ColorCode

Exploring Social and Psychological Dimensions of Color

Joana Filipa Duarte Condeço

Thesis to obtain the Master of Science Degree in

Information Systems and Computer Engineering

Supervisors: Prof. Sandra Pereira Gama
Prof. Daniel Jorge Viegas Gonçalves

Examination Committee

Chairperson: Prof. Rui Filipe Fernandes Prada
Supervisor: Prof. Sandra Pereira Gama
Member of the Committee: Prof. Manuel João Caneira Monteiro da Fonseca

May 2018

Abstract

Color representation has a great impact on the way we perceive information. The psychological influence of colors has been studied to improve the usage of color in the most various contexts, such as marketing and interface design. On the other hand, personality is a field of psychology which studies human character and behavior. The relation between personality and color preference has been theorized and studied. However, current research presents a lack of conclusive results. In this work, we cross personality data, not only with self-reported color preferences, but also with neurophysiological responses, overcoming the limitations of the current research.

During this study, we used the widely known personality questionnaire NEO PI-R for Extraversion measurement, and a set of color stimuli to gather both self-reported preference and neurophysiological responses. The signals used were Heart Rate (HR) as a measure of arousal, and alpha wave asymmetry – gathered through Electroencephalography (EEG) – as a measure of valence.

The results of this study suggested some trends on general preference and excitation levels, as well as on the influence of personality in color preference, and the influence of arousal level on the preference between different Extraversion groups. These trends contribute for the improvement of color uses, in particular in the Human-Computer Interaction (HCI) field, based on personality.

Keywords: Color, Extraversion, Human-Computer Interaction, Neurophysiological data, Personality types

Resumo

A representação da cor tem um grande impacto na forma como percecionamos informação. A influência psicológica das cores tem sido estudada para melhorar o uso da cor em diversos contextos, como por exemplo, marketing ou no desenho de interfaces. Por outro lado, a personalidade é um dos campos da psicologia dedicado a estudar o carácter e comportamento humano. A relação entre a personalidade e a preferência de cor tem sido teorizada e estudada. No entanto, o estado da arte actual apresenta falta de resultados conclusivos. Neste trabalho, cruzamos dados de personalidade, não só com as preferências de cor dadas por auto-avaliação dos utilizadores, como também com as suas respostas neurofisiológicas, de forma a ultrapassar as limitações do estado da arte actual.

Durante este estudo, utilizámos o amplamente conhecido questionário de personalidade NEO PI-R para a medida de Extroversão, bem como uma conjunto de estímulos de cor para obter tanto a auto-avaliação das preferências como as respostas neurofisiolofícas. Os sinais utilizados foram a Frequência Cardíaca como medida de excitação, e a assimetria de ondas cerebrais alfa - obtida através de Electroencefalografia - como medida de valência.

Os resultados deste estudo sugerem algumas tendências em relação à preferência geral e níveis de excitação, bem como da influência da personalidade e preferência da cor, e a influência da excitação na preferência entre diferentes grupos de Extroversão. Estas tendências contribuem para a melhoria do uso da cor baseado na personalidade, em particular na área de Interacção Pessoa-Máquina.

Palavras-chave: Cor, Extroversão, Interacção Humano-Computador, Dados neurofisiológicos, Tipos de personalidade

Acknowledgments

I would like to start by thanking my supervisors, for proposing this colorful study and inspiring me with their work in the Human-Computer Interaction field. In special, to Prof. Sandra Gama, thank you for believing in me and my work, and for accompanying me throughout my Master's journey.

I would also like to thank my colleagues Daniel Mendes, Daniel Pires, Inês Santos, Maurício Sousa, Rafael dos Anjos and Tomás Alves for helping me in different aspects of this work. To the specialists in psychology Patrícia Alves-Oliveira and Patrícia Arriaga, thank you for helping me with the choice of the personality questionnaire.

To my father, Carlos Condeço, thank you for inspiring me everyday and showing me that I can do anything if I believe in myself. You are the best father (and mother) on Earth and surroundings. To Marina Sousa, thank you for all the love and support. I feel very lucky to have you and João as part of my family.

And since family is much more than blood, I want to thank my chosen family, Michelle Correia, Danilo Rodrigues and little Dobby, for making every single day better after the long hours of work, and fulfilling my life with joy and love.

Speaking of love, my sweetheart João Pereira, I feel that I can not thank you enough. Thank you for lighting my path in my darkest moments, and making me see how colorful this world is. Thank you for believing in me, supporting me, and showing me how to love myself, with all my qualities and flaws.

To my girls, Ana Salta, Beatriz Santos, Sílvia Timóteo, Soraia Meneses Alarcão and Vânia Mendonça, thank you for working hard, laughing, crying and panicking with me during the last year. To Prof. Luísa Coheur, thank you for supporting me in my first university years, and for giving me the chance to work in the most amazing room, with the most amazing people ever.

A special thank you to Soraia and Vânia, the best proofreaders I could know. You inspire me everyday with your hard work and dedication.

To all my friends who supported me in the best and the hardest times, in particular Bernardo Santos, João Murtinheira, Nudrat Asif and Rui Fabiano, thank you for everything, from the bottom of my heart.

Lastly, I want to thank all the users who participated in this study, making it possible.

Thank you.
Obrigada.

To my father, my hero

Contents

List of Figures

List of Tables

List of Acronyms

ADC Analog-to-Digital Converter

BVP Blood Volume Pulse

CIE Commission Internationale d'Éclairage

CMA Circumplex Model of Affect

CPT Color Preference Test

ECG Electrocardiography

EDA Electrodermal Activity

EEG Electroencephalography

EMG Electromyography

HCI Human-Computer Interaction

HDF Hierarchical Data Format

HR Heart Rate

HSB Hue-Saturation-Brightness

HSL Hue-Saturation-Lightness

HSV Hue-Saturation-Value

IAPS International Affective Picture System

InfoVis Information Visualization

IST Instituto Superior Técnico

LCT Lüscher Color Test

MBTI Myers-Briggs Type Indicator®

VIMMI Visualization and Intelligent Multimodal Interfaces

1

Introduction

Vision allows us to perceive, through light, the world around us and, for this reason, color has been extensively studied. Different color models were created to reproduce, in the most diverse backgrounds, an identical representation of the colors of nature.

Color is frequently used in various contexts, with different cultural meanings, being able to trigger on us different emotions. Therefore, it is of extreme importance to know how to take advantage of color in each particular context.

The connection between color and psychology, more specifically emotions, has been explored and applied to settings which can benefit from certain emotional states. In Human-Computer Interaction (HCI), color can be an advantageous factor on affective priming.

Personality is, as well as emotions, a field of psychology, and has been explored since ancient Greece, when the first personality model emerged. Thenceforth, psychologists have been trying to define a solid representation for human character and behavior. Several personality models have been created, joined and refined by personality theorists, in an attempt to reach this representation.

Various theories on the relation of color with personality have been created, serving as basis to some personality tests based on color. Since these theories and tests have been used to define individual personality, some studies were performed to prove their validation. Using different personality tests, color tests and metrics, the results diverged. However, all of them had an aspect in common – subjects were asked about their preferences.

Color in different contexts may have different meanings and showing preference for a color may include an association to a concrete context. For instance, color in clothing may be strongly connected to culture and social image, not only with the effect of the color in people's emotions.

We believe that, to study the influence of personality on color preference, we need to go beyond the rational self-report. Along this document, we will explore color and personality, as well as the studies performed to date. We propose a study that will try to clarify not only previous doubts but also answer to some questions formulated by us. Understanding the relation between color and personality may change the color usage in different contexts, in particular in HCI. Using the effects of color on different personality types may be the future of user-adapted interfaces.

1.1 Goals

Our goal with this work is **to study the influence of personality in color preference**. We focus on the relation between Introversion/Extraversion and color in terms of preference and emotional reaction.

For this end, we needed to perform an extensive user study. This allowed us to collect personality and color preference data for a robust statistical analysis. With the results of our study, we were able to study the correlation between personality and color preference, from which we intended to derive new guidelines for color uses, particularly in HCI.

1.2 Contribution

The main contribution of our work consists in finding the following suggestions on the study results:

- General trends on the most and least liked colors, and more and less arousing colors, in general.
- Differences between self-reported and valence data.
- Influence of some Extraversion facets' scores in color preference and arousal level.
- Differences of most and least liked color between Extraversion groups.
- Influence of arousal in the color preference, per Extraversion group.
- Influence of Extraversion on the amount of colors liked and the preference intensity, by Extraversion group.

Despite the lack of statistically significant results, the main reason why we were not able to derive new guidelines, we believe the tendencies found may be studied and used to improve color usages, in particular in HCI, in general and based on personality.

1.3 Document Structure

This document is structured as follows: In Chapter 2, we explore the theoretical background regarding color, personality and affective neuroscience, in order to provide a context to our research and related work. In this chapter, we explain the main color and personality models studied over time, how humans perceive color and how Introversion/Extraversion influences individuals' lives nowadays, and some applications of affective neuroscience in HCI.

Following, in Chapter 3, we present our review and analysis of the state of the art regarding the influence of personality in color preference and choices. The discussion on the advantages and disadvantages of the existing studies lead us to the Chapter 4, where we present our approach to study this influence. In this chapter, we explore in detail the data required for the solution, the interface developed to collect the data, as well as the apparatus and the software used.

In Chapter 5, we explain how our data collection was built. We explain in detail how we gathered and organized the data, how we processed it and the interpretations used. Throughout this chapter, we present our initial ideas and all the choices we made to overcome the challenges that emerged.

With the data collected, processed and interpreted, we proceed to the statistical analysis needed to reach conclusions on the influence of Extraversion on color preference, described on Chapter 6. With these results, we draw conclusions, presenting them on Chapter 7, and giving directions for future work on the researched topic.

2

Background

To explore the psychological effect of color, more specifically the relation between color and personality, a theoretical research must be done. In this chapter, we will introduce those which we consider to be the most important theoretical aspects of color, personality and emotion classification through neurophysiological signals.

2.1 Color

Color helps us distinguish the world around us. For instance, in the natural world, it may ease the task of looking for oranges in an orange tree. However, this help goes far beyond nature – color is also very important in the technological world. It helps drivers differentiate traffic signs and lights and it accelerates the process of identifying features in a visualization. This is the reason why color vision is so important in our lives.

2.1.1 Color Perception

Color is a sensation produced in the brain in response to the incidence of light on the retina of the eye [43]. In the retina there are cones and rods, which are sensitive to light – cones are active at normal light levels while rods are sensitive at low levels of light. Since rods are overstimulated in all but the dimmest light, their influence on color perception can be ignored. Hence, color vision in mainly based on cones' function. There are three types of cones, which are responsive to short, medium and long wavelengths [65].

The trichromacy theory states that the color vision depends on the activity of those three types of cones, each with different spectral sensitivities, and we need at least three different wavelengths (approximately represented by red, green and blue) to match any wavelength [66].

However, some people do not have all the cone types – that impairment is called color blindness. The most common forms of color blindness are caused by lack of the long-wavelength sensitive cones, protanopia, or the medium-wave sensitive cones, deuteranopia. Since they affect the perception of the colors red and green respectively, both of them result in an inability to distinguish those colors.

[1] http://www.color-blindness.com/coblis1-color-blindness-simulator/

Figure 2.1: Simulation of default color vision (left), protanopia (center) and deuteranopia (right) on a color wheel[1].

2.1.2 Color Representation

The evolution of technology also brought the need of representing the world as it exists using graphic displays. This representation was complemented with the basis of colorimetry, which states that it is possible to match any colored light mixing three colored lights, called primary colors [65].

Since color vision relies on the cone sensitivity to the three types of wavelengths [66], the RGB model was created and named after its primaries – the colors red, green and blue [43, 65].

Considering that gamut is the set of colors that can be reproduced by a device, the gamut of a monitor is defined by the color space formed with the coordinates of the RGB primaries. That means that every color inside the gamut of a monitor is the light combination of those three primary colors [65].

The sum of all the RGB primaries using light results in white, however that is not true when we mix paints. For instance, mixing red and green lights results in yellow color, while if we use paints we get brown as the result. For this reason, it is said that lights mix additively, while paints mix subtractively.

At a very young age, people have contact with their very first color model – RYB, based on the primaries red, yellow and blue. Although it is a good model to learn and explore color mixing, the combination of these primaries generates a very limited gamut. For that reason, printers use the CMYK model (standing for cyan, magenta, yellow, and key - black), which is also a subtractive color model [36], but with a bigger gamut.

In Figure 2.2 we can observe the primaries used in these two complementary color representations, RGB and CMYK.

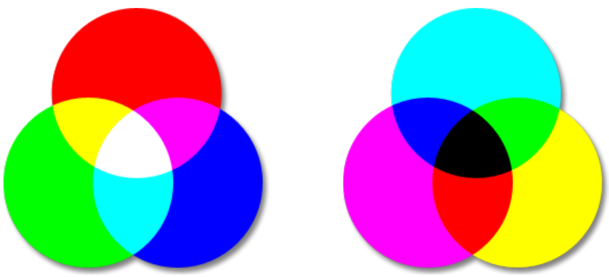

Figure 2.2: RGB (left) and CMYK (right) models' primaries and combinations[2].

Observing Figure 2.4, we can see that the gamut of a monitor is limited in comparison with all the colors that we can perceive, but the gamut of printing inks is even more restricted [65]. Although we can reproduce a larger amount of colors with the RGB model when compared to the CMYK model, it does not match our sense of the characteristics of color like lightness, saturation and hue.

To reduce the existing problem, three alternative color representations were created: Hue-Saturation-Value (HSV), Hue-Saturation-Brightness (HSB) and Hue-Saturation-Lightness (HSL). The two first ones are identical and they may be represented on a cylinder form, as we can see on Figure 2.3. On this representation, hue varies with the angle of the color wheel, saturation is the distance to the cylinder axis

[2] http://colorizer.org/

and value/brightness is represented by the height of the cylinder. In this representation, value/brightness is perceived as the amount of light of a color. On the other hand, lightness on the HSL representation corresponds to the amount of white in a color. Therefore, as we can observe on Figure 2.3, the representation is similar to HSV in terms of hue and saturation but the lightness significantly differ from value/brightness, making the bottom face of the cylinder black (no amount of white) and the opposite face white (maximum amount of white).

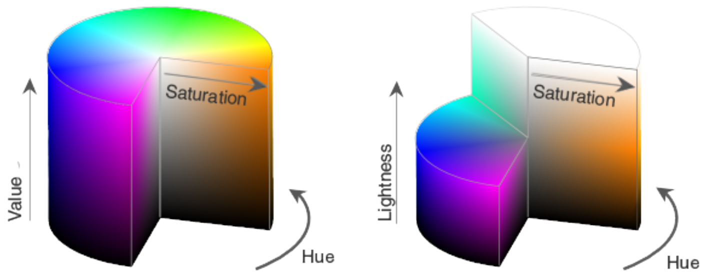

Figure 2.3: HSV (left) and HSL (right) representations[3].

These are useful tools, but not good enough when we need to specify, for instance, a colored light for rendering, since HSV conversion to RGB only produces a RGB specification [36]. For a more rigorous color specification, Commission Internationale d'Éclairage (CIE) created a system of color standards, in which gamut are included all the perceivable colors.

This system uses a set of imaginary primaries based on the tristimulus theory [66] called tristimulus values, XYZ, in which Y value corresponds to luminance. This characteristic is treated as a special information, since we often refer to lightness as an independent property of color in everyday speech. On the other hand, X and Z are usually represented by x and y in the chromaticity diagram.

The result of mixing two color lights represented by two points in the chromaticity diagram always lie within the straight line between those two points (see Figure 2.4). When we use a set of three color lights to specify a triangle in the diagram, any color within that triangle can be mixed using those color lights. Observing the image, an example of this is a gamut of a monitor being described as a triangle which vertices are representations of the red, green and blue lights.

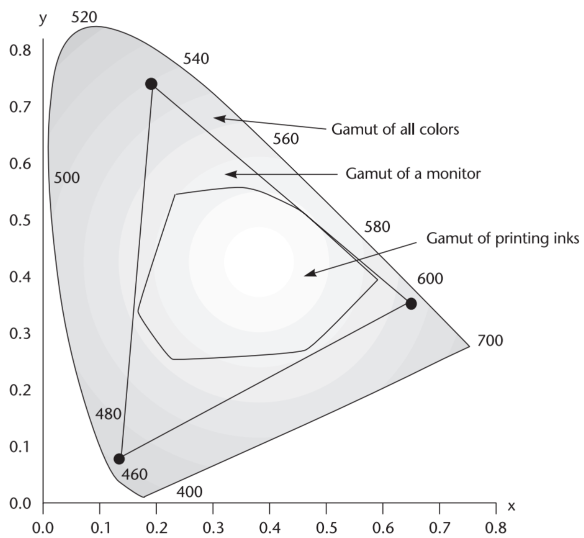

Figure 2.4: CIE chromaticity diagram [65].

[3] https://wisotop.de/hsv-und-hsl-farbmodell.php

Despite being an universal model, the CIE XYZ color space is not perceptually uniform. Due to this, the CIE produced a set of recommendations on the use of three uniform color spaces that are transformations of the XYZ color space - CIE L*a*b*, CIE L*C*h* and CIE L*u*v*. These three standards have some different properties that make them suitable for different tasks, so different industries can adopt the most appropriate standard [65].

CIE L*a*b* is based on the Opponent Process theory [35], which states that all colors can be perceived as a combination of three pairs of colors - blue/yellow, red/green and black/white. In this color model those pairs are the axes of the color representation. L is the grayscale axis, representative of the lightness, while a and b represent the pairs blue-yellow and red-green, respectively. Blue cancels yellow and red cancels green, so in the exact center of these pairs the color is gray, depending only on L.

CIE L*C*h* is similar to CIE L*a*b*, but it is represented by lightness (L), chroma (C) and the hue angle (h). It is also perceptually uniform and, as HSB and HSL, is based on how the human eye perceives color. Both CIE L*a*b* and CIE L*C*h* representations are present in Figure 2.5.

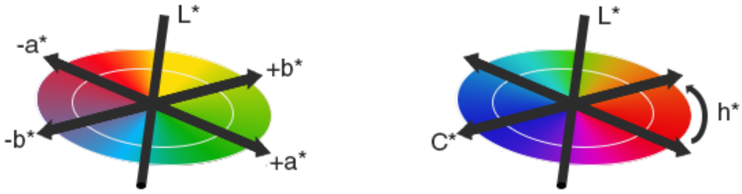

Figure 2.5: Representations of the CIE L*a*b* (left) and CIE L*C*h* (right)[4].

On the other hand, CIE L*u*v* was designed to balance the distances of the color differences in the color space. As we can see in Figure 2.6, the axes differ between models being u' and v' the result of the transformation of x and y, respectively.

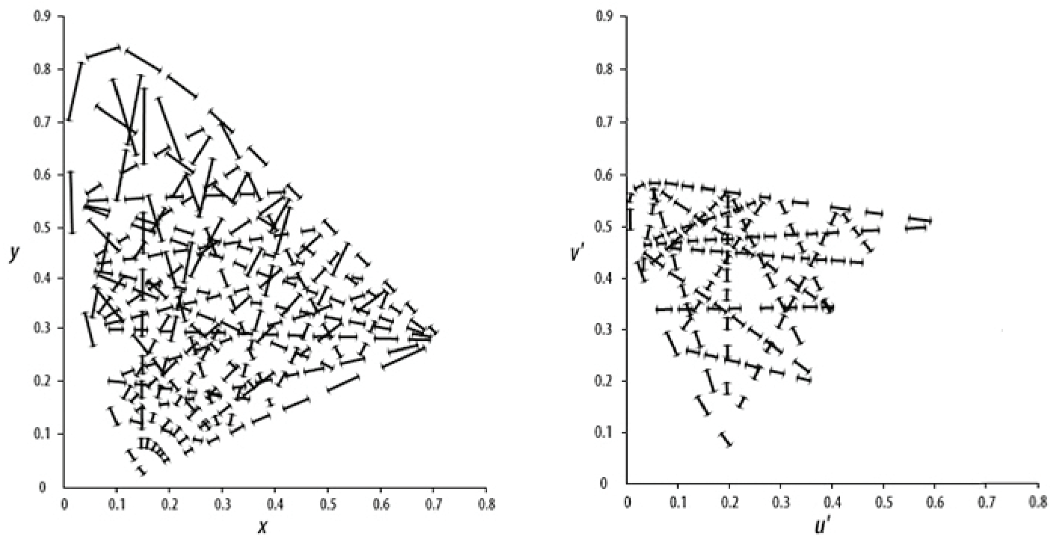

Figure 2.6: Comparison between CIE XYZ (left) and L*u*v (right) diagrams[5].

As shown, there is a variety of color representation models, specifically designed for use different uses. Although some of them are preferred for the color range which they may represent, all the diverse models are suitable for particular contexts.

[4]http://zschuessler.github.io/DeltaE/learn/
[5]http://dba.med.sc.edu/price/irf/Adobe_tg/models/cieluv.html

2.1.3 Color Applications

Although color is frequently used to improve the recognition process, its meaning goes far beyond this application. The color feature usually carries a cultural meaning and is capable of appealing and affecting people emotionally. All around the world, one color may have different meanings. Some examples of this are depicted in Figure 2.7.

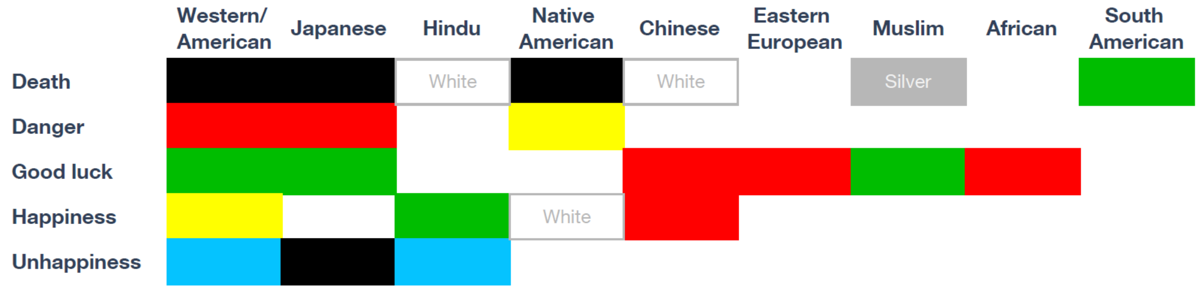

Figure 2.7: Color meanings in different cultures[6].

An example of application where color is studied in detail is marketing. Companies design their logos and color schemes with the goal of appealing to the consumers' interest and create a brand image associated to their products. To this end, it is important to analyze the cultural background and preferences of the target group, as well as focus on the message carried and emotional meaning of colors.

In Human-Computer Interaction (HCI), it is advantageous to provoke effect on certain affections such as attention, memory and performance [9]. The simplest, cheapest and most adaptable way of provoking such effect is through color. Previous studies concluded that color may specifically influence attention and, consequently, memory and performance, by increasing the users' excitement as response to a stimulus [22].

Several other studies have been building up the knowledge on the relation of color and emotions. As well as emotions, it has been proven that personality is a factor that influences users' ability to perceive information [33]. In psychology, the influence of personality on color preference has also been explored by psychologists and color theorists [3–6, 34, 46].

2.2 Personality

Communication and interaction are an essential pillar of life and since early times people try to explain the human behavior. Nowadays, we know that our actions are a result of different aspects such as our experience, individual goals, emotions and personality.

2.2.1 Personality Theories

Personality has been studied for a long time: many models were created in order to explore this dimension of the human character. These models can also be used to predict how humans are likely to behave in specific situations.

The first personality theory, known as the Four Temperaments theory, emerged in ancient Greece and was based on the Four Humor system explored by Greek medicine. This system defended that the proper function of the human body was based on four essential fluids – blood, phlegm, yellow bile, and black bile. This means that a healthy body contained all the fluids in a balanced way, while an imbalance of any of these fluids could be considered as a sign of illness.

[6]https://visual.ly/what-colors-mean-different-cultures

Nevertheless, Greek thinkers believed that these fluids were not only responsible for the nutrition and metabolism of the organism, but also had effect over mind, thoughts and emotions. That way, the imbalance of each fluid also provoked one of the four humors – Sanguine (blood), Phlegmatic (phlegm), Choleric (yellow bile) and Melancholic (black bile).

Each one of the four temperaments was directly related with a humor and had the following characteristics:

Sanguine People with a Sanguine temperament were consider as likely to be enthusiastic, optimistic, generous and sociable. They tended to be romantic and love beauty and the arts.

