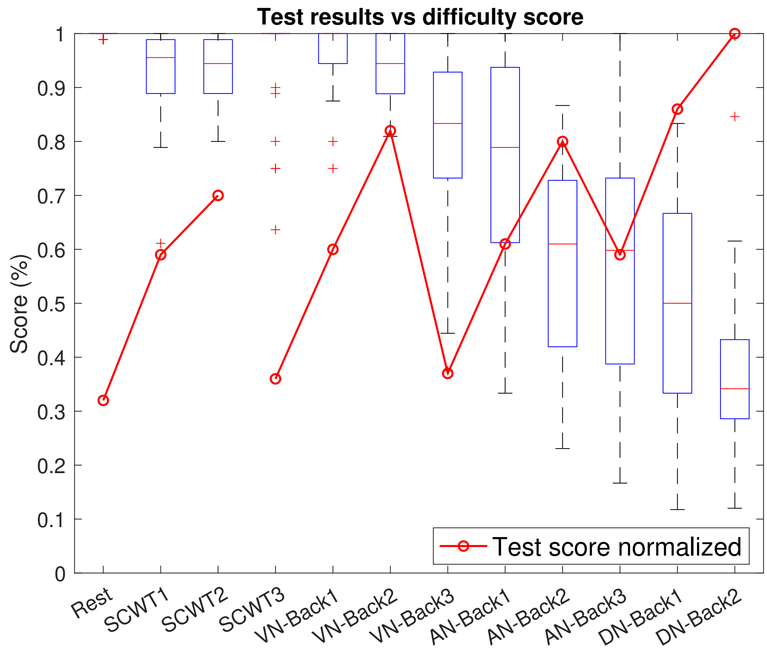

Figure 5.27: Test results vs. tests levels, with score over-plotted.

As for the reaction times, the plot with the TS is reported in Figure 5.27. In this case, the correlation results are negative since there is a lowering in performances with increasing difficulty.

From these graphs, there is a similarity in trends more in the SCWT than in the N-Back Test, even though, if considered separately, the three tests composing the N-Back reflect the trend of the green line. Again, this result may reflects more the lowering performance related to workload rather than stress, which still caused an additional decrease.

Chapter 6

Conclusion

6.1 Considerations

This study aimed to the identification of the features that globally best represent emotional arousal due to high stress and workload levels. The application was built thoughtfully considering the variation, on the physiological level, that needed to be evaluated: the Stroop test showed an increasing stress generation from different sources such as the short time given to answer, the noise of the clicking clock, the variation on the answer positioning and so on. In the same way, the N-Back test was planned to propose to the participants increasing difficulty through the three levels of each test, knowing that the tests weren't comparable on the same level: to confirm this, the TS showed an increasing value within each test, but a defined decrease in difficulty between the end of one test and the beginning of the following.

The analysis of the results was run through the correlation method, to try and find a similar trend in the features' variation among the whole pool of subjects during each phase of the test. From this analysis, the best-performing features were selected based on a threshold system, and the feature trends were also observed through a Box Plot representation. The amount of information that emerged from this study is surprising, even about aspects that weren't taken into consideration, such as the difference between visual and auditory memory, emerged from the test results described in the following section.

6.1.1 About the features

Out of the 35 extracted features through the signals analysis, 20 of them were classified through correlation as well-performing in the evaluation of global variations in physiological signals among subjects. Features, like the sympathovagal balance, showed their strength in the representation of Sympathetic Nervous System activation due to arousal, in particular in the Stroop test. More difficult instead was to define an appropriate reflection of the perceived workload, since the subjective perception factor plays an essential role, therefore it is difficult to distinguish in the global observation which subjects showed a pattern and which didn't. Moreover, the stress component affected the results of the N-Back test too. It is suggested a more thorough investigation with a more detailed techniques to take into count not only the global trend but the individual weight too, to better discriminate patterns of physiological response to external stimuli. Some examples could be the employ of ANOVA the statistical analysis or the development of a neural network.

6.1.2 About the test

Signal analysis and the test performances accurately reflected the test structure, demonstrating that the BiLoad achieved the generation of both stress and high workload, respectively through the Stroop Color and Word Test and the N-Back test.

Even though the majority of subjects weren't much affected by the acoustic stressor in the last SCWT level, it managed to put people under stress and generate appreciable variations in their signal and emotional state.