Phlegmatic Phlegmatic people were described as kind, sensitive, empathetic and calm. They were more likely to suffer depression caused by excessive apathy.

Choleric Known as bold, courageous, impatient and dramatic, people with a Choleric temperament were considered to be susceptible to extremism and fanaticism.

Melancholic Realism, efficiency and rigor were some of the characteristics recognized on Melancholic people, also considered to be more likely attached to material possessions, moody and withdrawn.

Although later biochemistry studies rejected the Four Humors theory, the Four Temperaments theory served as basis for other widely known personality models studied later. An example of this is the Eysenck's Personality Dimensions theory [24], used in the Maudsley Personality Inventory [25] and the Eysenck Personality Inventory [28].

Hans Eysenck developed this theory while working at a psychiatric hospital in London, where his job consisted on appraising the patients before the psychiatrists performed a full diagnose of their mental disorders. After compiling a list of questions and asking them to hundreds of soldiers treated for mental disorders, Eysenck found a pattern of connection between the answers and the patients' conditions. Based on these results, the psychologist found a group of personality traits that he then mapped using two main dimensions – Introversion/Extraversion and Neuroticism. Figure 2.8 shows the mapping of those personality traits in the axes Changeable and Emotional, respectively representative of the dimensions Introversion/Extraversion and Neuroticism.

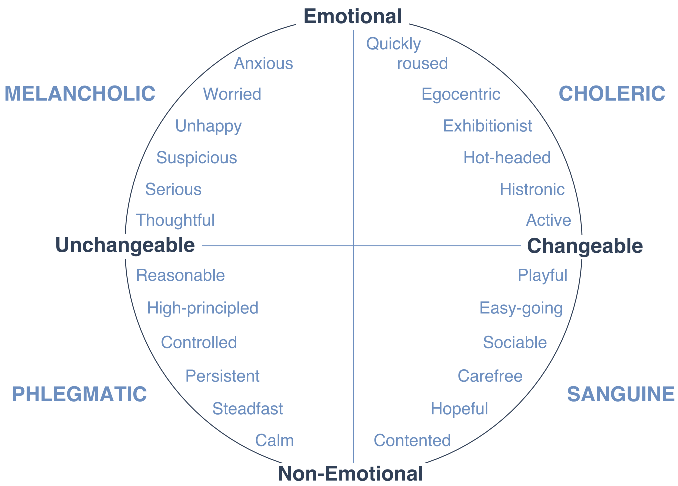

Figure 2.8: Personality traits mapped on Eysenck's personality dimensions and association with the Four Temperaments [26].

Eysenck also believed that these dimensions were correlated with human physiology, more specifically to the brain and nervous system. He also defined Introversion/Extraversion as being a dimension directly linked to arousal. The brain of Introverts is more excitable, making them more vulnerable to mood and self-protecting themselves from the outside world, causing a more reserved posture. On the other hand, Extraverts tend to seek outside stimulation to compensate the lack of brain arousal, tending also to be more lively and optimistic.

The second dimension, Neuroticism, is physically based on the nervous system, being an indicator of how nervous, anxious or stressed the person tends to be.

In later works, co-working with his wife Sybil Eysenck, the psychologist added Psychoticism as a third dimension to this model [27]. This dimension is related to people's tendency to be mentally unstable, resulting in a loner, aggressive and troublesome behavior.

Based on Eysenck's research and many other personality theories, an extensive study of personality traits converged into the most widely accepted model these days – the Five-Factor model [18]. This personality model is based on five factors – Extraversion (E), Agreeableness (A), Conscientiousness (C), Neuroticism (N), and Openness (O) –, and defends that all personality types may be defined as the vector resulting of the combination of these five dimensions.

The five factors of this model, also known as the Big Five or for the acronyms OCEAN and CANOE, are defined as follows [47]:

Extraversion A high value of Extraversion is usually associated to enthusiasm and interaction with other people, while a low value (also known as Introversion) indicates propensity to shyness, quietness and reservedness.

Agreeableness Involves aspects such as altruism, caring and emotional support at the positive end, while the negative end represents characteristics such as hostility, indifference to others and jealousy.

Conscientiousness Associated to impulse control, a high level of conscientiousness is representative of attributes such as orderliness, self-discipline and cautiousness.

Neuroticism The value of this dimension indicates the tendency to stress, anxiety and frustration, meaning that a lower level is related to relaxation and even-temper.

Openness to experience Also known as intellect, this dimension is related to traits such as imagination, adventurousness and liberalism.

However, not all the studies followed the same course. In the beginning of the last century Carl Jung created his own theory about psychological types [39]. He defended that each type of personality could be based on a type of attitude and a dominant function.

There are two major types of people – Introverts (I) and Extraverts (E) –, sorted by their attitudes towards life. An Extravert attitude is characterized by the energy exchanged with the external world, while the source and direction of an Introvert attitude is the person's internal world and thoughts.

The second factor of Jung's personality type model are the pairs of perceiving and judging functions. The perceiving functions – Sensation (S) and Intuition (N) –, represent the way the individual perceives information. A person who prefers the S function mainly believes in information gathered directly from their five senses, based on concrete facts. In contrast, a person who is more likely to perceive information via the unconscious or trust abstract and theoretical knowledge tend to prefer the N function. The judging functions – Thinking (T) and Feeling (F) – are based on the way how individuals make decisions. If a person's decisions rely mainly on what seems reasonable and logic, the function T is prevalent. The F function is preferred by the ones who make decisions based on their emotions and the empathy with a given situation.

Although these four functions are used in different situations, one of them tend to be dominant over the others. Therefore, Jung included eight psychological types in this model – IS, ES, IN, EN, IT, ET, IF and EF. These psychological types are broadly known and inspired the personality test that nowadays is, alongside the Five-Factor model, widely accepted and used in various topics - the Myers-Briggs Type Indicator® (MBTI) [48]. MBTI relies on sixteen different personality types, result of the combination of the attitude – Extraversion (E) or Introversion (I) –, the preferred perceiving function – Sensing (S) or Intuition (N) –, the preferred judging function – Thinking (T) or Feeling (F) – and a new dimension that identifies the individual preference for using the Perceiving (P) or the Judging (J) function. For instance, an Introvert who prefers the functions Intuition and Feeling and shows inclination to the Perceiving function is said to be the INFP type.

This model was created in order to help people make the best career decisions[7], using the indicator as an aptitude test, and nowadays personality tests and models are frequently used to find individual preferences in several ways, like relationships, hobbies and jobs.

2.2.2 Introversion and Extraversion

Throughout history, many psychologists and psychiatrists defined Introversion and Extraversion with slightly different meanings and effects on personality. Here, we will explore it based on modern psychology.

Extraversion is present in most personality classification systems, being considered as a continuum [39]. An individual who shows a high value of Extraversion is commonly called Extravert, while a person with a low value of this dimension is considered as Introvert.

Extraverts tend to enjoy interacting with people and the exterior, since they seek energy in the outer world. On the other hand, recharging in their own world, Introverts have tendency to be more reserved. Therefore, as Introverts get easily tired from socializing, spending time alone drains Extraverts' energy. This can be explained by a difference of sensitivity between individuals. Introverts are more sensitive than Extraverts to sensory stimulation, getting overstimulated in social environments, while Extraverts need those environments to feel stimulated.

As enthusiasts, Extraverts tend to be talkative and enjoy sharing their experiences, while Introverts are more likely to prefer thinking over talking and observing over interacting. Many introverts even have ease in socializing but they prefer not to, in order to not drain their energy quickly.

Many studies have been made to discover more differences between these complete opposites [59]. It has been proven that Introverts tend to be more alert, showing stronger reactions to stimulus change, but Extraverts are the ones who commonly have faster reaction times. A difference of preferential learning methods was also shown. While Introverts learn better by reception learning, Extraverts benefit more of discovering and exploring.

However, it is quite unfair to label individuals only as extremes of the continuum that Extraversion is. Not everyone is a natural partygoer, nor a completely closed person. Some people have their own characteristics of Introvert and Extravert, laying in the middle of the Extraversion dimension – these individuals are called Ambiverts [13].

Ambiverts are more likely to enjoy an active social life but they also need their own time and space. They love to observe and listen but also love to share their passions and experiences. Usually, meeting new people is not a concern to Ambiverts, but they tend to feel more comfortable among friends.

Ambiversion has also been observed by Jung [39], but it became a more discussed topic recently. Recent studies with Ambivert individuals have been performed in order to explore the advantages of being in the middle of the Extraversion continuum. One of the discussed studies was developed by Adam Grant and showed that Ambiverts are likely to have a better performance in sales compared to both Introverts and Extraverts [32]. This is because they express sufficient assertiveness and enthusiasm to

[7]http://www.myersbriggs.org/my-mbti-personality-type/mbti-basics/original-research.htm

persuade and close a sale, but are more inclined to listen to customers' interests without appearing too excited or overconfident.

In short, Extraversion is a very interesting and representative dimension in personality, worthy of deeper studies on correlations with other areas of interest. To study and have a better understanding of Extraversion, using it in favor of human studies and psychology, may add value to several contexts.

2.3 Preference and Arousal

Emotions are mental states provoked by external stimuli or thoughts, which play a crucial role in human cognition due to their presence and influence in human life. They have extreme importance in intelligence, perception and the decision-making process [16, 52], and they can be represented categorically or dimensionally.

The categorical perspective states that there is a set of primary emotions, resulting from neutral selection, considered universal for being independent of culture and personal experiences [15]. Many theorists have worked on defining these basic emotions, being Plutchik's [53] one of the most widely accepted. The author defends that every single emotion may be based on eight primary emotions – acceptance, anger, curiosity, disgust, fear, joy, sadness, and surprise.

On the other hand, the dimensional perspective maps emotions into two dimensions – valence and arousal. Valence represents the pleasantness of an emotion, while the excitement or activation of an emotion is represented by arousal [42, 45].

When studying emotions through the dimensional perspective, one of the most common model used is Russell's Circumplex Model of Affect (CMA) [54, 58]. In Figure 2.9, we can see how this model maps emotions into valence and arousal.

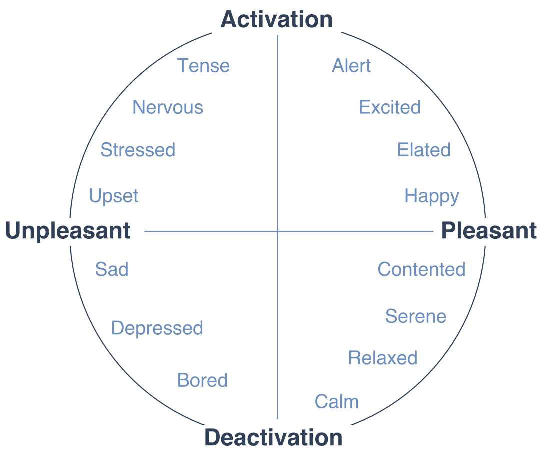

Figure 2.9: Emotions mapped into valence and arousal, following the CMA [54].

Neurophysiological signals are commonly used in emotion related research, being Electroencephalography (EEG), Electrodermal Activity (EDA) and Heart Rate (HR) three of the most used signals to measure emotions.

EEG is a fast, inexpensive and noninvasive technique based on recordings of the electrical brain activity through electrodes placed on the head. From the signals recorded through EEG, the following brain waves may be derived [1]:

Delta waves (1-4 Hz) This type of waves is associated with the unconscious mind, occurring in deep meditation and dreamless sleep stages.

Theta waves (4-8 Hz) Associated with the subconscious mind, theta waves also occur during sleep and meditation, being related to dreams and creativity.

Alpha waves (8-13 Hz) Alpha waves are associated to an aware, relaxed state of mind. They commonly occur when a person is resting but awake.

Beta waves (13-30 Hz) Related to an active state of mind, beta waves are prominent during conscious thought and logical thinking, being also associated with arousal.

Gamma waves ($>$30 Hz) Brain waves of great frequency, these are associated with an hyper brain activity.

Previous research has proved that valence may be measured through alpha wave asymmetry (using EEG) [10, 17]. On the other hand, EDA and HR are widely used to measure arousal. In light of this, we will use these neurophysiological responses to measure preference and arousal.

3

Related Work

Over the last century, many psychologists and color theorists tried to find a connection between color and various psychological factors, such as emotions, mental disorders and personality types. In this section, we explore and discuss some of the theories and studies developed in order to find correlations between color preference and personality.

Faber Birren was a color specialist known for his extensive research work on color theory. In some of his publications [3–5], the author explores the relation of color preference with mental illness and personality.

In terms of this last relation, the most well-known theory correlates color temperature preferences with Introversion and Extraversion. Considering colors like red, orange and yellow as warm colors, and green, blue and violet as cool colors, Birren defended that warm colors are usually preferred by Extraverts while Introverts prefer cool colors. The author explored these relations based on the effect that color stimulation has on humans and the way we interact with the world.

Physically, warm colors tend to increase body tension and excitation. Based on that, Birren concluded that a preference for warm colors is associated with a personality that seeks stimulation from outside and nourishes a desire to be well-adjusted to the world. On the other hand, cool colors tranquilize individuals and are therefore preferred by people who seek energy on their inner world. In addition, the author states that people who like many colors are likely to live in harmony with the outer world, while indifference to color is significant of introspective tendencies, meaning also that Extraverts enjoy bright and colorful environments, whilst Introverts feel more comfortable in sedate and conservative environments.

Some associations between colors and specific personality characteristics were also made by Birren [6]. He affirmed, for instance, that yellow is an intellectual color, being associated with great intelligence; convivial individuals tend to be attracted to orange; green is related to superficial intelligence and fluent habits of speech; purple is usually preferred by artistic people; and pink is a favorite color for people for whom the life has been a tough and harsh struggle.

Later, some other studies were conducted in order to prove the veracity of Birren's theories. The Color Pyramid Test published by Heiss et al. [34] is a well known and used test, created by the Swiss psychologist Max Pfister in order to study affect and personality characteristics. Using papers of 24 different hues, the subject is asked to fill a pyramid with 15 free spaces six times – in the first three times they must decorate the pyramids as pretty as they can, while in the next three they have to create the pyramids as ugly as they can.

The six pyramids are then analyzed and conclusions are taken from the breadth and constancy of the color choices as well as the position of the colors on the pyramids. For instance, from a flexible color choice and a moderate breadth of colors, the test indicates that an individual is more likely to be cautious and have an inhibited behavior. Another example is how, from a symmetric color organization, the authors conclude that the person has a differentiated well-structured personality [61].

In 1951, Obonai and Matsuoka [50] initiated the creation of a personality test which, based on the principle that colors have affective-symbolic properties that reflect individual experiences and personality, attempts to map personality factors into colors. A few years later, a study was made [51] in order to validate this test in terms of the gradation of abnormality, masculinity and femininity, and Introversion and Extraversion. In terms of this last factor, the study was applied to 2057 subjects of junior and senior high schools, and was divided in two steps: 1) Having a set of colors and a set of stimulus words, the examinees chose the color suitable for each stimulus; 2) Through a personality questionnaire and a teacher's evaluation, the subjects had their personality traits determined, being then classified as Introverts or Extraverts.

For each stimulus word, the colors associated were then divided and compared in two main groups of Introvert and Extravert subjects. The study showed that Extraverts and Introverts had different tendencies on the color choice for each stimulus word, proving the influence of this personal factor in color association. Therefore, according to the authors, the curve of Introversion and Extraversion diagnosed by the test approximated the normal distribution curve, proving that the test has some validity and reliability as a method of personality diagnosis.

Although it was not based on Birren's theory of warm and cool colors [3–5], Ake Bjerstedt hypothesized that the preference for warm or cool colors could indicate different personality traits [7]. This hypothesis was studied by using an alternative version of the Color Pyramid Test [34]. This consisted in a reduction of the complexity of the original test, by predefining 160 pyramid patterns from which the subjects choose their preferred one.

On a group of 603 individuals, from preschool children to university students, the study showed a clear preference for warm color patterns among the younger ones and cool color patterns among the older ones. On a smaller group (30 university students), studied more intensively, Bjerstedt found that individuals preferring warm color patterns were more likely to have greater ease in activation, as well as they tended to be more open and receptive, while cool color selectors tended to be more closed and selective.

Choungourian also explored the hypothesis that Introverts prefer cool colors while Extraverts prefer warm colors [11], in order to prove Birren's theory [3–5]. The author performed a study with 120 students (between 17 and 28 years old), which consisted in performing three tests: 1) Questionnaire of the Extraversion scale of the Maudsley Personality Inventory [25], in order to get the personality traits of the the individuals; 2) A color blindness test; 3) The color preference stimuli test with the eight Ostwald hues [38] – red, orange, yellow, leaf green, sea green, turquoise, blue and purple.

The analysis was made by separating the subjects in terms of the Extraversion scale, in three main groups – Introvert, Normal and Extravert –, and gender. The results showed that, except for male Extraverts, all groups preferred cool colors. Even so, a tendency was registered for Extraverts to prefer warm colors more than Introverts. Although it was an interesting study, it was not enough for the author to take relevant conclusions about the color preference difference between Introverts and Extraverts.

One of the most widely known and used tests that relate color preference and personality traits is Lüscher Color Test (LCT) [46]. This test uses the assortment of predefined colors to analyze individual psychophysical state, as well as the ability to withstand stress, to perform, and to communicate.

The test consists in showing the subject a set of eight different colors, which can be grouped into basic colors – blue, green, red and yellow –, and auxiliary colors – violet, brown, black and gray. The subject is then asked to sort these colors by order of preference and the result is calculated based on the color arrangement of the answer and the meaning of each color, defined by Lüscher.

For the complete performance of the test, the subjects must answer it twice, but some researchers opt for the shortened version in their studies, performing it once. The reliability of the test has been supported by many and questioned by others, in different aspects. Some of those studies will be introduced later in this section.

In 1975, Robinson performed a study to know if color preference could be correlated to Extraversion [56]. The study was designed based on Birren's color temperature hypothesis [3–5], testing the same hypothesis for LCT about the traits Concentric and Ex-centric.

Forty subjects (from University of Colorado) were part of this study composed by four phases: 1) The Eysenck Personality Inventory [28], to obtain the Extraversion score; 2) A color preference test with 28 colors of varying hues and saturations, arranged in pairs, where for each pair the subjects had to indicate their favorite one; 3) The shortened version of the LCT; 4) A self-rating test, where the subjects were asked to rate themselves in a scale of Extraversion, from zero to ten.

With the results of this study, Robinson showed that there is a significant relation between the color choices and Extraversion, supporting Birren's hypothesis. In terms of personality score, the results were also positive on the relation between Eysenck Personality Inventory scores and the self-score made by the subjects on Introversion and Extraversion. Lastly, the author claims that no differences, correlations or trends were found about Introversion or Extraversion based on the LCT.

Faced with opposite opinions about the validity of the LCT, Donelly developed a study [21] as an attempt to establish the concurrent validity of the test by correlating scores with those on the Taylor-Johnson Temperament Analysis [63]. For this, 98 college students performed the shortened version of the LCT, followed by the Taylor-Johnson Temperament Analysis.

Based on the results of this study, no relation was found between the personality descriptions present at the results of the color test and the temperament analysis descriptions. However, the results showed 81% of agreement between the color selection and the way the subjects responded to the temperament scale, showing that personality influences color preferences.

Following Donnelly's study, Braün and Bonta developed a research whose purpose was to provide an evaluation of the LCT, comparing the results with previous studies performed in America and Europe, in an attempt to prove the cross-cultural worth of the test [8]. This study consisted in two rounds of a strict administration of the LCT, the first one with 125 subjects and a second one (21 days later) with 88 subjects, all of them Canadian students.

An evaluation of the Lüscher colors was also performed by an analysis and rating of the colors in accordance with the Munsell and ISCC-NBS color-naming system. Regarding this analysis, the authors concluded that some of the colors were not pure colors and not all were named following this widely accepted color system.

In terms of color preferences, the results showed some cross-cultural and sex differences. In general, Canadian male students prefer slightly the color red over blue, contrasting with Americans' preference for yellow obtained by Donnelly [20], and the Europeans' like for red expected by Lüscher. The results also showed that the most chosen color for males was red, while yellow was preferred by females.

This led the authors to conclude that LCT is not adequate as a standard diagnostic tool, based on the cross-cultural and sex differences. However, Braün and Bonta state that the creation of a solid personality test based on color preference would be a worthy objective.

Another study on the relation of personality and color preferences was developed by David and

Miriam Stimpson [62], in order to consolidate the knowledge about this relation. The goal of this study was to focus on this relation in a "normal group", being all the 150 subjects American undergraduate students with no psychiatric or cross-cultural implications.

Two tasks were completed by the subjects. In the first one, the subjects performed the Edwards Personal Preference Schedule [23], in order to obtain the score of their personality traits. Next, they were asked to sort a set of eight color plates corresponding to the hues of Lüscher [46], according to their individual order of preference. With this test, the authors found no correlations between color preference and the personality variables present in the personality model used.

Five years later, Nelson et al. published a study on the correlation between color preference and stimulus seeking [49]. This study was mainly based on the hypothesis that the preference for color red is more common among seekers of high stimulation, while low stimulation seekers are more likely to prefer blue, following Robinson's work [56].

For this, 170 students aged between 12 and 14 years old, not color blind, were submitted to two tests. For the color preference test, the subjects were divided in six classes and two cards colored in red and blue were showed simultaneously to all the subjects in each class, having to choose between the two colors individually. To avoid influencing the color choice by position, left/right positions of the colors were alternate over classes. The Sensation Seeking Scale [67] was used as the second test of this study.

The results showed that more than 70% of the subjects preferred blue over red. It was also shown that the high and low scores on the Sensation Seeking Scale are respectively associated to red and blue. From these results, the authors concluded that red has stimulating and exciting properties, while blue has less stimulating, calming properties.

In a more specific study, Radeloff tried to find a correlation between psychological types and color preference in clothing [55]. The study was made using 111 college-age women and used as personality test the MBTI [48], focusing only on the attitude type and the dominant function. For the color preference test, 20 pieces of fabric were used, precategorized into four seasonal hue categories – Summer, Winter, Spring and Autumn[1]–, three value categories – light, medium and dark – and three chroma categories – intense, medium and dull. For all the pieces of fabric, an uniform gray matte background was used as well as a daylight incandescent bulb, in order to standardize the colors. The evaluation was made by asking the subjects to rate each piece of fabric as "I love this", "I sort of like this" or "I hate this".

The results of this study showed that about 75% of the subjects were Extraverts and, overall, cool colors are preferred much more than warm colors. This study showed that, unlike what was expected by Birren's theory [3–5] on color temperature and Extraversion, the Extraverts did not choose warmer and brighter colors, leading the author to conclude that this hypothesis is not valid in clothing matters.

More recently, Rosenbloom performed a study [57] on the relation between sensation seeking and color preferences. This study had the goal to test, among hypotheses of other natures, the hypothesis that high sensation seekers would prefer red (considered an arousing color) and sensation avoiders would prefer blue (a relaxing color).

In an attempt to relate the two main aspects of the study, 60 Israeli students with age ranked from 22 to 32 performed a sensation seeking test and a color preference test. For the first one, the authors chose the Sensation Seeking Scale [67]; for the second one, a clear painting of a human figure was used, accompanied by the colors gray, red, orange, brown, yellow, green, blue, and gray, with which the subjects had to color the figure.

The results showed a positive correlation between the scores of the Thrill and Adventure Seeking factor of the Sensation Seeking Scale and color preference. Those subjects who chose arousing colors for coloring the human figure were likely to have high scores of Thrill and Adventure Seeking, while

[1] This division was not performed in terms of the hue angle, but the tendencies for each season.

the ones who lower scores preferred to use nonarousing colors. However, no correlations were found between the color preference and the remaining factors of the Sensation Seeking Scale.

3.1 Discussion

Based on the analysis of the previously mentioned theories and studies, which rely on the relation between color and personality, there are some points to discuss about the chosen methodologies and conclusions.

Despite being widely known and inspiring many color theorists and psychologists, Birren based his theories mainly on color associations with psychology. These might have been the first big steps on this very interesting relation, but the lack of studies to prove Birren's theories leads to a deficiency of scientific validity.

The same happens regarding the Color Pyramid Test and Lüscher Color Test (LCT). These are studies used in psychology and psychiatry that were accepted before the existence of scientific tests to prove their reliability.

The study performed by Obonai and Matsuoka was of great importance, since the authors could conclude their hypothesis was correct. However, this study did not include color preference, but color associations only.

Based on color preference, and looking forward to prove the theory that the preference for warm or cool colors could indicate different personality traits, Bjerstedt's choice to use the Color Pyramid Test might have been a risky decision. However, the author was able to successfully prove his hypothesis. Following this study, it would have been interesting to map the different traits into personality dimensions, using a personality model. This way, Bjerstedt could also explore Birren's theory that relates warm and cool colors with Extraversion and Introversion. Some years later Choungourian failed to prove that theory, but a similar study performed by Robinson successfully verified its validity.

Both obtaining good results between color preferences and personality, Robinson and Donnelly showed that the LCT is not valid in terms of correlation between the color choices and the personality descriptions given as result. It is important to note that these studies did not explore other variables of the test, such as psychological disorders.

Following the invalidation of the LCT, Braün and Bonta also proved that the test results are not consistent in terms of the culture and gender of the subjects. These factors should not cause differences in personality, since they are not part of any of the most widely accepted personality models.

In summary, we can conclude that the LCT should not be used to make the correlation between color preference and personality. This conclusion was already expected taking into account the test *per se*. Considering that the result is always based on the order of the colors, a person who likes two or more colors equally may have various personality profiles, only based on the random order of choice for those colors, which is not a valid assumption.

Trying to find a connection between color preference and personality, David and Miriam Stimpson developed a study, based on none of the discussed theories, from which they concluded that the two factors were not related. We believe that the failure of the hypothesis might have been based on the personality model used for testing personality, which did not include the most used dimensions, such as Extraversion and Neuroticism.

In a more specific context, Radeloff also failed correlating color preference and personality. Besides the fact that the choice of colors in clothing is also affected by culture and fashion trends, the study is based on colors grouped by season and not by hue angle. We do not consider as relevant the fact that the study was performed only with female subjects, since no personality model defines gender as a dimension.

With similar studies, both Nelson et al. and Rosenbloom succeeded on proving that warm and cool colors are related to higher and lower levels of sensation seeking. Even considering that the color preference test of the study developed by Rosenbloom was not the best – since the association might have been based on habits or clothing preferences –, the results were conclusive.

These two studies also completely support Birren's theory on color temperature. Although the authors did not use Extraversion as dimension of study, there is a solid relation between Extraversion and sensation seeking, since these factors are both based on the interaction with the outer world.

In general, many of the studies show that the association between color preference and personality, more specifically Introversion and Extraversion, may be through color temperature. Some authors left notes in their conclusions about repeating the tests with some condition changes. This is due to the lack of rigor in some aspects, like the environment conditions or the colors used for the tests.

What we believe to be the most important factor of all the studies explored in this document is the way the subjects expressed their preference for the colors. With or without a context, the individuals were led to choose the colors rationally. The rational factor includes not include only the preference, but also the cultural and social context, as well as the emotional state. To focus on the preference only, the study would have to include detection and analysis of the effect of specific stimuli on the subjects. The detection could have been made using measurement of neurophysiological signals as responses to the color stimuli.

To study the relation between those measured values and personality dimensions, a validated personality test should be performed, followed by an extensive statistic analysis. This analysis should be rigorous and include statistically significant results, in order to confidently find a correlation between color preference and personality. Finding this correlation, some new guidelines could be derived and used in HCI, in order to encourage its usage through color.

4

Approach

In order to find an existing relation between personality – more precisely Introversion/Extraversion – and color preference, we needed to collect data related to both topics. In this chapter, we describe our approach, including the required data, the developed interface, and the equipment and software used.

4.1 Required Data

As discussed before, psychologists and color theorists developed studies trying to explore this relation but not all the results were conclusive. Based on the fact that the previous studies were not consistent in terms of personality tests nor in terms of color preference tests, we consider that the inconclusive results could be a consequence of a bad choice on how to gather data. Having this in mind, for our study we decided to collect the following data:

Profiling data To get a more detailed approach of the results of this study, we decided to collect some profiling information, not directly related to the subjects' personality, nor their color preference. Besides common demographic questions, such as Gender and Nationality, we also needed to gather data about each participant's self-reported Extraversion group, as well as their perception of color. In order to achieve the latter, we chose to perform a reduced version of the Ishihara color blindness test [37].

In a color study, user profiling and color blindness tests are crucial to have a detailed statistical treatment. In case of color blind users, the data should be discarded or used as case studies.

Extraversion score data To collect data regarding personality, it was essential to use a validated personality test, based on a widely accepted personality model. For our work, we chose the NEO Personality Inventory - Revised (NEO PI-R) [14], which is based on the Five-Factor model [18], already presented in Chapter 2. This inventory is composed by 240 questions – 48 per personality dimension –, and it divides each dimension in six different facets. Extraversion, which is the only dimension we intend to study, is divided in the following facets:

Warmth is the facet of Extraversion of most relevance to issues of interpersonal intimacy. Warmer people are affectionate and friendly, and easily form close attachments to others. Low scorers are more formal, reserved, and distant in manner than high scorers.

Gregariousness is related to the enjoyment people feel for being in company of others. Gregarious people seek social stimulation, while low scorers tend to be loners and avoid crowds.

Assertiveness is connected to domination, strength of will and social ascendance. High scorers on this facet speak without hesitation and often become group leaders, while low scorers prefer to keep in the background and let others talk.

Activity is representative of movement and energy. High Activity scorers are seen in rapid tempo and need to keep themselves busy, living a fast-paced life. Low scorers are more leisurely and relaxed in tempo, though they are not necessarily sluggish or lazy.

Excitement Seeking high scorers crave excitement and stimulation, enjoying noisy environments. On the other hand, low scorers feel little need for thrills and prefer a lifestyle that high scorers might find boring.

Positive Emotion assesses the tendency to experience positive emotions such as joy, happiness, love, and excitement. High scorers are cheerful and optimistic, while low scorers are less exuberant and high-spirited, not necessarily meaning they are unhappy.

The length of the NEO PI-R inventory is a major benefit of this test, since while performing longer tests the subjects tend to describe their actions and feelings more accurately, not being biased by what they think they are expected to answer. Another reason why we chose to use this test is because it allows us to go deeper on the analysis of Extraversion, trying to find the influence of specific facets on color preference.

Self-reported color preference data Although we were motivated to use neurophysiological data to indicate users' color preference, we considered important for this study to collect self-reported color preference data to explore the relation between rational and emotional responses to a set of colors. Besides this, collecting the self-reported preference allows us to draw a parallel with the previous works.

Neurophysicological color preference data To collect the color preference through neurophysiological data, we needed to record the participants' responses to a set of colors. To this end, we decided to gather both the neurophysiological data and self-reported data at the same time, as response to each color of the set.

As explained in Chapter 2, valence represents the pleasantness of an emotion, while the excitement or activation of an emotion is represented by arousal. Since our goal is to know the subjects' preference for each color, our main focus for this work was the valence of the emotion felt when presented to each stimulus, which can be measured using alpha wave asymmetry on frontal channels of the brain, measured through Electroencephalography (EEG).

Previous research has supported valence theories that an increase in alpha wave activity is inversely related to cerebral activation, and higher activation of the right hemisphere is associated to withdrawal reactions [10, 17]. Therefore, greater alpha wave activity in the left hemisphere means a shift in cerebral activation to the right hemisphere and, consequently, a withdrawal reaction.

Although it was not our main focus, we would like to discuss the excitation levels based on personality, so we decided to collect information about the arousal induced by the color stimuli. For that, we chose to record the participants' Heart Rate (HR), which is a well known indicator of the emotional arousal.

Another widely known indicator of arousal is Electrodermal Activity (EDA), but during the pilot tests phase we found the majority of the values were missing, when recording EDA with the used device. Since previous studies also shown that EDA values are not significant when using other arousal measures [60], we chose to use HR only to extract arousal scores.

The gathering of these data types, as well as the their processing methods will be described in detail in Chapter 5.

4.2 Interface

In order to collect the data needed to explore the influence of personality in color preferences, we developed an interface to be used during the user testing phase.

This interface consisted in a website developed in HTML, PHP, CSS and Javascript, using a Bootstrap blog theme[1], carefully changed and adapted to our needs, resulting in a coherent, simple and pleasant interface.

With the intention of collecting the users' profiling, personality and color preference data, we divided this test into 5 phases, being three of them questionnaires, one of them dedicated to the BITalino installation, and one merely informative. Each phase included an introduction screen explaining the task. When reaching these pages, the user was asked to warn the researcher, who would then explain the next task. We chose to to this predominantly through conversation, to bring the researcher and each user closer, in order to make them feel more comfortable. Comfort is very important during the user tests, since some tasks require physical touch and our goal is to minimize emotional changes induced by the test conditions.

Since the personality questionnaire used was the Portuguese version of NEO PI-R, meaning that the test could only be completed by Portuguese speakers, the full test was developed in Portuguese. Also, all the questions presented in all questionnaires are mandatory, meaning that the test pops up an error window when the users miss a question.

In the following subsections, we will present in detail the developed interface, including not only the looks but also the choices made regarding the data collection.

4.2.1 Introduction

The first screen of the interface shown to the user is the introduction screen. This screen is the first example of the layout of the informative pages that are shown before proceeding to each task of the test. These pages include a header, a brief text and a footer, as shown in Figure 4.1.

The header, which occupies a considerable space in the screen, contains a background image and the title and subtitle corresponding to the step of the test. We wanted the header background, presented in every information page and profiling and personality questionnaire pages, to be representative of the theme, so we chose an online colorful picture, free to modify and use in websites and presentations[2]. In terms of text, in this page we have the name of this work, *ColorCode*, followed by its subtitle, "Exploring Social and Psychological Dimensions of Color". On the remaining pages, the text contains the name of the task and a short informal sentence explaining its goal.

This screen also contains a brief text, which complements the introduction speech given by the researcher. We chose to use the text to complement the speech instead of describing all the steps of the test with the purpose of bringing the user closer to the researcher, making them feel more comfortable, which is an important step during this study.

[1] http://startbootstrap.com/template-overviews/clean-blog
[2] https://www.freepik.com/free-vector/colorful-background-for-holi-festival_1051050.htm

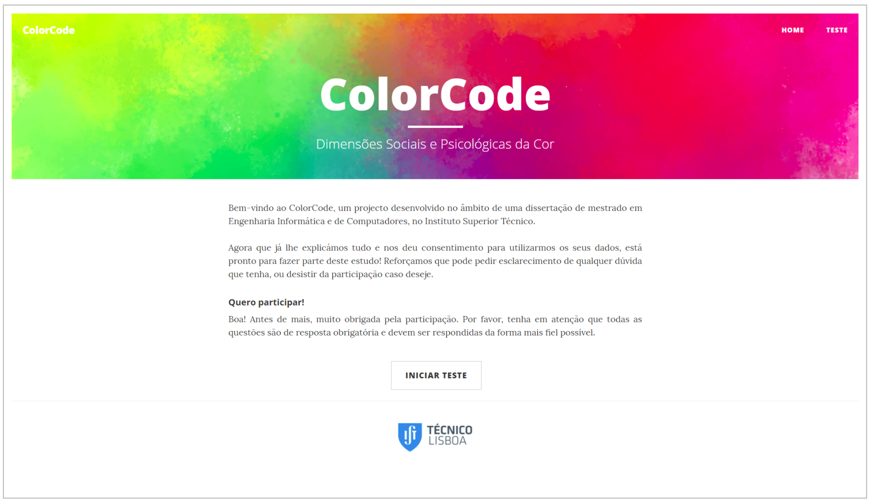

Figure 4.1: Main page of the ColorCode website.

Observing this page, we can see that there are two ways of proceeding to the test's first page: Clicking on the button "Iniciar Teste" ("Start Test"), or clicking on "Teste" ("Test"), on the top right corner of the page. Note that this second link, just like "Home", is present in all the further webpages, being a shortcut to come back to the beginning.

Lastly, in this page, as well as in all the other informative pages, and the profiling and personality questionnaires, we included a footer with the Instituto Superior Técnico (IST) logo, serving as shortcut to the institute website[3].

4.2.2 Profiling

The first task of the test consists on filling a profiling questionnaire, shown on Figure 4.2, which includes profiling questions, as well as the Ishihara color blindness test [37].

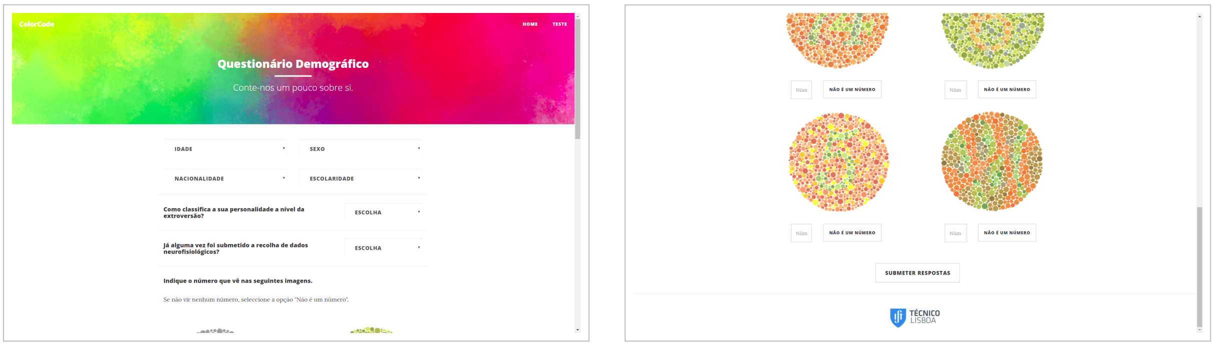

Figure 4.2: Profiling questionnaire: Profiling questions (left) and Ishihara color blindness test (right).

In the first half of the profiling questionnaire, we included basic profiling questions, presented as drop-down multiple choice questions. This type of questions has the benefits of multiple choice questions, while keeping the page clean. Besides keeping the page simple and tidy, multiple choice questions are very useful in questionnaires like this, since they ease the data division and further treatment, also reducing the time the users take to answer the questions.

[3]https://tecnico.ulisboa.pt/

For the fields which can be represented with one word only (for instance, age or nationality), we chose to use that word as an non-selectable predefined option, simplifying the fields organization through the page. For the longer questions, we wrote the question putting the drop-down field by its side, with “Select” as the non-selectable predefined option.

The questions included on the profiling questionnaire are:

Age This question may have been answered by choosing one of the following groups of age: **a)** < 18, **b)** 18 to 25, **c)** 26 to 35, **d)** 36 to 45, **e)** 46 to 55, **f)** 56 to 65, and **g)** > 65.

Gender Users were able to chose between: **a)** Female, or **b)** Male.

Nationality A list of 238 countries in Portuguese[4] was used to this end, being possible to search for the wished country using the keyboard.

Education Like age, the education level was divided into the following groups groups: **a)** No education, **b)** Basic education - 1st cycle, **c)** Basic education - 2nd cycle, **d)** Basic education - 3rd cycle, **e)** Secondary education, **f)** Licentiate’s degree **g)** Master’s degree, and **h)** Doctorate. As this test was held in Portugal and we expected the biggest part of the participants to be Portuguese, we chose to follow the Portuguese education system[5]. For users who did not study in Portugal, the equivalent education level should be chosen.

How do you rank your personality in term of Extraversion? For this question, the participants could choose between the following possible answers: **a)** Introvert, **b)** Ambivert, **c)** Extravert, and **d)** I do not know. Our goal with this question was to understand if the users had a previous idea of their personality type in terms of Extraversion, and how correct they were.

Have you ever had your neurophysiological responses collected? For this question, the users were able to chose between: **a)** Yes, and **b)** No.

The second part of the profiling questionnaire consisted in a six plate version of Ishihara’s color blindness test [37]. This test was created to detect deficiencies in color vision, having its full length 24 plates. Following the author’s instructions, the six plates of the simplified version of the test shall be: **Plate 1)** Number 1, **Plate 2)** Number 2 or 3, **Plate 3)** Number 4, 5, 6, or 7, **Plate 4)** Number 8 or 9, **Plate 5)** Number 10, 11, 12, or 13, and **Plate 6)** Number 14 or 15.

Having this into account, we chose the plates 1, 2, 7, 8, 10, and 15, represented in order in Figure 4.3, for the simplified version present in our questionnaire.

As part of the profiling questionnaire, each of these plates is a question which can be answered by typing the number seen on the plate or selecting the button “It is not a number”.

4.2.3 Personality

After submitting the answers to the profiling questionnaire, the subjects were directed to the personality introduction page, which led them to the personality questionnaire. As previously explained in the Section 4.1, we decided to use the Portuguese version of the widely known personality test NEO PI-R [14].

From this extensive questionnaire, we selected the 48 questions relative to the Extraversion dimension. Following its original form, we presented those question to the subjects, to which they could answer in a five point Likert scale [30], from “Discordo Fortemente” (“Strongly Disagree”) to “Concordo Fortemente” (“Strongly Agree”). Since this test could not be changed due to copyright, this Likert scale served as basis for the Color Preference Test (CPT), in order to maintain the study coherence.

[4] https://github.com/cristianoascari/lista-de-paises-em-sql
[5] https://en.wikipedia.org/wiki/Education_in_Portugal

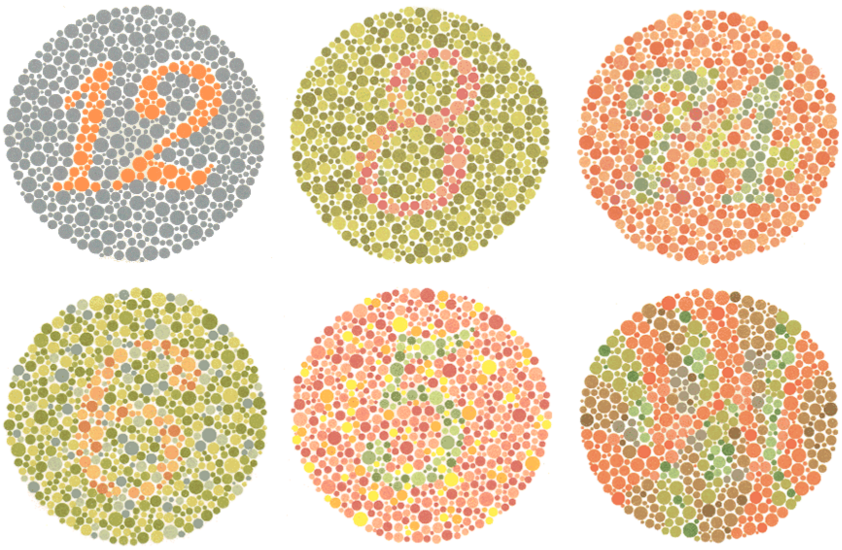

Figure 4.3: Ishihara's color blindness test - six plate version.

Each question was implemented as a multiple choice question, in which the users clicked the most adequate button to choose the most reliable answer. The questions were also numbered in ascending order, so we could easily identify missing answers and alert the users.

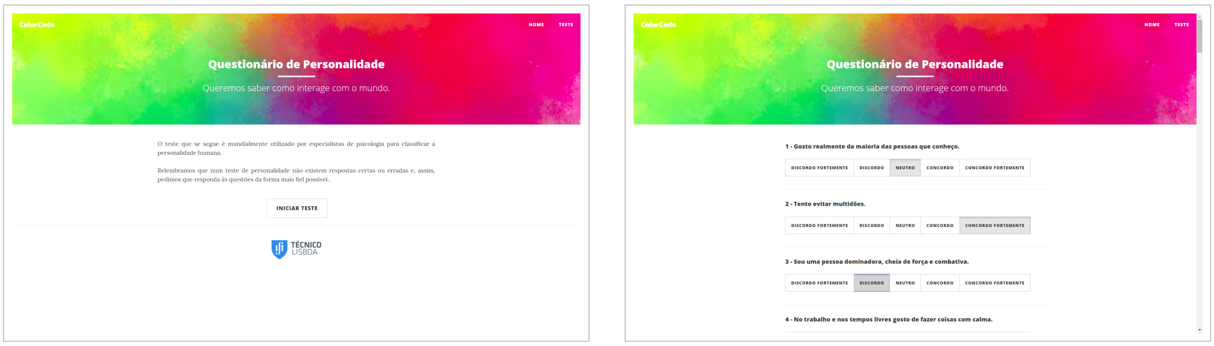

Figure 4.4: Personality introduction (left) and questionnaire (right).

4.2.4 BITalino Installation and Calibration

The BITalino installation and calibration phase is the only task which does not depend on the user. During this phase, the user is introduced to the neurophysiological signal recorder for the first time in the test.

The first part of this task consists in installing BITalino. When the user reaches the page shown in Figure 4.5, they are asked to do nothing until the installation is complete. This page is a simple introduction page, following the introduction layout already explained before.

This basic interface is ideal not to interfere with the test conductor's task, since it is not a new layout which could catch the user's attention and curiosity. When the installation is complete, the participant is asked to proceed to the calibration phase through the button "Iniciar Calibração" ("Start Calibration").

The calibration phase was of extreme importance in terms of data gathering, since it served as baseline for the extraction of arousal for this study, which is fundamental to calculate the HR change for each color, per user. Our main goal was to induce a relaxed emotion, having in mind not only their previous emotional state but also the study conditions. To this end, we chose to include in our calibration phase

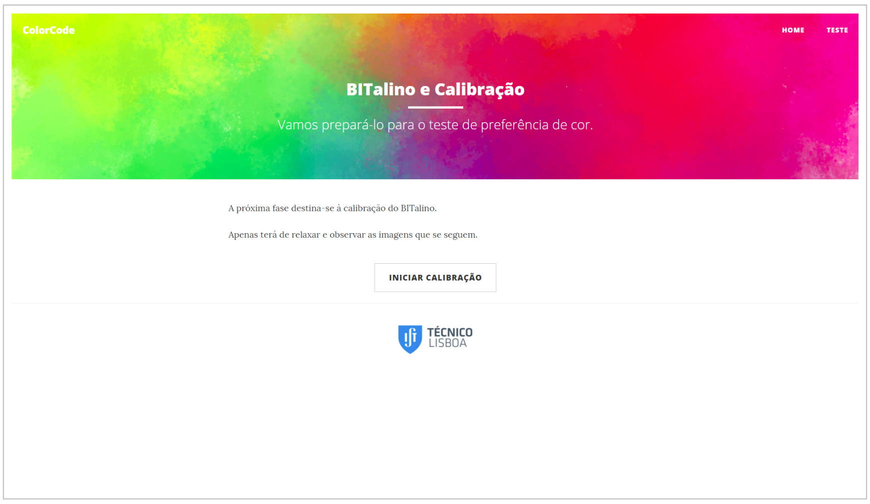

Figure 4.5: BITalino installation and calibration page.

images from the International Affective Picture System (IAPS), a research image dataset developed to provide a set of normative emotional stimuli for experimental investigations of emotion and attention[6]. This dataset includes a wide range of themes and each picture is categorized emotionally in terms of valence and arousal, in order to be used as a stimulus inductor.

The pictures used for this study (no. 5202, 5210, 5726, 5800, 5825, and 5829) are representative of flowers, landscapes and sunsets. Images containing nature and sunsets, among other contents, are considered pleasant and relaxing, in terms of valence and arousal respectively [40,41]. In Figure 4.6 we can see examples of the these types of pictures.

In light of this, during this calibration phase, each image of this set was shown for 20 seconds, and the set was displayed twice in loop, having this phase the total duration of 4 minutes. This way, each visual stimulus is long enough for the user to get used to the image while being induced to the relaxation state.

Figure 4.6: Examples of the types of images shown during the calibration phase.

4.2.5 Color Preference Test

Since our goal was to find the influence of personality (Extraversion) in color preferences, using not only self-reported preferences but also neurophysiological responses, our Color Preference Test (CPT) was thought and developed to support these two types of data gathering. Our goal was to design a test in which a set of colors were shown during a period, with the simplest interface, so the user would focus on the colors observed only, having also an easy way to rate the colors in terms of preference.

To this end, we decided to use a set of colors defined to be used in Information Visualization (InfoVis), which includes the colors: Red, orange, brown, yellow, green, cyan, blue, purple, pink, white, gray, and

[6] http://csea.phhp.ufl.edu/

black [65]. Although there are examples of the colors chosen for this purpose, several shades exist for each of these colors (except white and black), which led us to the next step – choosing the shade to use for each color. Since this was a test to be performed using a computer, we chose the colors based on the RGB color model.