On the contrary, the N-Back test showed through people's scores as if it subjected them to increasing difficulty, which was the purpose, but such linear growth was not expected between the different tests, considering that each of them started with a level considered *easy*. Regarding the Auditory N-Back, this result could be justified by the fact that acoustic memory may be weaker than visual[33]. In general, a possible explanation for this linear decrease in performances may be the increasing perceived workload due to the test duration in addition to possible increasing stress.

6.1.3 About the questionnaire

As demonstrated in the previous chapter, the test score (TS) given by the subject represents incredibly well the reaction times but not as well the performances, especially for the N-Back test.

There are two considerations to be made: what the subject indicated as TS may better represent the perceived *stress*, which didn't critically affect their performance in terms of reaction time since it was too short a period to have negative consequences on the person, considering also that in the *fight or flight* conditions, the reactiveness of the individual is not negatively affected.

On the contrary, workload showed a greater correlation to performances rather than reactivity: it affected enough peoples' test results since they couldn't follow the tasks very well, even at the levels they perceived as easy, maybe because they didn't get enough involved or lost focus.

Anyway, the inability to identify an unambiguous parameter to represent the test score did not affect this study, even though maybe, for future evaluations, it could be more appropriate to ask two separate questions to evaluate differently stress and workload. From the observed result, two possible questions to separately evaluate these two factors could be:

- How difficult was it to answer the question in terms of time?
- How difficult was it to get the correct answer?

6.2 Future development

This thesis is just the beginning of what can be achieved through biological signal analysis, both for stress and cognitive workload perception, but also in general, as emotional arousal and psychophysiological conditions reflect in people's signals.

The next steps that can be carried out starting from this thesis work are a deeper investigation of the level of correlation between the different features extracted from the signals and the emotional state, and the aiming at the automation of the process of Stress and high workload detection. To do so, one possibility is to implement a machine learning algorithm capable of interpreting parallel physiological signal variations indicative of the emotional states investigated, in particular through the features that in this study demonstrated their validity.

It could also be convenient the development of a more comfortable configuration, such as an accurate wearable device for the recording of these and other signals, to take in count as many indicators as possible so as not to miss out on anything. After that, this could be developed in real-time software to then find applications in the areas of safety and control of physiological parameters in a variety of fields: from hospitals to sports, from automotive to aviation to everyday life applications, such as assessment of work perception, in the workplace, study load in schools, and much more!

Bibliography

[1] Shan Gao and Lei Wang. «Effects of mental workload and risk perception on pilots' safety performance in adverse weather contexts». In: *International Conference on Human-Computer Interaction*. Springer. 2020, pp. 278–291 (cit. on pp. 1, 3).

[2] Karen Hovsepian, Mustafa Al'Absi, Emre Ertin, Thomas Kamarck, Motohiro Nakajima, and Santosh Kumar. «cStress: towards a gold standard for continuous stress assessment in the mobile environment». In: *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*. 2015, pp. 493–504 (cit. on pp. 1, 3, 5).

[3] Wolfram Boucsein. *Electrodermal activity*. Springer Science & Business Media, 2012 (cit. on pp. 1, 4, 5, 15–19, 48).

[4] Allan V Horwitz and Teresa L Scheid. *A handbook for the study of mental health: Social contexts, theories, and systems*. Cambridge University Press, 1999 (cit. on pp. 1, 4, 5).

[5] Cornelia Setz, Bert Arnrich, Johannes Schumm, Roberto La Marca, Gerhard Tröster, and Ulrike Ehlert. «Discriminating stress from cognitive load using a wearable EDA device». In: *IEEE Transactions on information technology in biomedicine* 14.2 (2009), pp. 410–417 (cit. on pp. 1, 5, 32).

[6] Win-Ken Beh, Yi-Hsuan Wu, et al. «MAUS: A Dataset for Mental Workload Assessmenton N-back Task Using Wearable Sensor». In: *arXiv preprint arXiv:2111.02561* (2021) (cit. on p. 3).

[7] Alice Baird, Lukas Stappen, Lukas Christ, Lea Schumann, Eva-Maria Meßner, and Björn W Schuller. «A physiologically-adapted gold standard for arousal during stress». In: *Proceedings of the 2nd on Multimodal Sentiment Analysis Challenge*. 2021, pp. 69–73 (cit. on pp. 3–5).

[8] George Fink. *Stress science: neuroendocrinology*. Academic Press, 2010 (cit. on p. 4).

[9] Simon Ollander. *Wearable sensor data fusion for human stress estimation*. 2015 (cit. on p. 4).