In light of this, the shade for each color, shown on Figure 4.7 was chosen as following:

- **White** and **black** already defined
- For **red**, **green** and **blue**, we chose the pure primary colors used on RGB representation.
- For **orange** and **cyan** we used their pure code as secondary colors of this model.
- Since magenta is the third secondary color, being considered between two of the colors used on this test – **pink** and **purple** –, we chose not to use this secondary color, but a shade of each of these colors different enough not to confuse the user.
- The shade of **brown** was picked using the same line of thinking, by choosing a shade significantly different from the remaining colors.
- The **gray** shade we chose to use was the basic HTML gray.

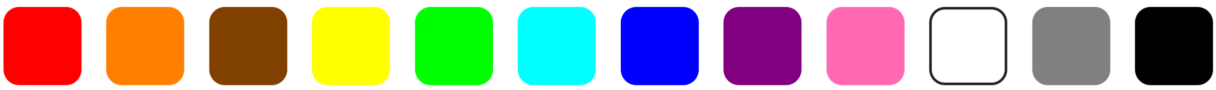

Figure 4.7: Colors used in the Color Preference Test.

The next step was to define how the test would be presented to the user. To this end, the first thing to take into account was the total duration of the test and, consequently, the time available to the CPT. To get a satisfactory amount of users, we limited the test's duration to the maximum of 30 minutes. To know the maximum length of the CPT, we had to estimate the duration of the user welcoming, the explanation scripts and the other tasks. After experiments, we defined the maximum duration of 5 minutes for this task.

Knowing that the visual reaction time is much shorter than 1 second [2], we decided to use the first second as tolerance period, also including this period on the full duration of each stimulus.

Although our goal was to use images as stimuli, nothing besides color is represented, implying that there are no people, objects or situations shown which would require more time to be noticed. A recent survey on using EEG signals for emotion recognition [1] reported that, among a considerable number of works using images as stimuli, the duration of each stimulus varied between 1.5 and 48 seconds. However, we had to take into consideration that, to compute HR, it is necessary to record at least two heart beats[7]. Knowing that an adult normal resting HR varies between 60 and 100 beats per minute[8], the minimum recording time needed to compute this feature is 2 seconds.

To give us room for maneuver for the data treatment, we considered the stimulus period of 5 seconds per color. In these 5 seconds, the color occupied the full screen, with nothing besides a white circle in the middle of the screen, as shown in Figure 4.8. This circle was added to the interface in order to avoid ocular movements and artifacts on the EEG recordings, which could affect the resulting data.

The circle color is gray and its fill is white. We chose both these colors for not having any hue. It was necessary to fill the circle because we wanted to show it in all the screens of the CPT, but both white and gray colors were used as color stimuli. Having a circle filled with other color allows the user to easily

[7] Information gathered from the BioSPPy code.
[8] https://www.health.harvard.edu/blog/resting-heart-rate-can-reflect-current-future-health-201606179806

Figure 4.8: Color in full-screen, when the stimulus is induced.

detect the circle in the screen. An important note is that, since it is presented in every single step of the CPT, it is expectable that the circle does not affect the stimuli.

After these 5 seconds of stimulus, a preference scale appears on the lower half of the screen. In this scale, the user may rate the presented color from “Strongly Dislike” to “Strongly Like”, similarly to the scale used for the personality questionnaire. Ten seconds is the duration that the user has to rate each stimulus, which is enough time to click in the preference rate, without thinking too much about the emotion felt. The selection is shown on Figure 4.9 and the recorded answer is the last selection during the 10-second period.

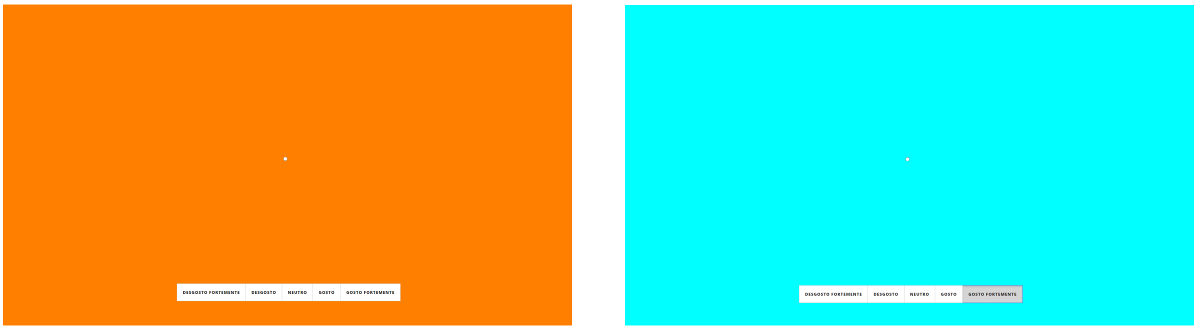

Figure 4.9: Self-report screen, before (left) and after (right) selecting the preference.

Since the duration per color is fixed, it was important to let the user know how much time was left to rate the observed color in terms of preference. To achieve this end, a countdown was added during the last 5 seconds of the self-reporting period. For this countdown, we chose to use white and gray font in darker and lighter colors, respectively, as we can observe in Figure 4.10. This way, we avoid the possibility of the user not seeing the countdown.

A black screen (see Figure 4.11) is presented between colors, to separate the stimuli. This screen is shown for a period of 5 seconds and, as well as each self-report screen, included a countdown of the remaining time until the next stimulus.

This black screen was added after performing the pilot tests (which will be described in detail in Chapter 5), since the afterimage effect was detected between some colors. Since we are using the RGB color model – an additive model –, black represents the absence of color, being the best color to minimize the visual artifacts between stimuli.

Having this said, the final version of this test on the interface consisted in 12 colors shown for 15 seconds each, separated by a black screen for 5 seconds, resulting on a fixed duration of 4 minutes.

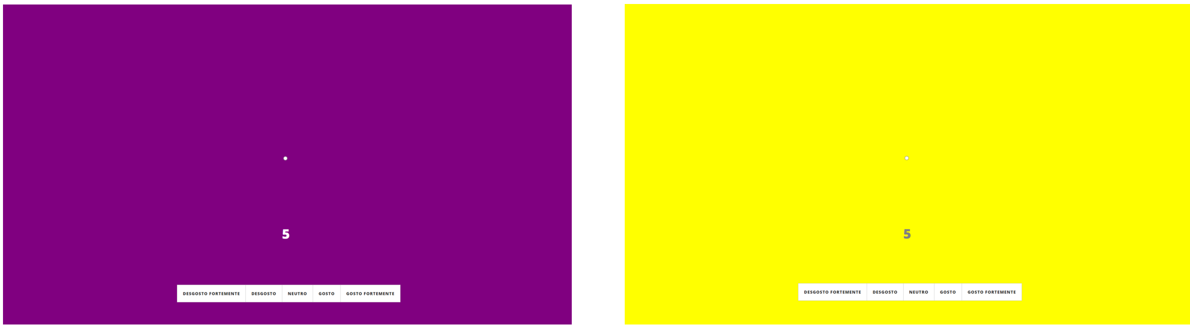

Figure 4.10: Countdown for the self-report screen, in dark (left) and light (right) colors.

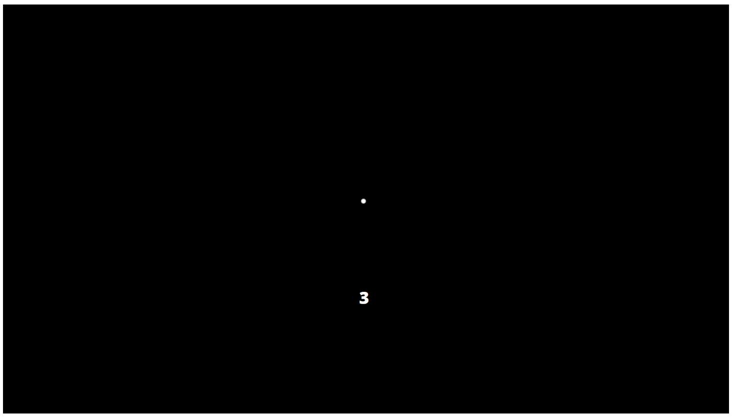

Figure 4.11: Black screen, used as separation between colors.

The order of the stimuli inducted is random, in order to avoid emotional responses influenced by the previous stimulus. This is true for all the colors except black, which is always shown as the first color. Making this decision, we were looking forward to avoid the user's misconceptions possibly caused by the separation black screen. On the other hand, a repetitive viewing of the same color may influence the resulting values from neurophysiological recordings.

4.2.6 Appreciation and Results

The last phase of the test was merely appreciative and informative. When the user reached the appreciation page, their test was considered as complete. This was when we thanked the subject for participating in the study, giving them the option of reviewing their personality test result. This step was not mandatory, being that the main reason why the personality test results are shown at the end of the test.

Also, the test was designed to be complete in 30 minutes maximum. Since the duration of this optional task could depend not only on the users' reading speed, but also their interest and curiosity, we chose to not include this last step in the maximum time limit.

The users who opted to see the results of their personality test could reach the last page by clicking on the "Ver Resultados" ("See Results") button, as shown in Figure 4.12, and proceed to the results page which can be seen in Figure 4.13.

The results page follows the main layout of the interface. In this page the user could review not only their Extraversion score, but also their score for each Extraversion facet. We chose to show this information by using horizontal percentage bar charts, in which the blue colored part represents the user's score for each facet. We also added to these bars the percentage value of each facet, so the user would know the exact value scored.

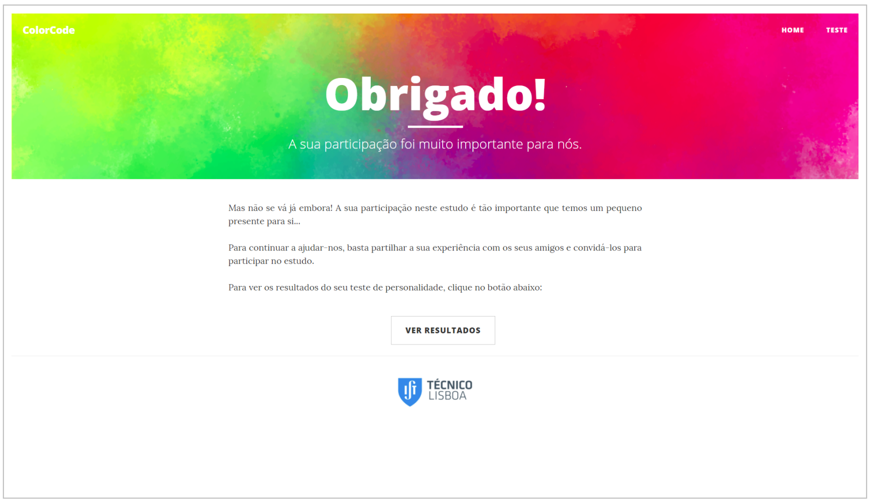

Figure 4.12: Appreciation page, shown when the user finished the test.

Complementing the bars, we included in this page the descriptions of the Extraversion values, based on the NEO PI-R Manual [14, 44]. In this text, not only the general definition of Extraversion is explained, but some examples are given on the meaning of higher and lower scores.

The same is applied to each facet of this dimension. As they are subscores of Extraversion, the facets are represented with the name on the left side of each bar. When clicking each facet name, a small description of the facet is shown, so the user can know what each score means in terms of personality traits.

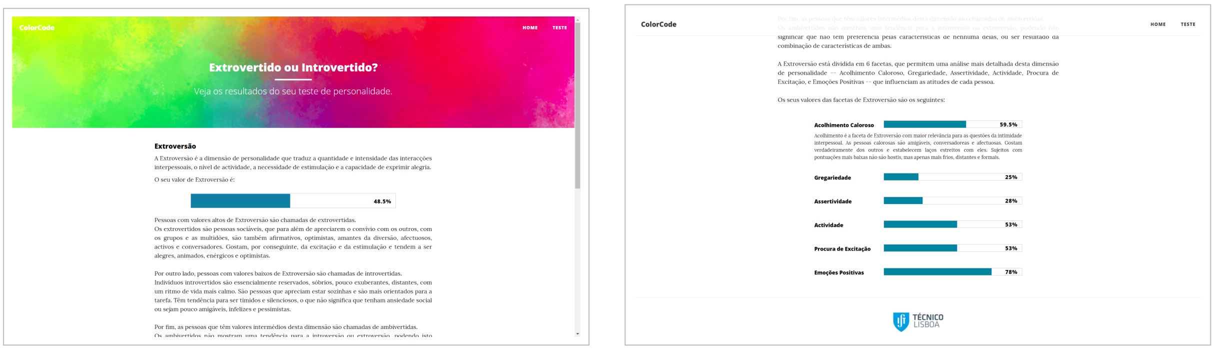

Figure 4.13: Personality test results: Extraversion score (left) and detailed facet score (right).

As this is the last screen, the test is given as fully completed. This interface was created in order to make the user follow the whole test in one way only, giving them no choice on the tasks order.

4.3 Apparatus

As this work consists on studying social and psychological dimensions of color to derive new guidelines for color uses, particularly in HCI, the development of the interface, the execution of the user tests and the data processing required specialized apparatus.

The first step of our approach was the implementation of the interface and the writing of the protocols. To this end, we used a personal laptop. This laptop model was ASUS K550JK-DM013, with a processor Intel Core i7-4710HQ 2.50 GHz, 8GB of RAM (2x4GB, DDR3L 1600 MHz SDRAM), and NVIDIA GeForce GTX850M 2GB DDR3 VRAM graphics card.

This computer was also used during the tests, for neurophysiological data gathering, while each participant performed the test using a PC containing a processor Intel Core i5-650 3.20 GHz, 4GB of RAM, and AMD Radeon HQ5800 graphics card. As this test was based on visual stimuli, we used a LG IPS Monitor 24MP55, calibrated in terms of color.

We used the toolkit BITalino[9], connected to the first mentioned computer, to collect the neurophysiological recordings, since it allows gathering physiological measures in a less intrusive way. Therefore, we used these devices as shown on the architecture scheme presented on Figure 4.14.

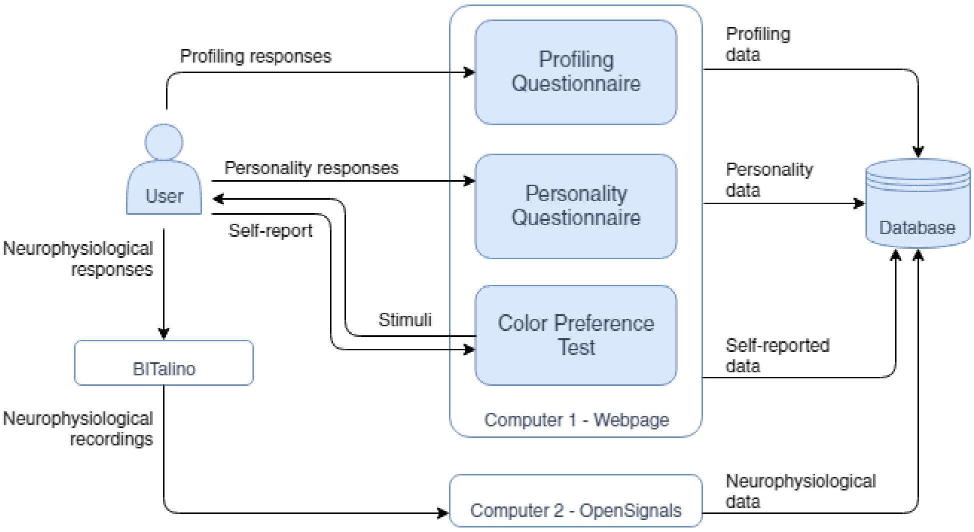

Figure 4.14: Architecture of the proposed solution.

The BITalino (r)evolution Plugged Kit BT, used in our work, includes Electromyography (EMG), Electrocardiography (ECG), Electrodermal Activity (EDA), Electroencephalography (EEG) and pulse sensors among other components (see Figure 4.15). In this research, we focused on collecting neurophysiological responses which could be used to compute emotional valence and arousal. To this end, we used the pulse sensor to collect Blood Volume Pulse (BVP) – later converted to HR, which is a well known indicator of arousal –, and EEG, from which emotional valence may be extracted.

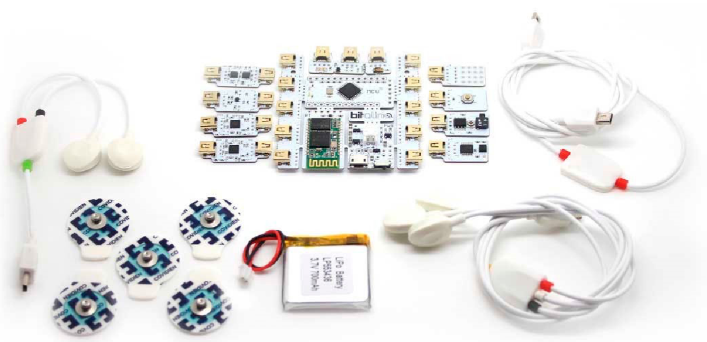

Figure 4.15: BITalino (r)evolution Plugged Kit BT [10].

This setup was used for all the participants to maintain the coherence of the study. For HR, we used the pulse sensor (composed by an ear clip and a sensor connection cable) clipped to the user's right earlobe. For EEG, we connected three gelled self-adhesive disposable Ag/AgCl electrodes to a 3-lead accessory, linked to an electroencephalography sensor, which is plugged to the BITalino board using a sensor cable.

[9]http://bitalino.com/en/

The 3-lead accessory identifies the EEG electrodes with the colors black (negative), red (positive), and white (reference). This allows us to work with a bipolar configuration, in which the red and black electrodes detect the electrical potentials in the specific scalp region with respect to the white electrode, which is placed in a region of low muscular activity. The resulting signal is the amplified difference between these two signals ($red - black$), eliminating the noise detected by the third electrode.

In light of this, the black and red electrodes were applied on the left and right sides of the user's forehead, respectively, and the reference one behind the ear, a zone of neutral brain activity. Since BITalino measures EEG using differential amplification, it is important that both black and red electrodes are placed in the same place of the user's head, but in opposite brain hemispheres, to measure the brain activity difference between hemispheres. In order to do this, we used the nose and the eyebrows as reference points, applying each electrode one finger above the side eyebrow and one finger from the nose line.

The BITalino board was connected via Bluetooth to the laptop to save the recordings. This computer was also used to process all the data collected, as well as to make the statistical treatment and complete the dissertation writing.

4.4 Software

In the course of our work, specialized software was required. In terms of the neurophysiological recordings, we used the OpenSignals (r)evolution[11], developed to visualize and record signals from BITalino. We used the Public Build 2017-08-28 version of this software, shown on Figure 4.16 .

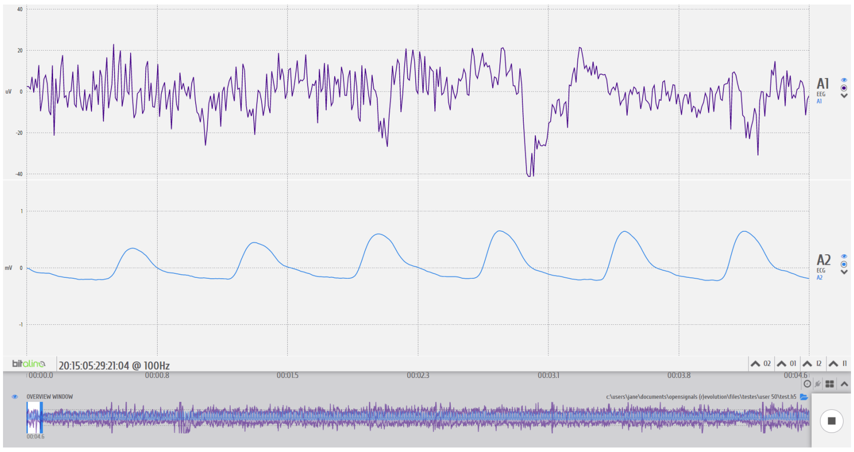

Figure 4.16: OpenSignals (r)evolution software.

With the OpenSignals software, we were able to record the EEG and BVP signals, saving each recording in H5 and plain text form. It also allows to visualize the signals while recording – which eased the detection of bad positioning of the electrodes – and after saving the files.

After gathering all the recordings, we used the BioSPPy toolbox, (version 0.4.0). This toolbox written in Python was developed to process biosignals such as BVP, ECG, EDA, EEG, EMG and Respiration, allowing feature extraction from these signals. For this work, we used it to extract the alpha waves from EEG and HR from BVP.

[11] http://bitalino.com/datasheets/OpenSignals_Datasheet.pdf

Lastly, for data organization and statistical analysis, we used Microsoft Excel 2013[12] and IBM SPSS Statistics 25[13]. Since our work leads to the derivation of new guidelines for HCI uses, a detailed statistical analysis was required. These software programs were used to normalize and organize the data, apply specialized statistical tests, and drawing charts representative of our results.

[12] https://products.office.com/en-us/microsoft-excel-2013
[13] http://www-01.ibm.com/support/docview.wss?uid=swg24043678

5

Data Collection

In the last chapter, we presented the data types we intended to collect for this work. In short, we decided the data we would need to study the influence of Extraversion in color preference included the users' profiling data, their Extraversion scores through a validated personality inventory, and their color preference through self-report and neurophysiological responses.

In this chapter we explain in detail all the data collection, including the ways used to gather, organize, process and interpret the data.

5.1 Data Gathering

Collecting a detailed database from a considerable amount of users, in order to make the statistical treatment as robust as possible, getting more reliable results for this study, is a crucial step during the development of our work. Designing and testing a solid study protocol is the first step of the data collection. This way, we can guarantee constant conditions between users. Our protocol was created while studying the theoretical background and the related work, defining the data types needed and developing the interface used for the user tests.

5.1.1 Pilot Tests

Before proceeding with the user tests, we decided to perform pilot tests. Since the study protocol was based on data gathering only, the pilot tests were designed as an evaluation through user participation [19], in order to detect flaws and improve the protocol and the developed interface. To this end, a Pilot Testing Protocol (Appendix A) was created and followed during this phase.

This new protocol included a full description of the room conditions, material, consent form, description of each phase of the test, including the explanation scripts to be used, and the usability metrics used to appraise the system. For this system appraisal, we divided the test in tasks and set individual effectiveness, efficiency and satisfaction metrics, as shown in Table 5.1.

The main goal of this pilot testing phase was to prevent future misunderstandings and delays during the user testing phase, as well as to estimate the duration of the user tests. In order to do that, we focused on the number of mistakes and time (by observation), respectively as effectiveness and efficiency metrics. Minimum, target and optimal values were defined as precondition for both these metrics, for all the tasks except for the ones performed by the researcher and for the ones designed to be taken within a fixed period. These values define, respectively, the worst, target and optimal cases expected for each task.

Two pilot tests were performed. Both testers were Computer Engineering master students specialized in Multimedia, meaning they were able to evaluate the system with a critical posture, taking basic usability rules into account while performing the test. They were asked to comment not only the tests included in the study, but the whole procedure, including the researcher's posture and speech, as well as the room conditions. Following the comments, the testers were encouraged to rate each task in terms of satisfaction. To follow the same design as the rating questions used in rest of the test, this consisted in a Likert scale from 0 (strongly dislike) to 4 (strongly like).

Task	Type	Metric	Worst	Target	Optimal	Result
Profiling	Effectiveness	Num. of mistakes	2	0	0	0
	Efficiency	Time	90 sec	60 sec	45 sec	65 sec
	Satisfaction	Rating	-	-	-	4
Personality	Effectiveness	Num. of mistakes	4	2	0	0
	Efficiency	Time	7 min	5 min	3 min	5 min
	Satisfaction	Rating	-	-	-	3
BITalino Installation	Efficiency	Time	10 min	8 min	6 min	11 min
	Satisfaction	Rating	-	-	-	4
Color Preference	Effectiveness	Num. of mistakes	4	2	0	0
	Satisfaction	Rating	-	-	-	3.5
Total	Effectiveness	Num. of mistakes	5	2	0	0
	Efficiency	Time	45 min	30 min	20 min	35 min
	Satisfaction	Rating	-	-	-	3.6

Table 5.1: Pilot testing metrics.

In terms of the procedure, both testers agreed with the choice of the present room conditions and scripts, making only minimal adjustments to some speech details about each task.

The users were able to finish all the tasks without any question, making no mistakes. No comments were made about the first task, having both testers given the highest rate of satisfaction when asked to evaluate the task. In terms of the personality questionnaire, one of the users found some grammatical errors. In this case, nothing was planned to be amended, since the questionnaire is validated and copyrighted, being applied according to its Portuguese version.

With respect to the Color Preference Test (CPT), both users noticed that, for some colors, the transition was unpleasant. This was due to the so-called afterimage effect. An afterimage is an image that continues appearing in one's vision after the exposure to the original stimulus has ended. In this case, the user's vision adapts to the first stimulus, which may affect the way the user perceives the following one. For some transitions, due to the continuous stimulus for each color (15 seconds each), the actual color was influenced by the previous one, taking more than the stimulus time (5 seconds) to recover and perceive the color correctly. This brought us a problem, since the first 5 seconds of each color were considered as the stimulus period and our goal was gathering the neurophysiological responses as well as the preference rating for each specific color stimulus.

´ Based on the results of the pilot tests, a few changes were made to the test and study protocol:

- The explanation scripts were simplified. This change would ease the interaction between the researcher and the user, while reducing the test duration.
- The data for EDA was excluded from the study, causing a reduction of the test length. This decision was made also based on the results of a previous study developed by Santos [60], in which the author found that the EDA values are not relevant for the classification of arousal when using BVP.
- To avoid the afterimage effect in the transition between colors, a black screen was added to be presented for five seconds between colors. This way, we are inducing a minimal visual stimulation, which also allows us to reduce the influence of the previous color on the actual neurophysiological response.

Beside the problems which led us to these changes, some BITalino installation issues were faced. However, even with a time period considered as the worst case in terms of the efficiency for this task, the full length of the test was near the defined target, including all the comments (planned for the pilot tests only).

5.1.2 User Tests

When developing a coherent study, it is essential to guarantee that the conditions are the same between users. To achieve this, a Study Protocol was created, describing all the steps each user test should follow. The first version of this protocol was created while defining the goal of this study. This allowed us to take into consideration the challenges in terms of data gathering, being also a great help in the construction of the interface and database.

The provision of the space for the tests, as well as the creation and improvement of the interface used, led to several changes on the study protocol before the pilot tests phase. After performing these tests, final adjustments were made on the interface and the whole user testing process, leading to the final version of the Study Protocol, presented on Appendix B.

Planning and Conditions

The user testing phase occurred between November 26th 2017 and January 26th 2018, in room 1-23 in IST, Taguspark campus.

During the first days of the user testing phase, around 2-3 tests were scheduled per day, to avoid delays and give the researcher time to get used to the study protocol and data organization. This has changed as the researcher felt more comfortable guiding the tests, which improved continuously the efficiency during the tests.

Belonging to the Visualization and Intelligent Multimodal Interfaces (VIMMI) research group, the room where the user tests took place was already prepared for a controlled laboratory environment, having controlled temperature and lighting conditions. These were crucial points of the protocol, since the test included recordings of reactions to visual stimuli and personal information sharing. As well as these aspects, it was also important to prevent the presence of others during the tests, which was also successfully achieved.

Users

The user testing was a very challenging phase in terms of users. Since this work is a study with results of extreme importance for HCI and interface design, it was crucial to collect a considerable amount of data to make the statistical treatment as robust as possible. Therefore, the testing phase counted with 50 users – equal amount of women and men.

To cover broad diversity of data was also a goal of ours, and since this work consisted in studying the influence of personality in color preference, excluding cultural and social factors, we wanted to focus on getting the same amount of female and male users. This was achieved by managing the user scheduling as follows: in the beginning of the study, we accepted every user who was interested in participate; when we achieved 25 male users we focused on accepting female users only, since the time we had to finish the data gathering was reduced.

During the tests, we noticed excitement and interest from the users regarding the whole test, specially the BITalino installation and the Extraversion results, shown at the end of the test. We thought it was also important to show our gratitude to each participant, being each user rewarded with a sweet. This, as well as the relevance and procedure of this study, led the users to share the word, which brought even more users to us. Besides this, no other method was used to gather users.

General Remarks

In general, all of the tests went as expected. Each test was completed within the estimated time – half of an hour –, and none of the tests were interrupted, neither any of the users quit during the test. For the tests, the final version of the interface, presented in Section 4.2, was used. We did not face any challenges in terms of interface, which was praised by many participants.

Besides liking the interface, most of the users understood the tasks correctly, completing all the questionnaires with no doubts and making no mistakes. This allowed us to gather the majority of the needed data successfully, aside from certain exceptions that were a slightly more challenging.

Challenges

As mentioned above, a few challenges were faced during the data gathering through user testing. Among these challenges are included some missing data values and mistakes that could be easily detected by the study researcher.

For instance, checking the profiling database data, the researcher identified a male user was who chose ”Female” as gender; or a few cases of researchers who chose ”Lower secondary education” – also known as the second cycle of basic education –, instead of ”Masters” – frequently called second cycle of superior education. In these cases, the mistake was immediately detected and, at the end of the test, the researcher questioned the users about the validity of the answers.

The most challenging issues faced occurred during the BITalino installation and calibration phase. Since the neurophysiological responses detection was unique for each user, there was no way to anticipate these issues. In cases where no BVP was detected, the researcher kept trying to adjust the electrode in question or the pulse sensor, in order to find the right position.

During the CPT, we detected an issue related to the implementation of the system: When there was a missing value of self-report preferences, none of the values was saved in the database. The researcher noticed this was a problem when a user failed rating the color black, mistaking it as a black division screen. Once this was noticed, and knowing how the whole system worked, the researcher proceeded with a manual registration of the self-reports for all the users. Except in those cases where the users misinterpreted the black color as a separation screen, all of the missing values happened because the users waited until the end of the countdown to rate the color. In these cases, the values were noted by observation, otherwise, the researcher waited for the end of the CPT to ask the user how they would rate the color black.

Lastly, one single user was affected by slow Internet connection speed, which caused an initial wait in the beginning of each task. This was only an issue in terms of the neurophysiological responses recordings, but once the researcher noticed that the previous pages took more than the expected time to load, the loading time of the neutral images and the CPT pages was clocked and then taken into account during the data synchronization, explained in Section 5.2.

5.1.3 Database

The majority of the data was directly collected through the interface, while the subjects answered the questionnaires during the user tests. To this end, and in order to create the best data organization, we created a MySQL database, dividing the data in several tables.

In light of this, we used the following organization for the MySQL tables:

Profiling The profiling table included the attributes: a unique **user identifier** (**user id**) used to anonymize the participants and connect all the information about each user between tables, **age**, **gender**, **nationality**, **education level**, the **Extraversion group** the user considered themselves to belong to, and if the user was subjected to **neurophysiological data** recordings before.

Ishihara This table is composed by the **user id**, and the **answers to the six plates** of the Ishihara's color blindness test [37].

Warmth, Gregariousness, Assertiveness, Activity, Excitement Seeking and Positive Emotion These six tables are similar and contain the **user id** and the **score of each of the eight questions** belonging to each Extraversion facet.

Colors This table has the **user id** and the **self-reported score** of each color shown in the CPT.

Results The results table includes the results from some calculations based on the information saved on the other tables. That being said, the attributes presented on this table are the **user id**, the **result of the color blindness test**, the **Extraversion score** computed from the scores of every question present on the personality inventory belonging to this dimension, and **each facet's individual score**.

Auxiliar table This last table was added to save some additional information about the test, of extreme relevance to process the neurophysiological data. The **user id**, the **starting time of the calibration phase**, the **starting time of the CPT**, and the **order the colors were shown to each user** are the attributes present on this table.

During the data processing phase, other tables were created in order to organize the neurophysiological data, for further use in statistical analysis. The tables were separated **by neurophysiological signal type**, including the **user id** and the **users' score for each color**. Lastly, new tables for each neurophysiological response type, containing the **normalized neurophysiological data** were created, which we explain in detail later in this document. For the statistical analysis, all the crucial variables were added to a single table, in order to study correlations and differences between groups of data.

5.2 Data Processing

After collecting the responses saved from the user tests through the interface and the subjects' neurophysiological responses, we proceeded to processing the latter in order to extract the data needed in their final form. To this end, we needed to understand the data types collected, design a way to select the data, extract the required features and organize the final data.

5.2.1 Data Types

When developing a study of this importance for interface design and HCI enhancement, it is crucial to understand the data, from its initial raw type until its intended form, including all the intermediate steps. During this work, we had to deal with data stemming from OpenSignals, which we never used before. The first step was interpreting and understanding the data originated from this software, and identifying the data we needed.

OpenSignals saves two types of raw files containing the information for each recording[1] – an H5 file (*.h5) and a plain text file (*.txt). The first one is saved in Hierarchical Data Format (HDF), frequently used in engineering and medical fields, and can be opened with OpenSignals, with which we may visualize the recorded data. For this work, we used the raw data in the plain text format, in which we can find the information we need for this study in text form. Among other information, we identified on the raw file the features necessary for the development this work: recording date, recording time, device, and the samples recorded from the selected channels. Being only informative, the first three mentioned features were represented in standard formats, without the need of further treatment. On the other hand, the channels' recordings required a detailed understanding.

In Figure 5.1, we can observe how the raw data is saved in the plain text file. Below the information header, we can see a total of seven columns. As described in the header, the first column represents the sample sequence number generated on the device, the second and third columns contain the state of the digital input, the following two columns represent the state of digital output and the last three the raw data sampled by the device through A1, A2 and A3 channels. In our case, we chose to connect the EEG, EDA and Pulse sensors to the channels A1, A2 and A3, which was then reduced to A1 and A2 when we removed the EDA responses from the study.

```
test.txt
# OpenSignals Text File Format
# {"20:15:05:29:21:04": {"sensor": ["EEGBITREV", "ECGBIT"], "device name": "20:15:05:29:21:04",
"column": ["nSeq", "I1", "I2", "O1", "O2", "A1", "A2"], "sync interval": 2, "time": "15:28:35.423",
"comments": "", "device connection": "BTH20:15:05:29:21:04", "channels": [1, 2], "keywords": "",
"mode": "bitalino_rev", "digital IO": [0, 0, 1, 1], "firmware version": 1281, "device":
"bitalino_rev", "position": 0, "sampling rate": "bitalino_rev", "label": ["A1", "A2"], "resolution":
[4, 1, 1, 1, 1, 10, 10], "date": "2017-11-26", "special": [{}, {}]}}
# EndOfHeader
0   0   0   0   0   537 533
1   0   0   0   0   583 525
2   0   0   0   0   557 519
3   0   0   0   0   523 512
4   0   0   0   0   465 509
5   0   0   0   0   526 507
6   0   0   0   0   545 505
```

Figure 5.1: Example of a raw file, output of neurophysiological recordings using OpenSignals.

The values sampled from each channel are the result of the conversion of the analog signal to the digital form. To process the data in its Analog-to-Digital Converter (ADC) form, the first step is to divide the raw data files into a file for each channel. In order to convert the ADC values to EEG values (in microvolt), the following transfer function[2] must be applied:

$$EEG(\mu V) = \frac{(\frac{ADC}{2^n} - \frac{1}{2}) \cdot VCC}{G_{EEG}} \times 10^6, \tag{5.1}$$

[1] http://bitalino.com/datasheets/OpenSignals_File_Formats.pdf

[2] The transfer function for each type of data can be found in each specific sensor data sheet, provided by the BITalino team.

Where:

$$\begin{aligned} ADC &\text{ - Value sampled from the channel,} \\ n &\text{ - Number of bits of the channel}^3, \\ VCC &= 3.3V \text{(operating voltage),} \\ G_{EEG} &= 40000 \text{(sensor gain)} \end{aligned}$$

On the other hand, the ADC values of BVP are already ready to be converted to HR. In order to get a baseline for arousal, these BVP values were also recorded while inducing a neutral emotional state in the user.

Therefore, the whole processing of the data gathered using OpenSignals was made by reading and writing on plain text files, requiring only a text editor access and modify them. All the data collected from the questionnaires are saved in a database using MySQL. After saving all the users' answers, these tables were downloaded as CSV files, which allowed us to easily import and organize the data using Microsoft Excel.

5.2.2 Neutral Data Selection

After dividing the raw files into a file per channel, we proceeded with the selection of neutral and color preference data. The first step is to synchronize the signal recorded using OpenSignals and the data gathered through the interface. This synchronization is crucial for the data treatment, since the data was gathered in two different computers – the questionnaires' data was saved in the database while each user completed their test, and the neurophysiological responses were recorded by the researcher using the OpenSignals software.

As explained in Subsection 5.1.3, in order to perform this synchronization, we registered the time of the calibration and color preference tests. This way, we would be able to synchronize the starting points, by subtracting the test starting time to the recording starting time. The result of this subtraction would be the amount of seconds to discard from the files corresponding to the recordings of each channel. Having the amount of time to delete in seconds and knowing the selected sampling rate was 100Hz (100 samples per second) for all of the tests, we know the amount of samples to be deleted from the files by calculating:

$$\text{Number of samples} = \text{Time in seconds} \times 100. \tag{5.2}$$

After performing this synchronization, we continued the data selection. It is important to take into account that, from the beginning of the BITalino calibration and the CPT, the gathering period was the same for all of the users. This is due to the fact that the users were exposed to the same neutral images and colors for the same amount of time, predefined when creating the interface (see Section 4.2).

For the neutral data selection, we followed the division explained on Figure 5.2. As shown, we decided to use a period of 60 seconds (6000 samples) per user. This length was set to compute the baseline value, being for that reason much bigger than the length used for each color stimulus, as explained in the next subsection.

Explaining this choice in a more detailed way, we needed the baseline to be as solid as possible for each user. Since the neurophysiological responses were recorded and sampled while the user was relaxing observing a set of neutral images, some emotions could be induced in some user by specific images. In order to avoid this, in that 60 second period, three to four different images were shown to the user, giving us a stronger baseline for HR.

After the synchronization was completed, and knowing that the BVP file's first 24000 samples (equivalent to 240 seconds of recordings) are representative of the calibration phase, we proceeded to the last step of the neutral data selection. The first 18000 samples were deleted from the file, as well as all of the samples following the 6000 samples (60 seconds) selected to be used.

The resulting files for all the users' neutral data have 6000 lines.

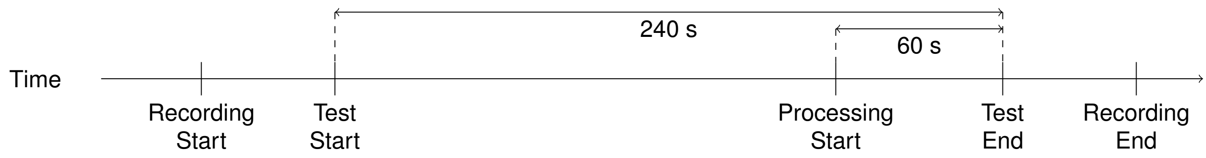

Figure 5.2: Time division for the neutral data.

5.2.3 Color Preference Data Selection

After selecting the neutral data for further interpretation, we proceeded with the color preference data selection, this time for EEG and BVP. As a first step, the same synchronization technique described before for the neutral data selection was planned to be used for this selection.

In contrast to the neutral data, we intend to select from the color preference file several sample intervals, corresponding to the colors shown to the users as stimuli. This being said, after performing the synchronization, we moved forward with the division by color. As planned since the interface design, this division was eased by the fact that all the colors were sequentially shown to all the users during the same period of time. As previously described in Section 4.2, the CPT has a total duration of 4 minutes. As shown in Figure 5.3, the 240 seconds length of neurophysiological recordings corresponding to the color preference test are equally divided into the 12 chosen colors, 20 seconds each.

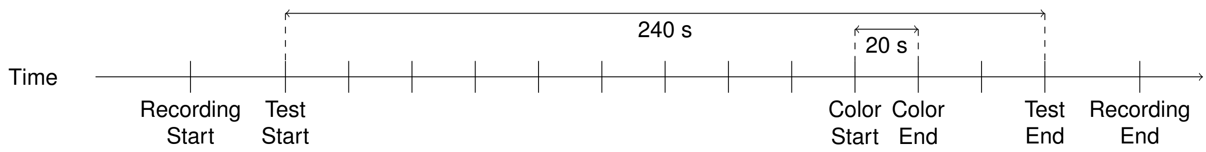

Figure 5.3: Time division for the color preference data.

After dividing the EEG and BVP files by colors, the resulting 2000 samples were treated following the division shown on Figure 5.4. As this figure shows, and as explained in Section 4.2, each color in the CPT is divided in three phases:

1. **5 seconds** in which the user was focused on a white dot in the center of the monitor, perceiving the stimulus;
2. **10 seconds** when the user classified the observed color having into account their preference;
3. **5 seconds** of black screen, time during which the user could neutralize their vision and prepare for the next stimulus.

Our idea for this division was to record the first 5 seconds of each visual stimulus with a minimal number of artifacts, to process and study only 2 seconds of that period. Although this was enough time to compute EEG, while testing this possibility, we found that 2 seconds were not sufficient to process BVP into HR. Due to this, and having the possibility to extend this time interval, we decided to process the last 4 seconds of this first phase.

This being said, we discarded the first 100 samples per color, saving the next 400 and deleting the rest.

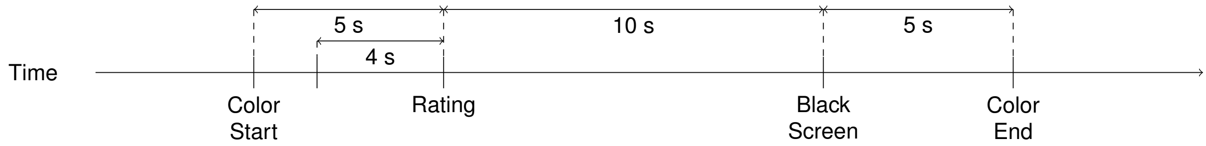

Figure 5.4: Time division for each color data.

5.2.4 Challenges

Although everything was carefully planned, we faced a main challenge during this phase that we were not able to detect with the pilot tests. During the data analysis, we realized that the times saved in the OpenSignals files and the ones recorded by the interface during the user tests were not coincident, neither the difference was constant. With this problem, we were not able to perform the synchronization as expected.

The issue was detected when we realized that, for some cases, the difference between starting times was longer than ten minutes, when the recording was shorter than five minutes. In some other cases, this difference was negative, pointing that the users started the test before the researcher started recording the neurophysiological data through OpenSignals. It is important to reinforce that the User Test Protocol (Appendix B) was followed equally for all the tests, meaning that this difference could never be negative.

The solution found for this challenge consisted in using OpenSignals to find the stimuli starting time manually. For neutral data, this task was simplified since we intended to get a relaxation state baseline of 1 minute from a 4 minute stimuli phase.

As explained in Subsection 5.2.2, we initially planned to use the last minute of the neutral stimuli phase to this end. Our solution for synchronizing the neutral files was not only based on the file per se, but also on the researcher's knowledge on the user tests conditions and performance.

During the calibration phase of the user tests, the researcher was responsible for starting and ending the recording, respectively before the user started observing the images and after the neutral stimulus ceased. Therefore, we know that each recording was stopped 1 to 10 seconds after the end of this phase. Having this in mind, we could roughly find the starting time, by subtracting 250 seconds to the total time – 10 seconds corresponding to the discarded time, plus 240 seconds, which was the period during which the user was observing the neutral images.

This solved the problem for the majority of the cases, leaving just a few exceptions to deal with. For some users, the recording's total duration was shorter than 250 seconds. For these users, the time exceeding from the 240 second stimulus period was roughly divided in half, being considered as the periods before and after the images.

Since this issue brought us the impossibility of knowing the exact moment when the user started observing the neutral images, but having the synchronization performed as explained, we then decided to also change the period selected for processing, choosing the period between 150 and 210 seconds. This way, we have a margin of error of 30 seconds, but we still guarantee that the user was exposed to the relaxation stimuli for a considerable period of time.

Apart from these cases, we witnessed a situation when the recording was suddenly stopped by OpenSignals, having caused a loss of the beginning of the calibration phase. In this case, we subtracted 70 seconds to the end of the file – 10 seconds of discarded time, plus 60 seconds to process –, and processed the 60 seconds starting from that moment.

However, these synchronization techniques were not applicable to the color preference data selection, since we have a much smaller margin of error per color. Again, we decided to perform the synchronization manually. To this end, we used OpenSignals to read the recordings, so we could visualize the signal.

During the CPT, each color was divided as previously described in Figure 5.4, including three steps: the stimulus period, followed by a color classification period, ending with a black screen to neutralize the user's vision before the next stimulus. Before the test, the researcher asked the participant to try not to blink and move their eyes during the stimulus period, in order to avoid artifacts in the EEG recordings caused by ocular movements.

Having this information, we planned our solution. Knowing that the ocular movements are easily detected using OpenSignals signal visualization, and that people tend to blink before and after the period during which they are asked not to blink, we focused on the 5-second stimulus period of each color. Finding the pattern composed by these 5 seconds, followed by 15 seconds of possible blinking, repeatedly for 12 times, we identified the moment when it happened for the first time, which led us to the starting time of the CPT. Having this information, we proceeded with the synchronization, performing the planned selection afterward.

However, there were some exceptional cases in which we could not find the 5-second stimulus period clearly. This could be taken as an issue, but in our work it was also an advantage. Not recognizing the ocular movements of a participant means that their movements did not generate major artifacts on the EEG signal. Having this said, in these cases we could consider the first 15 seconds of each color as color stimulus period, having more room for maneuver. In light of this, we found the patterns and, consequently, the starting moment of these files, proceeding with the synchronization and the planned selection.

Although this was a major challenge during our work, we successfully overcame it, guaranteeing the data selection, regarding both neutral and color preference, for all subjects.

5.2.5 Feature Extraction

As discussed in Chapter 4, we decided to use HR to obtain arousal, and the asymmetry of the frontal alpha brain waves to obtain valence.

To extract the intended features from the neurophysiological data collected until this point of our work, we used the BioSPPy toolbox[4]. Using this toolbox, we can obtain HR from BVP, and the theta, alpha high, alpha low, beta and gamma brain waves from EEG.

For arousal, we have chosen to compute the HR change from a baseline value. To this end, the first step was to extract this feature from each user's BVP files – resulting from the neutral and color selection –, using BioSPPy. Then, we calculated the median value from the values resulting from the output files. The last step was to find the difference between the baseline value and each color value, per user, using the expression:

$$differenceValue = colorValue - baselineValue \tag{5.3}$$

For valence, we performed slight changes to the BioSPPy script. Since our intention was to extract the alpha waves, we looked up on how the range of frequencies for each brain wave was specified and added a new range. Since the alpha low waves have average power in the 8 to 10Hz frequency band, and alpha high waves in the 10 to 13Hz frequency band, we added the extraction of alpha brain waves with average power in the 8 to 13Hz frequency band.

Using this modified version of BioSPPy, we were able to obtain an output corresponding to the whole range of frequencies of the alpha brain waves. Following what we previously did for the HR values, we calculated the median value for each color, considering it as the value of the alpha right-left asymmetry, representative of approach/withdrawal responses [10].

[4]https://github.com/PIA-Group/BioSPPy

The choice of using a median value for these final values was made having in mind that median is the measure of central location which is less affected by outliers [29]. Having these values of HR difference and alpha brain waves asymmetry, per color and per user, we could proceed with the data interpretation.

5.3 Data Interpretation

When creating a data collection like the one used in this work, we need to know how to interpret the data after gathering and processing it. In this subsection we explain how to understand and apply the final changes needed in data, in order to prepare it for further statistical analysis.

5.3.1 Color blindness

As previously mentioned in Section 4.2, the profiling questionnaire included a six-plate version of Ishihara's color blindness test [37] to find any color blind subjects.

For the chosen plates, the answers given by a person with no color vision deficiencies are: 12, 8, 74, 6, 5, and "It is not a number". In case the user fails any of these plates, writing a different number as answer, they are registered in the database as being color blind. This process happens during the filling of the questionnaires, through the interface.

In case of color blindness, the participant's data would be discarded or used as a study case.

5.3.2 Personality Score

During the gathering of the scores of the personality questionnaire through the interface used for our user tests, we needed to assign a value to each answer given. To this end, we followed the NEO PI-R manual [14], in which the authors explain how to compute the value for each question.

This test assigns, to each personality dimension, a score between 0 and 5. In order to reach this value, the questions are divided by facet of each dimension. To each question, the subject gives an answer between "Strongly Disagree" and "Strongly Agree". For some questions, the score is given by mapping the answers directly from 0 to 4 (0 - "Strongly Disagree", 1 - "Disagree", 2 - "Neutral", 3 - "Agree", and 4 - "Strongly Agree"), while for the remaining ones the score is reversed, according to the following equation:

$$reversedScore = 4 - score$$

After collecting the score for all the questions, reversing the ones referred in the manual, the score assigned to each facet is given by the average of scores of the questions belonging to that facet. Similarly, the final score of the dimension is given by the average of the scores of the dimension's questions.