[10] Roberto Zangróniz, Arturo Martıénez-Rodrigo, José Manuel Pastor, Marıéa T López, and Antonio Fernández-Caballero. «Electrodermal activity sensor for classification of calm/distress condition». In: *Sensors* 17.10 (2017), p. 2324 (cit. on pp. 5, 19, 46).

[11] Bo Zhang. «Stress recognition from heterogeneous data». PhD thesis. Université de Lorraine, 2017 (cit. on pp. 5, 10).

[12] Kalliopi Kyriakou et al. «Detecting moments of stress from measurements of wearable physiological sensors». In: *Sensors* 19.17 (2019), p. 3805 (cit. on p. 5).

[13] Miquel Fullana, J.E. Dunsmoor, K.R.J. Schruers, H.S. Savage, Dominik Bach, and Ben Harrison. «Human fear conditioning: From neuroscience to the clinic». In: *Behaviour Research and Therapy* 124 (Nov. 2019), p. 103528 (cit. on p. 5).

[14] Antonio Cano Vindel, Juan José Miguel Tobal, Héctor González Ordi, and Iciar Iruarrizaga Dıéez. «HIPERVENTILACIÓN Y EXPERIENCIA DE ANSIEDAD.» In: *Ansiedad y Estrés* 13 (2007) (cit. on p. 5).

[15] Christopher D Wickens. «Multiple resources and performance prediction». In: *Theoretical issues in ergonomics science* 3.2 (2002), pp. 159–177 (cit. on p. 6).

[16] Mark S Young, Karel A Brookhuis, Christopher D Wickens, and Peter A Hancock. «State of science: mental workload in ergonomics». In: *Ergonomics* 58.1 (2015), pp. 1–17 (cit. on p. 6).

[17] Barry H Kantowitz. «Attention and mental workload». In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 44. 21. SAGE Publications Sage CA: Los Angeles, CA. 2000, pp. 3–456 (cit. on p. 7).

[18] Essam Debie, Raul Fernandez Rojas, Justin Fidock, Michael Barlow, Kathryn Kasmarik, Sreenatha Anavatti, Matt Garratt, and Hussein A Abbass. «Multimodal fusion for objective assessment of cognitive workload: a review». In: *IEEE transactions on cybernetics* 51.3 (2019), pp. 1542–1555 (cit. on p. 7).

[19] Brice Rebsamen, Kenneth Kwok, and Trevor B Penney. «Evaluation of cognitive workload from EEG during a mental arithmetic task». In: *Proceedings of the human factors and ergonomics society annual meeting*. Vol. 55. 1. SAGE Publications Sage CA: Los Angeles, CA. 2011, pp. 1342–1345 (cit. on p. 7).

[20] Fátima Pereira da Silva. «Mental workload, task demand and driving performance: what relation?» In: *Procedia-Social and Behavioral Sciences* 162 (2014), pp. 310–319 (cit. on p. 7).

[21] Klaus von Berlepsch Andreas Kistler Charles Mariauzouls. «Fingertip Temperature as an Indicator for Sympathetic Responses». In: *International Journal of Psychophysiology* 29 (1998), pp. 35–41 (cit. on pp. 11, 26).

[22] Kalliopi Kyriakou et al. «Detecting moments of stress from measurements of wearable physiological sensors». In: *Sensors* 19.17 (2019), p. 3805 (cit. on p. 11).

[23] E Neumann and Robert Walter Blanton. «The early history of electrodermal research.» In: *Psychophysiology* 6 4 (1970), pp. 453–75 (cit. on p. 12).

[24] C W Darrow. «Sensory, secretory, and electrical changes in the skin following bodily excitation.» In: *Journal of Experimental Psychology* 10.3 (1927), pp. 197–226 (cit. on p. 12).

[25] Michael E. Dawson, Anne M. Schell, and Diane L. Filion. «The Electrodermal System». In: *Handbook of Psychophysiology*. Ed. by John T. Cacioppo, Louis G. Tassinary, and Gary G.Editors Berntson. 4th ed. Cambridge Handbooks in Psychology. Cambridge University Press, 2016, pp. 217–243 (cit. on pp. 12, 13, 18).

[26] Marieke van Dooren, J. J. de Vries, and Joris H. Janssen. «Emotional sweating across the body: Comparing 16 different skin conductance measurement locations». In: *Physiology & Behavior* 106 (2012), pp. 298–304 (cit. on pp. 12, 14).

[27] M S Nomikos, Edward M. Opton, James R. Averill, and Richard S. Lazarus. «Surprise versus suspense in the production of stress reaction.» In: *Journal of personality and social psychology* 8 2 (1968), pp. 204–8 (cit. on p. 17).