In order to ease these calculations, we saved the scores to the questions by table of facet, as explained in Subsection 5.1.3, meaning that after collecting this information, the final scores of Extraversion and its facets – Warmth, Gregariousness, Assertiveness, Activity, Excitement Seeking, and Positive Emotion – were easily calculated in average.

We then divided the scores into three main groups – Introverts, Extraverts and Ambiverts –, in which Introverts have low Extraversion scores, Extraverts have high Extraversion scores and Ambiverts have their scores between Introverts and Extraverts. For this division, we followed a technique used by several authors [12, 31, 32]. Having the mean value of Extraversion between all users, the Ambivert group includes the participants with scores between mean value ±standard deviation. Below the lower value, the subjects are considered as Introverts, and above the higher value as Extraverts.

5.3.3 Valence and Arousal

With the neurophysiological data processed and organized, we proceeded to interpret it, mapping it into valence and arousal scales. Since we are using a standard scale along the development of this work, in which the Extraversion score and the self-reported color preference are already mapped, we chose to use this scale for the data normalization – from 0 (minimum value) to 4 (maximum value).

Since we are dealing with neurophysiological responses, which may vary from subject to subject, we chose to normalize the data per user, having into account their individual responses. To this end, and having the final values of the Alpha waves values extracted from EEG and the HR difference for each color, we started by identifying each user's minimum and maximum values for both these features. We then proceeded to normalize the values applying the following function:

$$normalizedValue = \frac{(value - userMinimum) \times 4}{userMaximum - userMinimum}. \tag{5.4}$$

Although we are aware that some data may be affected by the use of this normalization, since the user may have no color which they strongly like or strongly dislike, we believe this is the best approach because it depends on each user values only, preventing the data impairment caused by other users' neurophysiological responses. Also, normalizing the values on this stage of the data interpretation, before the statistical analysis, allows us to have more reliable values to compute valence and arousal median values per color.

6

Results

With the data gathered, processed and interpreted, we proceeded to the statistical analysis. In this step of our work, we study the data by observing trends and applying widely used statistical tests, to explore the influence of personality on color preferences.

In this chapter, we present the results of the statistical analysis performed on the user profiling, Extraversion, and color data, as well as the study made on the relation between Extraversion and color (preference and arousal).

6.1 User Profiling

We had 50 participants in the user testing phase of this work, with only one of them presenting some issued with the neurophysiological data collected. When processing the mentioned participant BVP recordings, using BioSPPy, it was not possible to compute HR due to the lack of heart beats detected in the signal. Since this feature was not possible to extract, we did not use this user's data for our study, making the decision of slightly compromising the even balance we had when gathering the data to get more reliable conclusions from our study. Nevertheless, we succeeded in guaranteeing a significant number of subjects participating in our user tests.

In light of this, the study counted with the participation of **49 subjects**, **49% females and 51% males**, with **ages between 18 and 45 years old** (only two participants were older than 35 years old). All the participants had Portuguese as their mother language, **having the great majority of the users Portuguese nationality** (around 90%), while the remaining subjects were Brazilian. In terms of education, around 78% of the subjects had an academic degree (Licenciate or Master), while 20% had a high-school degree, and only one participant had upper secondary education level. For the majority of the users (80%), this was the first participation in tests with neurophysiological recordings.

Regarding Ishihara's color blindness test [37], five subjects failed to answer the six plates correctly. However, a pattern was found among these subjects – all of them failed only one of the plates, considering the forth plate as not being a number instead of being the number 6, or finding the number 4 in the last plate, where no number should be found.

Following the guide instructions to detect color blindness, these mistakes could suggest these subjects had red-green deficiencies. The author states that, for the forth plate, the subjects with these deficiencies find no number, while for the last plate, they find the number 45. However, these types of deficiency would affect the users' capability of recognizing the numbers on the remaining plates correctly.

Since this was not true for the participants who failed the six-plate color blindness test, we did not consider these subjects as color blind, thus including their results in this study.

6.2 Extraversion

Extraversion, the dimension of personality proposed to study in this work, was collected in two forms: self-reported data, through the profiling questionnaire, and personality data, through answering the NEO PI-R reduced to Extraversion.

When asked to which Extraversion group they belonged, **29% of the subjects considered themselves as Introverts**, **37% saw themselves as Ambiverts**, **33% said they belonged to the Extraverts' group**, and **one participant did not know which group they belonged to**. However, these results proved to be different from the personality questionnaire results. After performing the division of the users by Extraversion groups, as explained on Section 5.3.2, the personality questionnaire shown that: **18% of the users were Introverts**, **65% were Ambiverts**, and **16% Extraverts**.

Having this said, comparing the self-reported and the personality data, we observed that **63% of the users got their Extraversion group right**, while **14% reported themselves to belong to a lower Extraversion group**, and **20% thought they belonged to a higher Extraversion group** than the actual one. **Among the Ambiverts, around 19% wrongly classified themselves as Introverts and 28% as Extraverts**.

The descriptive statistics on the scores resulting from the personality questionnaire, for Extraversion and the facets of this dimension, are shown on Table 6.1.

Score	Mean	Median	Std. Deviation	Minimum	Maximum
Extraversion	2.31	2.33	0.47	1.29	3.23
Warmth	2.69	2.75	0.50	1.50	3.32
Gregariousness	1.89	1.88	0.79	0.25	3.38
Assertiveness	1.91	1.75	0.66	0.62	3.12
Activity	2.01	2.00	0.49	1.00	3.00
Excitement Seeking	2.51	2.62	0.63	1.00	3.75
Positive Emotion	2.82	3.00	0.68	1.25	3.75

Table 6.1: Descriptive statistics for Extraversion and its facets.

These results suggest that:

- The facets of Extraversion with **higher scores are Warmth** (*Mdn* = 2.75, *M* = 2.69, *SD* = .50) and **Positive Emotion** (*Mdn* = 3.00, *M* = 2.82, *SD* = .68).
- The facets of Extraversion with **lower scores are Gregariousness** (*Mdn* = 1.88, *M* = 1.89, *SD* = .79) and **Assertiveness**(*Mdn* = 1.75, *M* = 1.91, *SD* = .66).

To confirm the validity of these results, further statistical analysis is required. Since studying Extraversion by itself was not our focus, we did not perform statistical tests on these results. However, we drew the boxplot presented in Figure 6.1 in order to analyze the tendencies on the collected data.

Observing the boxplot, we see that:

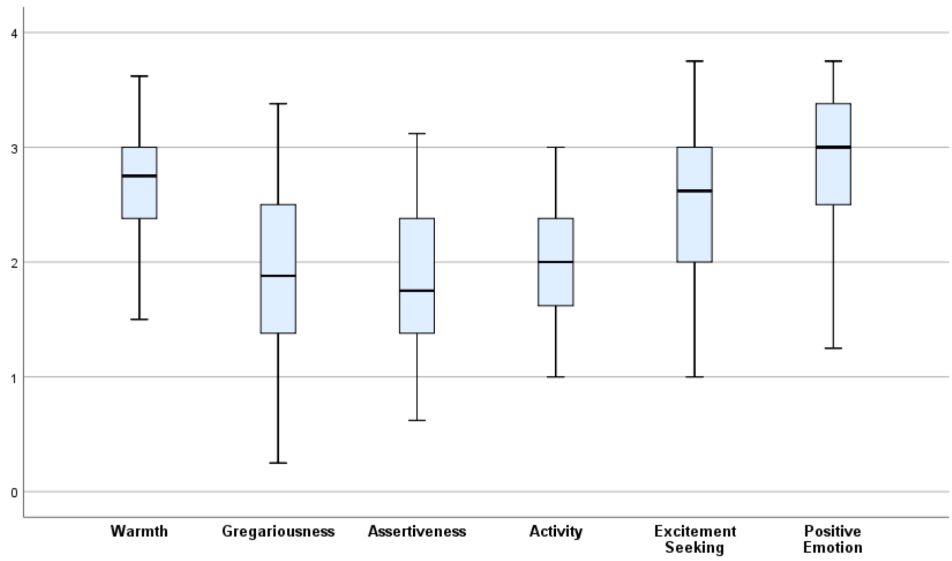

Figure 6.1: Comparison between the Extraversion facets' scores.

- **Warmth**, **Excitement Seeking** and **Positive Emotion are likely to have greater scores than Gregariousness**, **Assertiveness** and **Activity**.
- **Warmth** and **Activity** had the shortest ranges, suggesting that these **are the facets of Extraversion in which the subjects differ the least**.
- **The subjects are likely to differ the most in** the Extraversion facet **Gregariousness**, since it shows the widest score distribution. Hence, it is more difficult to draw conclusions on this facet.

Since our intention was to perform statistical tests to relate Extraversion (dimension and its facets) and colors, we performed the Shapiro-Wilk Test to verify if the data follows a normal distribution.

The results of the normality test, presented on Table 6.2, show that the Positive Emotion score is the only data which does not follow a normal distribution (p = .001). This means that, for further statistical analysis, nonparametric tests were used and their results will be presented ahead.

Score	Significance
Extraversion	.801
Warmth	.245
Gregariousness	.475
Assertiveness	.236
Activity	.440
Excitement Seeking	.194
Positive Emotion	.001

Table 6.2: Shapiro-Wilk normality test for Extraversion and its facets.

6.3 Color

With respect to color, we collected data referring to self-reported score, valence score through alpha wave asymmetry, and arousal score through HR. The descriptive statistics for each of these scores are shown on Tables 6.3, 6.6 and 6.8.

Color	Color Name	Mean	Median	Std. Deviation	Minimum	Maximum
	Red	2.22	2.00	1.09	0.00	4.00
	Orange	2.59	3.00	1.04	0.00	4.00
	Brown	2.04	2.00	1.02	0.00	4.00
	Yellow	1.53	1.00	1.21	0.00	4.00
	Green	1.55	1.00	1.21	0.00	4.00
	Cyan	2.61	3.00	1.17	0.00	4.00
	Blue	2.29	2.00	1.10	0.00	4.00
	Purple	2.47	3.00	1.08	0.00	4.00
	Pink	2.10	2.00	1.09	0.00	4.00
	White	2.63	3.00	1.06	0.00	4.00
	Gray	2.65	3.00	0.88	1.00	4.00
	Black	2.55	3.00	1.02	0.00	4.00

Table 6.3: Descriptive statistics for self-reported color preference scores.

The descriptive statistics for the color preference self-reported scores suggest that:

- **White** (*Mdn* = 3.00, *M* = 2.63, *SD* = 1.06), **gray** (*Mdn* = 3.00, *M* = 2.65, *SD* = .88) and **black** (*Mdn* = 3.00, *M* = 2.55, *SD* = 1.02), colors with no saturation, **are three of the most preferred colors**.
- Among the remaining colors, **orange** (*Mdn* = 3.00, *M* = 2.59, *SD* = 1.04), **cyan** (*Mdn* = 3.00, *M* = 2.61, *SD* = 1.17), and **purple** (*Mdn* = 3.00, *M* = 2.47, *SD* = 1.08) have the highest scores.
- The colors with **lower scores are yellow** (*Mdn* = 1.00, *M* = 1.53, *SD* = 1.21) and **green** (*Mdn* = 1.00, *M* = 1.55, *SD* = 1.21).

Further statistical analysis is required to confirm whether there is a significant difference among the scores, and which are the most liked and the least liked colors. In order to proceed with further analysis, a statistical test was performed in order to verify normality. The result of the Shapiro-Wilk test for self-reported data, valence (original and rounded scores), and arousal is shown in Table 6.4.

Since there was evidence against normality for all the colors in all samples – except for arousal scores of orange (*p* = .067) and brown (*p* = .071) –, a Friedman test was used, being the results presented on Table 6.5. We chose this test since it was developed to detect differences between two or more related samples which do not follow a normal distribution.

As shown, the Friedman test for self-reported color preference shows that there is a significant difference between the color scores ($\chi^2(11) = 69.292$, *p* = .000).

In terms of the valence scores, resulting from the alpha wave asymmetry, the descriptive statistics suggest that:

- **Black** (*Mdn* = 2.14, *M* = 2.16, *SD* = 1.47) is **the most liked color**.
- **Blue** (*Mdn* = 1.57, *M* = 1.82, *SD* = 1.24) and **white** (*Mdn* = 1.48, *M* = 1.60, *SD* = 1.29), are **the most preferred colors** after black.
- The least liked colors are **orange** (*Mdn* = 1.15, *M* = 1.74, *SD* = 1.39), and **cyan** (*Mdn* = 1.17, *M* = 1.40, *SD* = 1.03).
- **Cyan** is the only color which was not considered a favorite color.

Color	Color Name	Significance			
		Self-Reported	Valence	Valence (R)	Arousal
	Red	.000	.003	.000	.033
	Orange	.000	.000	.000	.067
	Brown	.001	.000	.000	.071
	Yellow	.000	.010	.000	.026
	Green	.000	.010	.001	.006
	Cyan	.000	.013	.000	.002
	Blue	.000	.005	.001	.023
	Purple	.000	.002	.000	.007
	Pink	.002	.002	.000	.014
	White	.000	.001	.000	.002
	Gray	.000	.001	.000	.013
	Black	.000	.000	.000	.000

Table 6.4: Shapiro-Wilk normality test by color, for self-reported, valence, valence rounded, and arousal scores.

	Self-reported	Valence	Arousal
Chi-Square	69.292	14.262	10.6000
Degrees of freedom	11	11	11
Significance	.000	.219	.477

Table 6.5: Friedman test for self-reported color preference, valence, and arousal scores.

Since we wanted to confirm the difference of scores between colors, we proceeded with further statistical analysis. Performing the Friedman test we found that **there is no statistically significant difference between the scores for valence, gathered through alpha wave asymmetry** ($p >$.05).

In order to find trends in our data, we built the boxplot shown on Figure 6.2.

Observing the boxplot, we can find that:

- Around 75% of the subjects scored the colors **red, brown, yellow, green, cyan, purple, pink and white as Neutral or below**.
- **More than 25% Strongly Like black**.
- There is a **slightly higher agreement for cyan and brown**.
- For the color **brown**, the central half of the distribution is **condensed between Dislike and near the lower value of Neutral** (1.5).

Both self-reported data and neurophysiological data were collected with the intention of comparing the color preference stated by the users and their neurophysilogical responses. Since the self-reported preference was ranked in a Likert scale, meaning the responses were discrete, and valence was measured in a continuous scale, we rounded the valence scores. This results of descriptive statistics are shown in Table 6.7.

The comparison of these results with the ones presented on Table 6.3 suggests that:

Color	Color Name	Mean	Median	Std. Deviation	Minimum	Maximum
	Red	1.66	1.40	1.20	0.00	4.00
	Orange	1.74	1.15	1.39	0.00	4.00
	Brown	1.50	1.27	1.24	0.00	4.00
	Yellow	1.48	1.33	1.14	0.00	4.00
	Green	1.59	1.44	1.14	0.00	4.00
	Cyan	1.40	1.17	1.03	0.00	3.64
	Blue	1.82	1.57	1.24	0.00	4.00
	Purple	1.46	1.33	1.18	0.00	4.00
	Pink	1.68	1.43	1.29	0.00	4.00
	White	1.60	1.48	1.29	0.00	4.00
	Gray	1.65	1.38	1.32	0.00	4.00
	Black	2.16	2.14	1.47	0.00	4.00

Table 6.6: Descriptive statistics for valence scores.

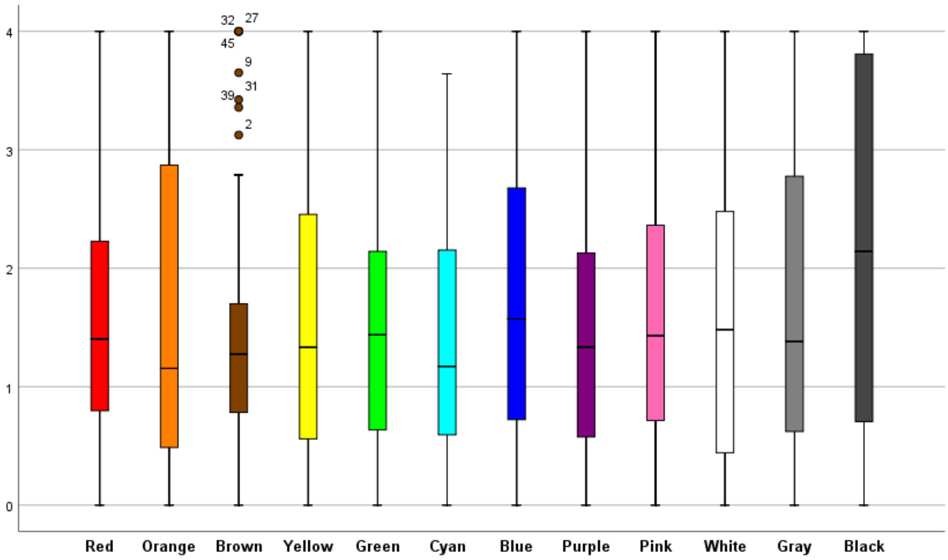

Figure 6.2: Comparison of the valence scores between colors.

- Yellow, green and blue have equal median scores of self-report and valence, suggesting that **the self-reported preference for yellow, green and blue is the most legitimate**.
- **Orange, cyan, purple, white and gray are the colors for which the self-reported preference is further from the neurophysiological preference**, since they show the biggest difference between median scores.

In order to confirm these results, and since the distributions present evidence against normality, a Wilcoxon signed ranks test was applied between the self-reported score and the valence score, per color. This is a nonparametric test designed to find differences between two sets of scores that come from the same subjects (dependent variables).

This test showed that **there is not a statistically significant difference** between the self-reported and valence scores for the colors **yellow** (Z = -.248, p = .804), **green** (Z = -.115, p = .909), **pink** (Z =

Color	Color Name	Mean	Median	Std. Deviation	Minimum	Maximum
	Red	1.65	1.00	1.234	0.00	4.00
	Orange	1.65	1.00	1.451	0.00	4.00
	Brown	1.49	1.00	1.277	0.00	4.00
	Yellow	1.49	1.00	1.175	0.00	4.00
	Green	1.59	1.00	1.171	0.00	4.00
	Cyan	1.43	1.00	1.155	0.00	4.00
	Blue	1.86	2.00	1.258	0.00	4.00
	Purple	1.39	1.00	1.204	0.00	4.00
	Pink	1.67	1.00	1.360	0.00	4.00
	White	1.57	1.00	1.384	0.00	4.00
	Gray	1.65	1.00	1.362	0.00	4.00
	Black	2.14	2.00	1.514	0.00	4.00

Table 6.7: Descriptive statistics valence round scores.

-1.758, p = .079), and **black** (Z = -1.605, p = .108).

From the results with statistically significant difference, **blue is the color with closest scores of self-reported preference and valence** (Z = -2.010, p = .044, r = -.287). On the other hand, **cyan is the color with biggest difference between scores** (Z = -4.246, p = .000, r = -.607). The test shows that, for all the colors proved to have a statistically significant difference, the users tended to **self-report their preference higher comparing to their feelings**.

Color	Color Name	Mean	Median	Std. Deviation	Minimum	Maximum
	Red	2.20	2.26	1.23	0.00	4.00
	Orange	2.04	2.00	1.22	0.00	4.00
	Brown	1.85	1.69	1.12	0.00	4.00
	Yellow	1.90	1.85	1.19	0.00	4.00
	Green	1.76	1.72	1.28	0.00	4.00
	Cyan	1.81	1.74	1.32	0.00	4.00
	Blue	1.68	1.59	1.23	0.00	4.00
	Purple	1.79	1.59	1.29	0.00	4.00
	Pink	1.91	1.84	1.30	0.00	4.00
	White	1.54	1.14	1.17	0.00	4.00
	Gray	1.90	1.98	1.26	0.00	4.00
	Black	2.19	2.07	1.49	0.00	4.00

Table 6.8: Descriptive statistics for arousal scores.

Lastly, the descriptive statistics resulting from the color arousal scores, through HR, suggest that:

- **Red** (Mdn = 2.26, M = 2.20, SD = 1.23) is **the color which induces most arousal**.

- After red, **orange** (*Mdn* = 2.00, *M* = 2.04, *SD* = 1.22) and **black** (*Mdn* = 2.07, *M* = 2.19, *SD* = 1.49), are **the most exciting colors**.
- The color with lowest arousing properties is **white** (*Mdn* = 1.59, *M* = 1.68, *SD* = 1.23).
- **Blue** (*Mdn* = 1.59, *M* = 1.68, *SD* = 1.23), and **purple** (*Mdn* = 1.14, *M* = 1.40, *SD* = 1.03) are the **most calming colors** after white.

To confirm these results, we went further on statistical analysis, by performing a Friedman test. Since the results showed that the difference between arousal scores is not statistically significant ($p > .05$), we tried to find trends on the data visualizing the boxplot presented in Figure 6.3. The graph shows us that:

- **All the colors are someone's more or less arousing color**.
- Around **half of the subjects scored white between 0 and 1**, suggesting this is considered the least arousing color.

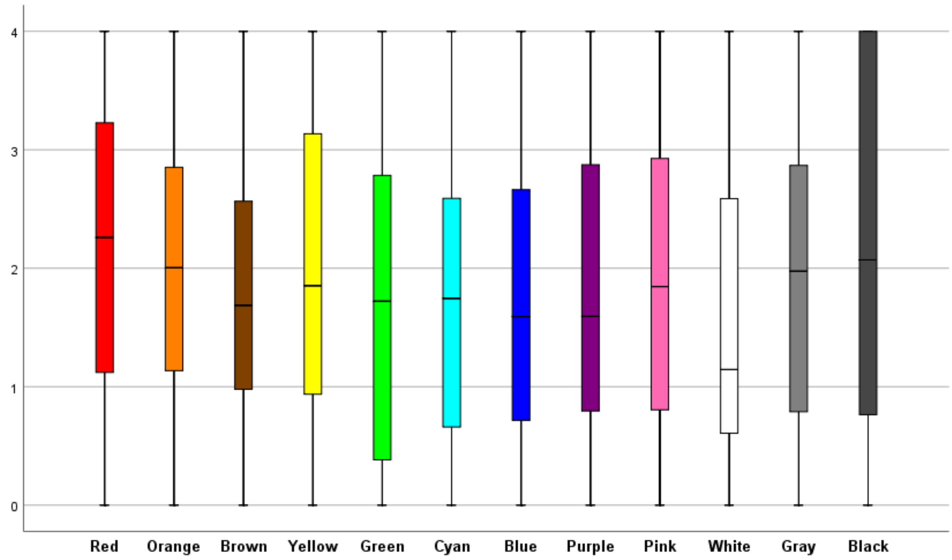

Figure 6.3: Comparison of the arousing properties of the colors.

6.4 Extraversion and Color

Our main goal with this work was to study the influence of personality on color preference. The first step towards this goal was trying to find a correlation between Extraversion and the color scores. Recalling Tables 6.2 and 6.4, the Shapiro-Wilk tests performed showed that there was evidence against normality in the majority of our data. For that reason, we chose to apply Spearman's correlation test, since it does not require the normality of the variables' distribution and it is not very sensitive to outliers.

6.4.1 Self-reported Color Preference

Applying the test to find correlation between the Extraversion score and the self-reported score for the colors used in this study, the results we obtained (presented in Table 6.9) revealed no statistical significance on the correlations studied. However, since the personality questionnaire allowed us to calculate the score for each facet, we decided to go further with the statistical analysis by trying to find correlation between the self-reported score and the scores obtained for each facet.

When running a Spearman's correlation test for **Warmth and self-reported scores**, the test showed that:

Color	Color Name	Correlation Coefficient	Significance
	Red	.067	.646
	Orange	.196	.178
	Brown	.010	.948
	Yellow	.202	.163
	Green	-.055	.707
	Cyan	.058	.694
	Blue	-.058	.690
	Purple	.019	.899
	Pink	.164	.261
	White	.267	.063
	Gray	-.150	.304
	Black	.016	.916

Table 6.9: Spearman's correlation between Extraversion score and self-reported score for each color used in the study.

- There was a **medium, positive correlation between the Warmth and** the self-reported preference for the color **white** (r = .396, p = .005).
- A **medium, positive correlation between the Warmth score and the self-reported score for yellow** was found (r = .292, p = .042).

The remaining results of the Spearman's correlation test on Warmth and self-reported color preference showed no statistical significance ($p >$.05), meaning that it is not possible to take conclusions from the collected data.

Applying the same test to find correlation between the facets Gregariousness, Assertiveness and Activity, and the self-reported preference, the results showed not to be statistically significant for any color ($p >$.05).

The same resulted from the correlation test on Excitement Seeking and self-reported preference for all the colors except **gray** (r = -.290, p = .043), for which a **medium, negative correlation was found between the Excitement Seeking facet and** the self-reported preference.

The last facet of Extraversion we tested to be correlated with the self-reported preference scores was **Positive Emotion**. Except for all **the remaining colors, for which the results were not statistically significant** ($p >$.05), we found that:

- There is a **strong, positive correlation between Positive Emotion and** the self-reported preference for **white** (r = .487, p = .000).
- The score of this facet showed a **medium, positive correlation with** the self-reported **orange** (r = .365, p = .010).

In addition to the correlation tests, we decided to analyze the difference of the self-reported preference, per color, between Extraversion groups (Introvert, Ambivert and Extravert). Since the Shapiro-Wilk normality test showed that not all the samples followed a normal distribution ($p <$.05), we used the non-parametric Kruskal-Wallis H test, obtaining the results shown on Table 6.10 . This test is a rank-based test which can be used to determine if there are statistically significant differences between two or more groups of an independent variable on a dependent variable.

Color	Color Name	Kruskal-Wallis H	Significance
	Red	1.529	.466
	Orange	1.394	.498
	Brown	0.316	.854
	Yellow	3.057	.217
	Green	1.910	.385
	Cyan	0.868	.648
	Blue	2.233	.327
	Purple	0.745	.689
	Pink	1.567	.457
	White	0.544	.762
	Gray	0.749	.688
	Black	2.158	.340

Table 6.10: Kruskal-Wallis test for the self-reported scores between Extraversion groups, by color.

As presented on the table, the results show that **the differences are not statistically significant between Extraversion groups for any of the colors**.

6.4.2 Valence through Alpha Wave Asymmetry

We then ran Spearman's tests to find possible correlations between the scores of Extraversion and its facets, and the scores of valence given by the alpha wave asymmetry. When testing the correlation between the dimension scores and the color preference, as shown on Table 6.11, no statistically significance was revealed.

The tests showed that there is no statistical significance on the correlation between Warmth, Gregariousness, Excitement Seeking and Positive Emotion, and the preference score of any color tested. On the other hand, we found a **medium, positive correlation between Assertiveness score and the preference for color white** (r = .354, p = .013), and a **medium, negative correlation between Activity and black** (r = -.288, p = .045).

As well as for the self-reported color preference, we then tested the difference of valence scores between Extraversion groups, per color. The nonparametric Kruskal-Wallis H rank-based test was used due to the evidence against normality found in some of the samples (p <.05).

The results we obtained when performing the test, presented on Table 6.12, show that there was **no statistically significant difference of valence scores between Introverts, Ambiverts and Extraverts for any color except pink**. The mean ranks for this color were, respectively, 13.39, 29.20, and 21.25.

Since we cannot draw conclusions on these results, we tried to find tendencies in the data through visualization. The bar charts shown on Figure 6.4 show the median scores of each color for Introverts (left), Ambiverts (center), and Extraverts (right). Through these graphs, we intended to visualize general differences on the color preferences between these groups.

Based on this image, results suggest that:

- **Introverts' most liked color was cyan** (*Mdn* = 2.16), followed by **brown** (*Mdn* = 1.53), while **their least favorites were pink** (*Mdn* = 0.86), **gray** (*Mdn* = 0.89), and **orange** (*Mdn* = 0.93).

Color	Color Name	Correlation Coefficient	Significance
	Red	-.060	.682
	Orange	.148	.311
	Brown	-.129	.378
	Yellow	-.032	.829
	Green	.070	.632
	Cyan	-.182	.210
	Blue	-.024	.868
	Purple	.152	.297
	Pink	.140	.338
	White	.139	.339
	Gray	.048	.741
	Black	.189	.193

Table 6.11: Spearman's correlation between Extraversion score and valence score, by color.

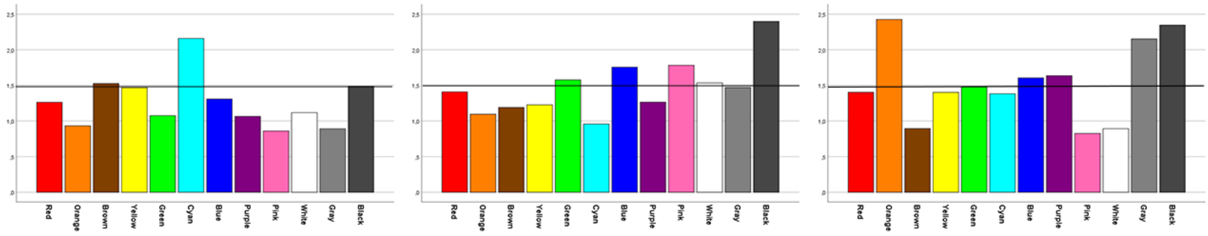

Figure 6.4: Median scores, per color, for Introverts (left), Ambiverts (center) and Extraverts (right).

- **Ambiverts' preferred colors** were **black** (*Mdn* = 2.40), **pink** (*Mdn* = 1.78), and **blue** (*Mdn* = 1.76), being **cyan** (*Mdn* = 0.96) and **orange** (*Mdn* = 1.10) **the least liked colors.**
- **Extraverts preferred orange** (*Mdn* = 2.43), **black** (*Mdn* = 2.35) and **gray** (*Mdn* = 1.47), **disliking brown** (*Mdn* = 0.89), **pink** (*Mdn* = 0.83), and **white** (*Mdn* = 0.89).
- The **highest score median values** belong to **orange, rated by Extraverts**, and **black, rated by Ambiverts**.
- The **lowest median scores** was given by **Introverts and Extraverts** to the color **pink**.
- Roughly comparing to the score value 1.50 – score that, when rounded to the Likert scale used for the self-reported preference, corresponds to neutral preference –, marked with a black line on the graphs:
  - **Introverts** scored **one color with a higher score**, **eight colors with a lower score**, and **three matches**.
  - For **Ambiverts**, **four colors were scored higher** than that value, **six were scored lower**, and **two matched that value**.
  - **Extraverts** liked **five colors more**, **six colors less**, and **one the same**.

To find trends on the comparison valence scores between Extraversion groups, by color, we drew the individual boxplots shown on Figure 6.5.

Observing the boxplots, we can detect the following tendencies:

Color	Color Name	Kruskal-Wallis H	Significance
	Red	0.141	.932
	Orange	2.781	.249
	Brown	1.773	.412
	Yellow	0.284	.868
	Green	1.732	.421
	Cyan	4.286	.117
	Blue	1.312	.119
	Purple	1.371	.504
	Pink	9.272	.010
	White	1.499	.473
	Gray	0.539	.764
	Black	1.400	.497

Table 6.12: Kruskal-Wallis test for valence scores between Extraversion groups.

Red

- The score **distribution for Introverts is limited comparing to Ambiverts and Extraverts**, suggesting a higher level of agreement.
- The **scores given by Introverts are mainly distributed between 0 (Strongly Dislike) and 2 (Neutral)**.

Orange

- Introverts' and Ambiverts' scores have a wider distribution than Extraverts' scores, pointing that **Extraverts agree more when rating orange**.
- Around **50% of the Introverts and Ambiverts rated the color below 1 (Dislike)**.
- Around **50% of the Extraverts scored above 2.5**, which is rounded to 3 **(Like)** on the Likert scale used for self-reported preference.
- The **median values showed to be increasing**, as the Extraversion score of each group.

Brown

- **Introverts show the widest distribution**, suggesting a smaller agreement.
- Although the **median score for Introverts is the highest** among the three groups, it shows that around **50% of the subjects rated the color below 1.5**, which marks the start of the **Neutral** score.
- **Introverts had the highest median score, followed by Ambiverts and Extraverts** (descending).

Yellow

- Although the **distribution for Ambiverts' scores is wider than Introverts' and Extraverts'**, the majority of the data is concentrated between 0 and 2.5, being the **median values identical** between groups.

Green

- Around **75% of the Introverts scored below 1.5 for the color green (below Neutral)**, which is close to both Ambiverts' and Extraverts' median score.
- **Ambiverts showed to have a smaller agreement** than Introverts and Extraverts.

Cyan

- Around **75% of the Introverts scored above 1 (Dislike)**, showing to be **the group which prefer the color the most**.

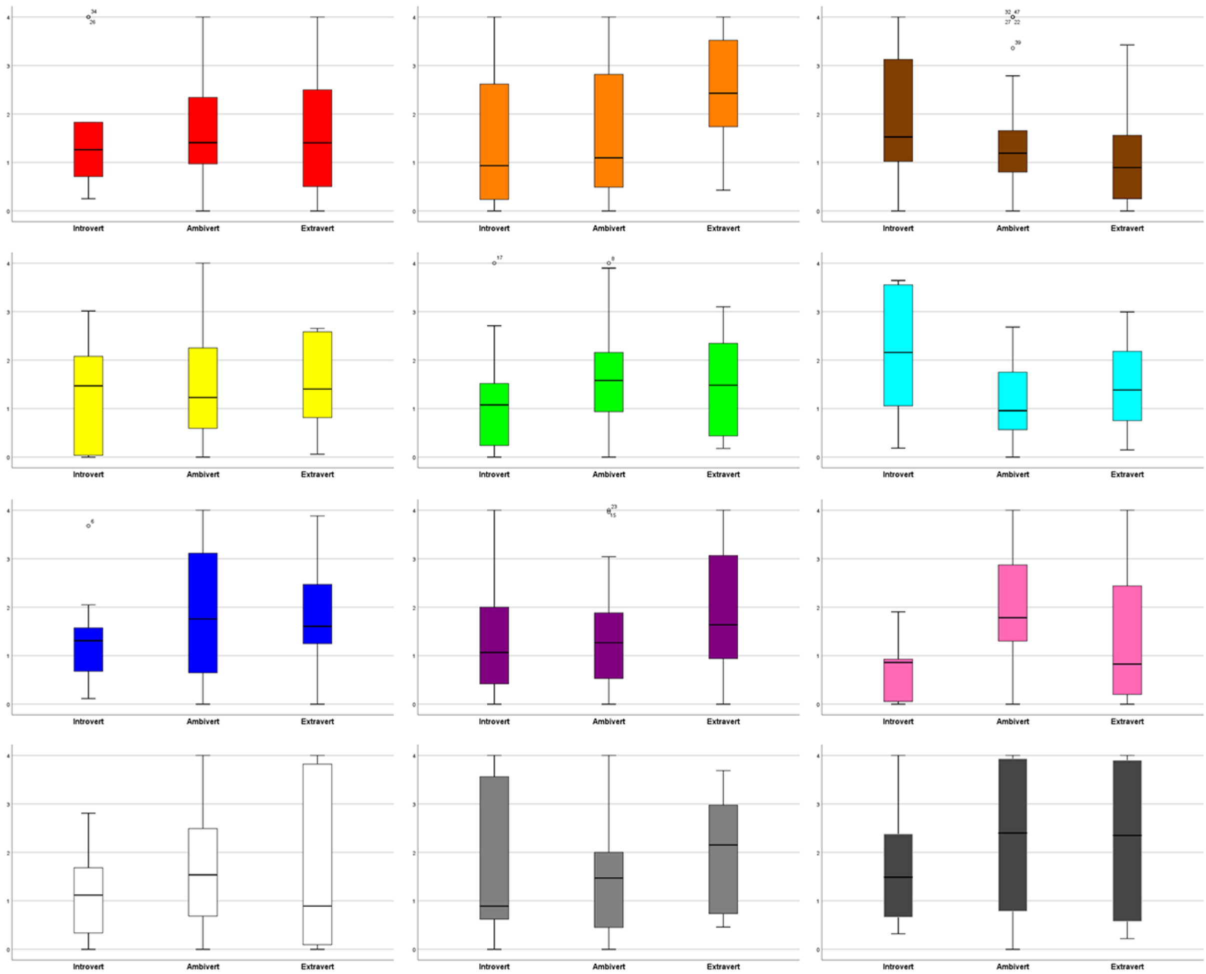

Figure 6.5: Boxplots representative of valence scores difference between Introverts, Ambiverts and Extraverts, per color.

- The **Ambivert and Extravert groups showed a bigger agreement**, rating the color between 0 and 3.

Blue
- **The median and extreme values for Ambivert and Extravert groups are identical**, but the **central 50% of the scores is more concentrated for the Extravert group**.
- The **Introvert group showed to like this color less**, since there is a higher agreement level derived by the data distribution between around 0 (Strongly Dislike) and 2 (Neutral).

Purple
- Although the distribution of scores for Introverts is wider, both **Introverts and Ambiverts have around 75% of the data distributed between 0 (Strongly Dislike) and 2 (Neutral)**.
- **Extravert subjects showed to agree the least** on the score for the color.
- The median score showed to be **lowest for Introvert and highest for Extraverts**.

Pink
- We can easily identify the statistically significant differences for this color. The **maximum score for Introverts is close to the median score for Ambiverts**, and both **Introverts' and Extraverts' median scores are similar**.

White
- Although the median value for the Extravert group is the lowest, around **25% of the subjects scored this color above 3.5 (rounded to 4, Strongly Like)**.
- **Introverts agreed more in terms of preference for this color**, showing a distribution between 0 (Strongly Dislike) and 3 (Like).

Gray
- For **Introverts,** the distribution was more **concentrated between 0 and 1 (Strongly Dislike and Dislike, around 50%)** and **above 3.5 (rounded to Strongly Like, around 25%)**.
- **Extraverts showed a higher level of agreement, for the central values**.
- There is a **increasing tendency for median values depending on the Extraversion scores** of each group.

Black
- For **Ambiverts and Extraverts, more than 25% of the scores were above 3.5 (Strongly Like)**.
- The **median value of scores is lower for Introverts** than for the other Extraversion groups.

Visualizing these graphs, we can also notice that:

- **Red, blue and pink are the colors which tend to be less liked by Introverts** (Neutral score or below).
- **Ambiverts tend to like brown and purple the least** (based on the highest scores, ignoring outliers).
- The colors with **lowest distribution values among Extraverts are yellow and cyan**.

6.4.3 Arousal through Heart Rate

At last, we analyzed the relations between Extraversion and the arousal scores given by HR. To this end, we started by testing correlations between the dimension and its facets and each color's arousal scores, applying the Spearman's test. As well as for self-reported preference and the color preference given by valence, and as shown on Table 6.13, the results were not statistically significant.

Color	Color Name	Correlation Coefficient	Significance
	Red	.137	.346
	Orange	.264	.066
	Brown	-.072	.623
	Yellow	.165	.257
	Green	.097	.507
	Cyan	.061	.676
	Blue	.115	.432
	Purple	.007	.963
	Pink	.046	.752
	White	.032	.828
	Gray	.099	.497
	Black	-.159	.276

Table 6.13: Spearman's correlation between Extraversion score and the arousal score for colors used in the study.

Applying the same correlation test between the facet **Warmth and the arousal score**, we found that:

- Between this **facet and the arousal score for orange**, there is a **medium, positive correlation** (r = .319, p = .025).
- **Yellow arousal score and Warmth have a medium, positive correlation** (r = .295, p = .039).

For all the other colors, there was no statistical significance. The same occurred for all the colors when testing the correlation between arousal and Gregariousness.

When using Spearman's correlation test on Assertiveness and arousal scores for all the colors, the only color that showed to be statistically significant was **black**. For this color, we found a **medium, negative correlation** (r = .295, p = .039).

For Activity and Excitement Seeking, no statistically significant results were found on the correlation between them and the arousal scores, for any color. Lastly, for **Positive Emotion**, a **medium, positive correlation was found between the facet and the arousal score for the color orange** (r = .352, p = .013). All the other colors scores had no statistically significant results.

Just as for the self-reported preference and the valence scores gathered through the asymmetry of the alpha waves, we ran a Kruskal-Wallis H rank test to verify differences between Introverts, Ambiverts and Extraverts scores in terms of arousal, per color. The results, shown on Table 6.14, show that there is no statistically significant difference for any of the colors ($p >$.05).

Color	Color Name	Kruskal-Wallis H	Significance
	Red	1.414	.493
	Orange	3.493	.174
	Brown	0.439	.803
	Yellow	0.612	.736
	Green	0.506	.776
	Cyan	0.267	.875
	Blue	1.339	.512
	Purple	4.827	.089
	Pink	1.567	.457
	White	0.544	.762
	Gray	0.749	.688
	Black	2.158	.340

Table 6.14: Kruskal-Wallis test for the arousal scores between Extraversion groups.

Since the Kruskal-Wallis H test results proved the differences of arousal between Extraversion were not statistically significant, we proceeded to analyze trends visually, through the boxplots shown on Figure 6.6. From these graphs we can see that:

Red
- The **minimum level of arousal felt by Extraverts has a higher score**, showing a higher level of agreement among Extraverts.
- The scores of **around half of the Extraverts** are **between 3 and 4**.
- It is **one of the more arousing colors for all the Extraversion groups**.

Orange
- Based on the score median value, **Extraverts feel more aroused by the color orange, followed by the Ambiverts and the Introverts** (increasing order).
- There is a **higher level of agreement on higher scores for Extraverts**.
- Around **25% of Introverts** got scores close to 0, suggesting they **considered this color very relaxing**.

Brown
- Based on the maximum and median scores, **Introverts and Extraverts did not get as**

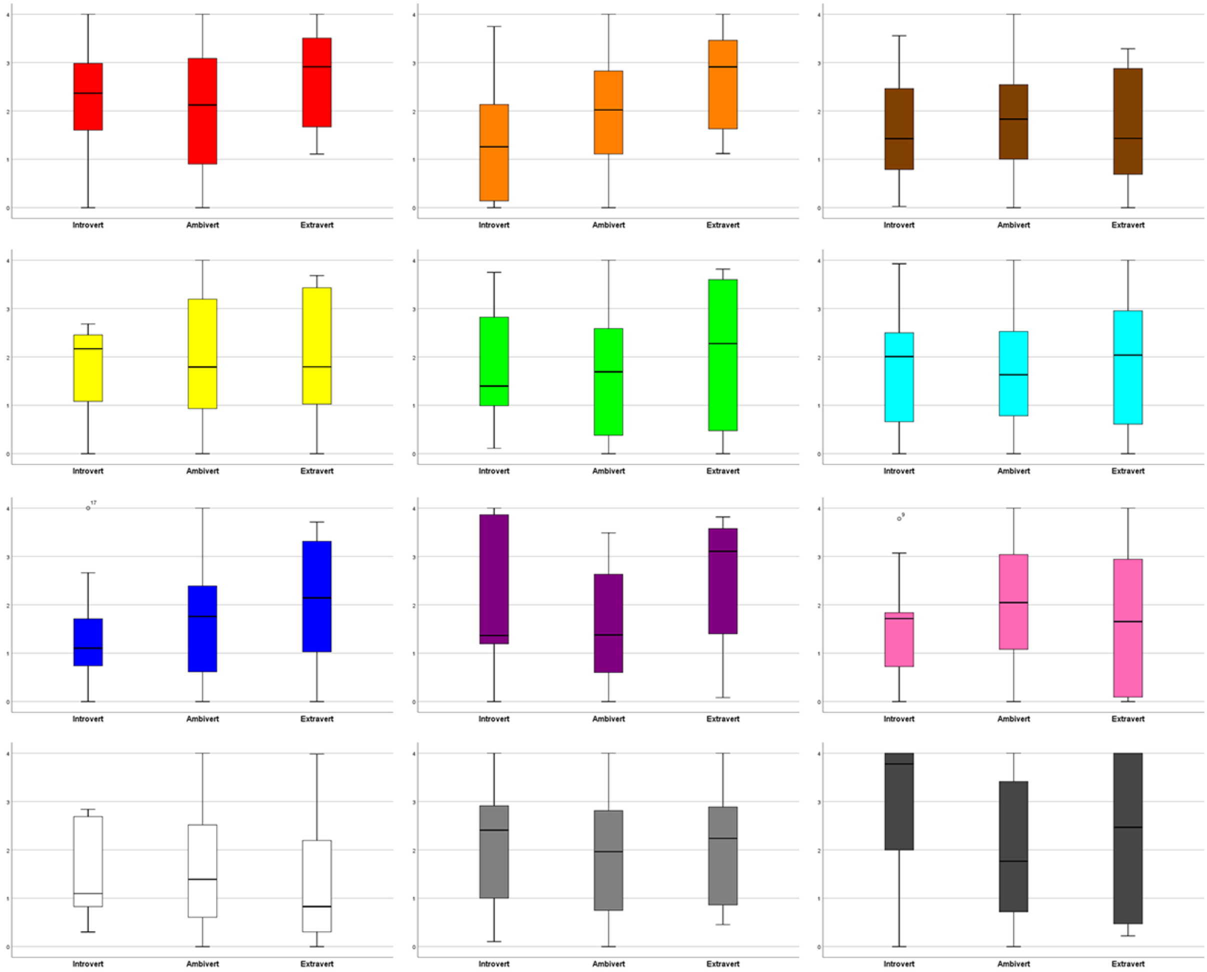

Figure 6.6: Boxplots representative of the heart rate difference between Introverts, Ambiverts and Extraverts, per color.

aroused as Ambiverts while observing this color.

Yellow
- Although the median value being higher for the Introvert group, **Introverts showed a higher level of agreement for lower values when reacting to this color in terms of arousal**.
- **Ambiverts and Extraverts scored this color similarly**.

Green
- The median scores of arousal between groups showed that there is a **increasing trend for this color**, suggesting that Introverts feel less aroused than Ambiverts when observing this color, who feel a lower arousal than Extraverts.
- More than half of the Introverts scored this color below 1 or above 3, suggesting a **lower agreement level**.

Cyan
- The users showed to **score this color identically** between groups of Extraversion.

Blue
- The median values showed a **increasing tendency between Extraversion groups** (Introverts < Ambiverts < Extraverts)
- **Introverts agreed more in finding this color less arousing** (excluding outliers).

Purple
- **Ambiverts found this color slightly less arousing** (based on the maximum score).

Pink
- Ignoring outliers, we can see a **higher agreement among Introverts**.

- Besides the different distributions, the **median scores were identical** between groups.

White
- The **maximum score for Introverts is below 3**, suggesting the subjects agreed more in lower scores.

Gray
- There is a **slightly higher agreement among Extraverts** for higher scores.
- **The distributions were similar between groups**, as well as the median score values.

Black
- All the groups showed a **discordance among users**, but around **50% of the scores recorded from Introverts were near the highest score**.

We can also observe that:

- **Introverts** showed to be **easily aroused by black and red**, **contrarily to white and blue**.
- Roughly, the **Ambivert group had Neutral median scores for all colors**.
- **Red, orange and purple are the most arousing colors for Extraverts**, while **white and brown are the least**.

6.5 Discussion

Although some results were not statistically significant, we worked on finding tendencies in the data. Having said this, it would be interesting to perform a study with a bigger, more diverse sample. Our sample had some limitations in terms of diversity, particularly in terms of age, nationality and Extraversion group. Since this study was performed in a university, the majority of our users were students, being that the reason why the majority are aged between 18 and 35 years old. In terms of nationality, the sample was restricted to Portuguese speakers, due to the personality questionnaire used being written in this language. The last limitation is the hardest to overcome, since the Extraversion group is assigned after completing a personality questionnaire, during the user tests. However, the more subjects participate in a study of this kind, the greater are the chances of getting a more balanced sample.

Regarding the distribution of our sample by Extraversion groups, the majority of the subjects were Ambiverts. This was expected, since Extraversion facets are very diverse and to get a higher score of Extraversion implies having a higher score in all of the facets, and the opposite for lower scores.

From this Ambivert group, almost half of the subjects classified themselves as Introverts or Extraverts. One of the reasons why this happens is because people are used to see Extraversion as a scale divided in two halves – Introverts and Extraverts. This was also noticed by the researcher, since during the user tests a large amount of subjects asked the meaning of “Ambivert”. Another cause is that rating personality on a scale is not an easy task, even more when people tend to associate Extraversion with social behaviors, underestimating facets like Activity and Positive Emotions. Lastly, each user is led to divide the Extraversion continuum in order to classify themselves. Since there is no validated division between the three groups, the division may differ between users.

Besides Extraversion, color is the main topic of this work. Therefore, in order to discuss the obtained results, we present Table 6.15, including the colors used and their representations using the models RGB and CIEL*C*h*. RGB was the model used to develop the interface, including the CPT, and CIEL*C*h* was chosen for being perceptually uniform and based on how the human eye perceives color [65]. This last model has a main advantage: it allows us to isolate the Hue channel, which we proposed to study within this work.