[28] Cornelia Setz, Bert Arnrich, Johannes Schumm, Roberto La Marca, Gerhard Tröster, and Ulrike Ehlert. «Discriminating stress from cognitive load using a wearable EDA device». In: *IEEE Transactions on information technology in biomedicine* 14.2 (2009), pp. 410–417 (cit. on pp. 18, 48).

[29] Shao Hanyu and Chen Xiaohui. «Motion artifact detection and reduction in PPG signals based on statistics analysis». In: *2017 29th Chinese control and decision conference (CCDC)*. IEEE. 2017, pp. 3114–3119 (cit. on p. 20).

[30] Francesco Rundo, Sabrina Conoci, Alessandro Ortis, and Sebastiano Battiato. «An advanced bio-inspired photoplethysmography (PPG) and ECG pattern recognition system for medical assessment». In: *Sensors* 18.2 (2018), p. 405 (cit. on pp. 20, 48).

[31] Fred Shaffer and Jay P Ginsberg. «An overview of heart rate variability metrics and norms». In: *Frontiers in public health* (2017), p. 258 (cit. on pp. 22, 24, 25).

[32] Kil-sang Yoo and Won-Hyung Lee. «Mental stress assessment based on pulse photoplethysmography». In: *2011 IEEE 15th International Symposium on Consumer Electronics (ISCE)*. IEEE. 2011, pp. 323–326 (cit. on pp. 23, 24).

[33] Elizabeth Hilton. «Differences in visual and auditory short-term memory». In: *IU South Bend Undergraduate Research Journal* 4 (2001), pp. 47–50 (cit. on pp. 32, 75).

[34] Maura Mitrushina, Kyle B Boone, Jill Razani, and Louis F D'Elia. *Handbook of normative data for neuropsychological assessment*. Oxford University Press, 2005 (cit. on pp. 32, 33).

[35] JHM Tulen, P Moleman, HG Van Steenis, and F Boomsma. «Characterization of stress reactions to the Stroop Color Word Test». In: *Pharmacology Biochemistry and Behavior* 32.1 (1989), pp. 9–15 (cit. on p. 32).

[36] Colin M MacLeod and Penny A MacDonald. «Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention». In: *Trends in cognitive sciences* 4.10 (2000), pp. 383–391 (cit. on p. 33).

[37] Susanne M Jaeggi, Martin Buschkuehl, Walter J Perrig, and Beat Meier. «The concurrent validity of the N-back task as a working memory measure». In: *Memory* 18.4 (2010), pp. 394–412 (cit. on pp. 33, 34).

[38] Michael Kane Andrew Conway. «The invention of n-back: An extremely brief history». In: (2016) (cit. on p. 33).

[39] Sevinc Kurt and Kelechi Kingsley Osueke. «The effects of color on the moods of college students». In: *sage Open* 4.1 (2014), p. 2158244014525423 (cit. on p. 38).

[40] James L. Szalma and Peter A. Hancock. «Noise effects on human performance: a meta-analytic synthesis.» In: *Psychological bulletin* 137 4 (2011), pp. 682–707 (cit. on p. 40).

[41] Alberto Greco, Gaetano Valenza, Antonio Lanata, Enzo Pasquale Scilingo, and Luca Citi. «cvxEDA: A convex optimization approach to electrodermal activity processing». In: *IEEE Transactions on Biomedical Engineering* 63.4 (2015), pp. 797–804 (cit. on pp. 46, 47).

[42] Hugo Gamboa. «Multi-modal behavioral biometrics based on HCI and electrophysiology». PhD thesis. Universidade Técnica de Lisboa, 2008 (cit. on p. 46).

[43] Mathias Benedek and Christian Kaernbach. «Decomposition of skin conductance data by means of nonnegative deconvolution». In: *psychophysiology* 47.4 (2010), pp. 647–658 (cit. on p. 48).

[44] Kai Li, Heinz Rüdiger, and Tjalf Ziemssen. «Spectral Analysis of Heart Rate Variability: Time Window Matters». In: *Frontiers in Neurology* 10 (2019) (cit. on p. 51).

[45] Daniel McDuff, Sarah Gontarek, and Rosalind Picard. «Remote measurement of cognitive stress via heart rate variability». In: *2014 36th annual international conference of the IEEE engineering in medicine and biology society*. IEEE. 2014, pp. 2957–2960 (cit. on p. 52).