When focusing on color preference – self-reported and obtained from neurophysiological recordings –, the data suggested that the results both varied depending on the color. For instance, in general, white and black were considered two of the most liked colors through both self-report and valence scores.

Color	Color Name	RGB (hexa)	RGB			CIEL*C*h*		
			Red	Green	Blue	Lightness	Chroma	Hue
	Red	#ff0000	255	0	0	53,23	104,58	40,00º
	Orange	#ff8000	255	128	0	67,05	85,53	59,95º
	Brown	#804000	128	64	0	34,51	50,83	61,86º
	Yellow	#ffff00	255	255	0	97,14	96,91	102,85º
	Green	#00ff00	0	255	0	87,74	119,78	136,02º
	Cyan	#00ffff	0	255	255	91,12	50,12	196,39º
	Blue	#0000ff	0	0	255	32,30	133,82	306,29º
	Purple	#800080	128	0	128	29,78	69,33	328,23º
	Pink	#ff69b4	255	105	180	65,48	65,13	350,58º
	White	#ffffff	255	255	255	100,00	0,00	0,00º
	Gray	#808080	128	128	128	53,59	0,00	0,00º
	Black	#000000	0	0	0	0,00	0,00	0,00º

Table 6.15: Colors used as stimuli, represented in different color models.

They are neutral colors, with both hue and chroma values 0, but opposite Lightness values. These results suggest that the general color preference is not based on the Lightness channel.

On the other hand, we can see that the most liked color – blue –, and some of the colors self-reported as most liked – gray and purple – have medium-low Lightness (30-50), while all the least liked colors – yellow and green through self-reported, orange and cyan through valence scores – have high Lightness levels (>70). We also found that orange and cyan, considered two of the most liked colors through self-report, were, in fact, the least liked colors, and nothing besides the Lightness level differs from the most liked colors.

Overall, the results suggest that Extraverts and Ambiverts like a larger amount of colors than Introverts, supporting Birren's theory [3–5]. It is also suggested that Extraverts' scores tend to be higher than Introverts'.

In terms of arousal, derived from HR, the results suggested that the most arousing colors are red, orange and black, while the least arousing colors are white, purple and blue. In terms of color preference, white and blue – and purple, though self-report – are least arousing and have higher scores in terms of preference, while orange is one of the most arousing and least liked colors. Although red was not considered one of the most liked, neither one of the least liked colors, the results suggest that people, in general, tend to prefer least arousing colors.

In Figure 6.7, each color with Chroma different from 0 is represented in terms of median score of arousal and Hue. Observing the chart, we can notice that, following the Hue degree, the arousal level starts high (red) and decreases until its lower level (blue and purple) and then starts increasing again. Since Hue is represented as a circumference, where 360º is equivalent to 0º, the color pink is closer to red. These results also suggest that cold colors are least arousing than warm colors.

The only color whose result stands out while observing Figure 6.7 is brown, being a warm color and one of the least arousing colors. Besides the Hue tendency found, this color has something in common with the other two least arousing colors (blue and purple) – the Lightness. Except for black, these are the three colors with the lowest levels of Lightness (brown = 34.51, blue = 32.30, and purple = 29.78), suggesting that colors with lower lightness are less arousing colors.

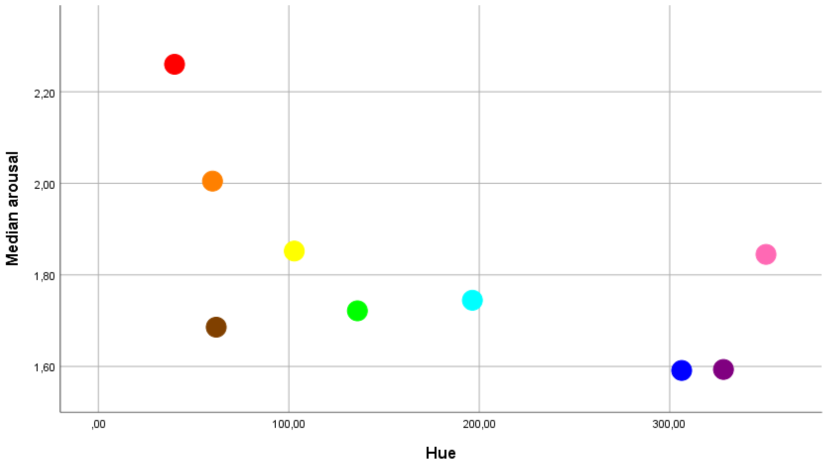

Figure 6.7: Median arousal levels of the colors by hue.

On the other hand, black (Lightness = 0) was considered the second more arousing color, following red. It is possible that this value is influenced by the fact that black was the first color shown to all the users. Since the first color is the first time the user deals with the test, understanding visually the explanation given by the researcher before, this can generate a higher level of anxiety. However, this result coincides with previous research, where black was proved to have the greatest arousal level, when comparing achromatic colors through self-reported ratings [64].

With respect to the correlations studied between Extraversion scores and color preference, the statistical tests revealed that the more assertive a person is, the more they like white, and the more active a person is, the less they like black. Since both are achromatic colors, the only difference between these two colors is Lightness. Having in mind that excitement seekers seek higher arousal levels, we expected to find a positive correlation in terms of preference between Excitement Seeking scores and colors with higher arousal level, which was not found.

Since the majority of the statistical tests results were not statistically significant, we observed trends in data. Orange and gray showed to be two of the most liked colors for Extraverts, while being two of the least liked colors for Introverts, and the opposite happens for brown. On the other hand, the most liked color by Introverts (cyan) is not disliked by Extraverts, and the second preferred color for Extraverts (black) is not disliked by Introverts. Having said this, we can not claim that Introverts and Extraverts dislike each others' most liked colors. However, the results suggest that Introverts and Extraverts do not share the same favorite colors, and their most liked colors – cyan for Introverts and orange for Extraverts – are opposite in the color wheel.

Comparing color preference between Extraversion groups, Introverts and Extraverts show opposite preference for some colors. For orange, brown and purple, the preferences show an increasing or decreasing tendency. However, the Ambiverts' preference for those colors is not always between the two extreme groups. The statistical analysis showed that the Ambivert group appreciate pink more than Introverts and Extraverts. Having this in mind, it would be interesting to go further on studying Ambiverts' preference.

Speaking in terms of color temperature, we can see that Extraverts' favorite color (orange) is a warm color, while Introverts' favorite color (cyan) is a cool color, which is consistent with Birren's theory that warm colors are usually preferred by Extraverts while Introverts prefer cool colors [3–5]. However, this is not true, for instance, for brown (warm color), which is one of both Introverts' most liked and Extraverts' least liked color.

In terms of arousal, the comparison between Extraversion groups revealed that all the groups considered red one of the most arousing colors, and white one of the least arousing colors. As well as preference, some colors have opposite arousing effects in Introverts and Extraverts, but there is not always an increasing or decreasing trend.

However, the results suggest some trends between the arousal levels and preferences by Extraversion groups. For instance, none of the colors that caused more arousal on Introverts is one of their favorites. Supporting this, there is a color (gray) which is one of the most arousing and least liked by Introverts. The opposite happens for Extraverts – their most liked color is one of the most arousing colors for the group, while white is considered a relaxing color and one of their least liked. Since Extraverts tend to seek for outside stimulation, while Introverts tend to protect themselves from that stimulation, this finding matches the characteristics of Extraversion.

7

Conclusions and Future Work

Color and personality have been widely studied, being of extreme importance in human life. While personality studies try to define a representation of human character and behavior, the research on color tries to understand visual perception and how color influences our lives.

Nowadays, graphic displays are widely used to accomplish the most varied tasks. Therefore, it is crucial to understand color and study the most appropriate ways of representing color in these devices, in order to present information clearly.

This work has allowed us to study valence and arousal of color stimuli, not only in a general way but also in terms of the personality dimension Extraversion. We found that users with different Extraversion levels have distinct emotional reactions to color stimuli. We also found that, depending on the personality traits, the approach/withdrawal reaction to the stimuli is related to the arousal level felt. In this case, we found that Extraverts' favorite colors cause them higher arousal, while Introverts tend to prefer more relaxing colors. Correlations between Extraversion facets and the preference for some colors were also found.

We could also verify that Introverts' most liked color is on the opposite side from Extraverts' on the color wheel. Arousal showed to be related to the Hue angle and, consequently, the color temperature – warm colors showed to be more arousing than cool colors. On the other hand, colors with lower Lightness level showed to be more relaxing (except for black).

In terms of the self-reported preference, the study showed that the responses were very different from the preference collected through neurophysiological responses. This reinforces that self-reports include not only subject's preference but also their culture and context.

Although the results suggest some tendencies, we could not draw conclusions based on the statistical analysis. In light of this, it would be interesting to continue the studies on the influence of personality in color preferences. Future works on this field of study should be based on a larger, more diverse sample.

It would be interesting to design the study with more options in terms of languages, to allow a wider cultural diversity. It is very important to have in mind that, to gather data in terms of Extraversion, a personality test shall be answered honestly, through the most accurate options. Therefore, a test performed in the subjects' native language may lead the researchers to getting more reliable results. In order to maintain the study coherence, the entire test should be performed in the users' native language, meaning that further studies should include several languages to cover more cultures.

Further studies should also include a wider diversity in terms of the other demographic information, such as age and education. Since self-reported color preference may be influenced by social and cultural factors, it would be interesting to underlie the studies on samples with more diverse information.

It would also be crucial to have a more balanced sample in terms of Introverts, Ambiverts and Extraverts. Since this is not easy to guarantee in advance, the wider the sample, the greater chances of getting the desired results. Regarding personality, it would also be interesting to focus on Ambiversion differences. For instance, dividing the Ambivert group into subgroups would allow the analysis between Ambiverts with higher and lower scores, as well as with Introverts and Extraverts, to study any preference and arousal similarities between the closest groups.

The study of the influence of personality in color preference can even be extended beyond Extraversion. Neuroticism, Openness to experience, Agreeableness and Conscientiousness are the other dimensions of the adopted personality model, being divided in six facets each, meaning this field show a wide range of options for research.

Our results suggested some trends in terms of preference and arousal when comparing single channels of the CIEL*C*h* color representation. To go further studying the influence of color properties, it would be interesting to compare reactions between colors changing one channel at a time and maintaining the two remaining channels with the same value. This way the researchers could explore, for instance, the influence of Hue itself, not depending on the other properties of the colors.

With the results obtained from further studies, new guidelines may be derived for varied HCI applications. Regardless of the application chosen, the adaptation to each user's personality should be made through a validated personality questionnaire. In contrast with emotion – a psychological topic widely researched for HCI uses –, personality is lifelong tailored. This means that the questionnaire should be answered from time to time, but sudden changes are not expected.

Bibliography

[1] Soraia M Alarcão and Manuel J Fonseca. Emotions recognition using eeg signals: a survey. *IEEE Transactions on Affective Computing*, 2017.

[2] Anupama Batra, Sangeeta Vyas, Dheeraj Jeph, et al. Color difference on simple visual reaction time in young volunteers. *International Journal of Clinical and Experimental Physiology*, 1(4):311, 2014.

[3] Faber Birren. *Color Psychology and Color Therapy: A Factual Study of the Influence of Color on Human Life.*, volume 25. LWW, 1950.

[4] Faber Birren. *Selling color to people*. University Books, 1956.

[5] Faber Birren. *Color*. University Books, 1963.

[6] Faber Birren. Color preference as a clue to personality. *Art Psychotherapy*, 1(1):13–16, 1973.

[7] Ake Bjerstedt. Warm-Cool Color Preferences As Potential Personality Indicators: Preliminary Note. *Perceptual and Motor Skills*, 10:31–34, 1960.

[8] CMJ Braun and JL Bonta. Cross-cultural validity, reliability, and stimulus characteristics of the Lüscher Color Test. *Journal of personality assessment*, 43(5):459–460, 1979.

[9] Scott Brave and Clifford Nass. Emotion in human–computer interaction. *Human-Computer Interaction*, page 53, 2003.

[10] Antoine Choppin. Eeg-based human interface for disabled individuals: Emotion expression with neural networks. *Unpublished master's thesis*, 2000.

[11] Assador Choungourian. Introversion-Extraversion and Color Preferences. (August 2014), 2010.

[12] Donald Cohen and James P Schmidt. Ambiversion: characteristics of midrange responders on the introversion-extraversion continuum. *Journal of Personality Assessment*, 43(5):514–516, 1979.

[13] Edmund S Conklin. The definition of introversion, extroversion and allied concepts. *The Journal of Abnormal Psychology and Social Psychology*, 17(4):367, 1923.

[14] Paul T Costa and Robert R MacCrae. *Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO-FFI): Professional manual*. Psychological Assessment Resources, Incorporated, 1992.

[15] O da Pos and Paul Green-Armytage. Facial expressions, colours and basic emotions. *Journal of the International Colour Association*, 2007.

[16] Antonio R. Damasio. *Descartes' Error: Emotion, Reason, and the Human Brain*. Harper Perennial, 1995.

[17] Richard J Davidson. Anterior cerebral asymmetry and the nature of emotion. *Brain and cognition*, 20(1):125–151, 1992.

[18] John M Digman. Personality structure: Emergence of the five-factor model. *Annual review of psychology*, 41(1):417–440, 1990.

[19] Alan Dix. Human-computer interaction. In *Encyclopedia of database systems*, pages 1327–1331. Springer, 2009.

[20] Frank A Donnelly. The Lüscher Color Test: reliability and selection preferences by college students. *Psychological reports*, 34(2):635–638, 1974.

[21] Frank A Donnelly. The Lüscher Color Test : A Validity Study. 44:17–18, 1977.

[22] Mariam Adawiah Dzulkifli and Muhammad Faiz Mustafar. The influence of colour on memory performance: A review. *The Malaysian journal of medical sciences: MJMS*, 20(2):3, 2013.

[23] Allen Louis Edwards. Edwards personal preference schedule. 1954.

[24] HJ Eysenck. *Dimensions of Personality*. The international library of psychology. Routledge & Kegan Paul, 1950.

[25] HJ Eysenck. Manual of the maudsley personality inventory, 1959.

[26] HJ Eysenck. *The Biological Basis of Personality*. Transaction Publishers, 2006.

[27] HJ Eysenck and SBG Eysenck. *Psychoticism as a dimension of personality*. Crane, Russak, 1977.

[28] HJ Eysenck and SGB Eysenck. The eysenck personality inventory. 1965.

[29] Andy Field and Graham Hole. *How to design and report experiments*. Sage, 2002.

[30] Manuel J Fonseca, Pedro Campos, and Daniel Gonçalves. Introdução ao design de interfaces. *FCA Editora. ISBN*, pages 978–972, 2012.

[31] Stiliyan Yordanov Georgiev, Christina Venelinova Christov, and Dolja Toshkova Philipova. Ambiversion as independent personality characteristic. *Act. Nerv. Super. Rediviva*, 56(3-4):65–72, 2014.

[32] Adam M Grant. Rethinking the extraverted sales ideal: The ambivert advantage. *Psychological Science*, 24(6):1024–1030, 2013.

[33] Tear Marie Green and Brian Fisher. Towards the personal equation of interaction: The impact of personality factors on visual analytics interface interaction. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on*, pages 203–210. IEEE, 2010.

[34] Robert Heiss, Petra Halder, and Diether Höger. *Der Farbpyramidentest*. H. Huber Bern, Stuttgart, Wien, 1975.

[35] Ewald Hering. Outlines of a theory of the light sense. 1964.

[36] John F. Hughes, Andries Van Dam, Morgan MCGuire, David F. Sklar, James D. Foley, Steven K. Feiner, and Kurt Akeley. *Computer Graphics Principles and Practice*, volume 53. 2014.

[37] Shinobu Ishihara. *Tests for colour-blindness*. Kanehara Shuppan Company, 1960.

[38] Egbert Jacobson. Basic color: An interpretation of the ostwald color system. 1948.

[39] Carl Gustav Jung. *Psychological types*. Routledge, 2014.

[40] P Lang and Margaret M Bradley. The international affective picture system (iaps) in the study of emotion and attention. *Handbook of emotion elicitation and assessment*, 29, 2007.

[41] Peter J Lang. The emotion probe: Studies of motivation and attention. *American psychologist*, 50(5):372, 1995.

[42] P.J. Lang, M.M. Bradley, and B.N. Cuthbert. International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical report, NIMH Center for the Study of Emotion and Attention, 1997.

[43] Haim Levkowitz. Color Theory and Modeling for Computer Graphics, Visualization, and Multimedia Applications. *Springer*, 402:236, 1997.

[44] MP Lima and A Simões. Neo-pi-r manual profissional. *Lisboa: CEGOC*, 2000.

[45] Yisi Liu, Olga Sourina, and Minh Khoa Nguyen. Real-Time EEG-Based Emotion Recognition and Its Applications. *Transactions on Computational Science*, 2011.

[46] M Lüscher. *The Lüscher color test*. Pocket Books. Random House, 1969.

[47] RR McCrae and OP John. An introduction to the five-factor model and its applications. pages 175–215, 1992.

[48] Isabel Briggs Myers, K C Briggs, and Consulting Psychologists Press. *The Myers-Briggs Type Indicator: Supplementary manual*. Consulting Psychologists Press, 1976.

[49] J. Gordon Nelson, Marc T. Pelech, and Stephen F. Foster. Color Preference and Stimulation Seeking. *Perceptual and Motor Skills*, 59(3):913–914, 1984.

[50] T Obonai and T Matsuoka. Manual of the color symbolism personality test. *Tokyo: Kobunshya*, 1952.

[51] Tarao Obonai and Takeshi Matsuoka. The Color Symbolism Personality Test. *The Journal of General Psychology*, 55(April):229–239, 1956.

[52] Rosalind W Picard. Affective computing. 1995.

[53] R. Plutchik. The Nature of Emotions. *American Scientist*, 2001.

[54] Jonathan Posner, James A Russell, and Bradley S Peterson. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, 17(03):715–734, 2005.

[55] D. J. Radeloff. Psychological Types, Color Attributes, and Color Preferences of Clothing, Textiles, and Design Students. *Clothing and Textiles Research Journal*, 9(3):59–67, 1991.

[56] Carl Robinson. Color Preference As A Function Of Introversion And Extraversion. page 10016, 1975.

[57] Tova Rosenbloom. Color preferences of high and low sensation seekers. *Creativity Research Journal*, 18(2):229–235, 2006.

[58] J.A. Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39:1161–1178, 1980.

[59] DH Saklofske, HJ Eysenck, SBG Eysenck, RM Stelmack, and W Revelle. Extraversion–introversion. 2012.

[60] Inês Santos. CogniHue: Studying the cognitive effect of color in HCI. Master's thesis, Instituto Superior Técnico, Portugal, 2017.

[61] K Warner Schaie. On the relation of color and personality. *Journal of Projective Techniques and Personality Assessment*, 30(6):512–524, 1966.

[62] David V Stimpson and Miriam F Stimpson. Relation of personality characteristics and color preferences. *Perceptual and motor skills*, 49(1):60–62, 1979.

[63] Robert Milton Taylor et al. *Taylor-Johnson Temperament Analysis manual*. Los Angeles: Psychological Publications, 1968.

[64] Patricia Valdez and Albert Mehrabian. Effects of color on emotions. *Journal of experimental psychology: General*, 123(4):394, 1994.

[65] Colin Ware. *Information Visualization: Perception for Design*. 2004.

[66] Thomas Young. On the theory of light and colors. *Philosophical Transactions of the Royal Society*, 91(January):12–49, 1802.

[67] Marvin Zuckerman, Elizabeth A Kolin, Leah Price, and Ina Zoob. Development of a sensation-seeking scale. *Journal of consulting psychology*, 28(6):477, 1964.

Appendix A

ColorCode - Pilot Testing Protocol

Joana Condeço[1], Sandra Gama[2], and Daniel Gonçalves[3]

[1]joana.condeco@tecnico.ulisboa.pt
[2]sandra.gama@tecnico.ulisboa.pt
[3]daniel.goncalves@inesc-id.pt
[+]Instituto Superior Técnico – Taguspark, 2744-016 Porto Salvo, Portugal

The following document is the pilot test protocol used for the *ColorCode* project, which purpose is to explore the influence of personality (Extraversion) in color preferences. This protocol was made to test the system, in order to create a solid final study protocol. To achieve this end, the user test will be divided in four tasks, which will allow a better refinement of the specific steps of the test.
This document does not only include the steps intended for the study, but also usability metrics and specifications on the system testing.

Room Conditions

Since the study is based on human perception and reactions, controlled room conditions are a crucial aspect to guarantee its consistency. It is important that, during a user test, the person feels comfortable, in order to produce more accurate test results.

In order to achieve this, the tests will occur in a room without daylight, with fully controlled illumination and temperature. During each test, only the developer and the user shall be in the room. In case this is not possible, the presence of others must be unnoticeable, not disturbing the test. During the test, the room shall be quiet, except when the user has questions to ask.

All the tests will be performed using the same computer and monitor, in order to provide the same resolution and color calibration of the monitor, which is a crucial factor of the user tests. The calibration step shall be made before the tests, using, for instance, the color calibration settings available in some Operating Systems.

Material

Computers

- **Neurophysiological data collection:** Any computer with an Operating System compatible with the software OpenSignals (r)evolution (e.g., Win 32-bit v.2017), specialized for data acquisition and visualization.
- **Questionnaire data collection:** Any computer with an Operating System compatible with Google Chrome (e.g., version 60.0.3112.101) to access the questionnaire[1].

Neurophysiological responses

Toolkit BITalino[2], using:

- **General:**
  - Ethanol
  - Cotton discs
  - Ten20 conductive - neurodiagnostic electrode paste
- **Electroencephalography:**
  - Electroencephalography (EEG) sensor
  - 3-lead accessory
  - Gelled self-adhesive disposable Ag/AgCl electrodes (x3)

[1]http://web.ist.utl.pt/ist168624/colorcode.html
[2]http://bitalino.com/en/

– Sensor cable

- **Electro-dermal Activity:**
  – Electro-dermal Activity (EDA) sensor
  – 2-lead accessory
  – Gelled self-adhesive disposable Ag/AgCl electrodes (x2)
  – Sensor cable

- **Heart Rate: Pulse sensor composed by:**
  – Sensor connection cable (1m)
  – Ear clip

Introduction

The user test will start with a brief introduction to make the user more comfortable In this step, the test developer will introduce themselves and any other observer in the room. Next, the pilot test purpose and goal will be explained to the user, clarifying that the intention is to test the system and not the user (existing no right or wrong answers) and has the chance to stop any time.

The estimated duration of the test will also be mentioned, but since this is a pilot test, the duration will rely on the user's opinion about the system. Lastly, the user is asked to sign a consent form.

Greeting Script

Olá, [user's name].

O meu nome é [developer's name] e este teste piloto servirá para avaliar o sistema desenvolvido para a recolha de dados para a minha tese de mestrado, *Colorcode*. Com este trabalho, vamos explorar a influência que a personalidade tem nas preferências de cor das pessoas. Como tal, o teste incluirá um questionário demográfico, um questionário de personalidade e um questionário de preferências de cor. Para além deste último questionário, vamos recolher também respostas neurofisiológicas para avaliar as preferências de cor.

Para avaliar o sistema, o teste será dividido em quatro fases. Assim, poderemos avaliar individualmente cada fase do teste. No final de cada fase, será pedido que responda a um questionário de satisfação, onde poderá dar recomendações de melhoramento.

Não existem respostas certas ou erradas às questões colocadas ao longo do teste, e por isso pedimos que responda às questões da forma mais fiel possível. Em caso de dúvida, pode chamar-me a qualquer momento. Não será possível alterar respostas dadas em páginas anteriores, por isso pedimos que verifique as suas respostas antes de avançar entre testes.

O teste tem a duração estimada máxima de 45 minutos mas por se tratar de um teste piloto, o tempo pode variar de acordo com o desempenho do sistema. Caso deseje, tem também a opção de desistir a qualquer momento.

Para participar no teste, pedimos que assine esta declaração de consentimento.

[Give the user the consent form]

Consent Form

Eu, ____, declaro que aceito participar nos testes com utilizadores do projecto *ColorCode*, tendo conhecimento dos seguintes aspectos:

- Serão guardados os meus dados demográficos, de personalidade e de preferências de cor (auto-avaliação e respostas neurofisiológicas)
- Os meus dados serão tratados e utilizados de forma anónima, apenas para fins académicos.

________________ ____

Data Assinatura

2/4

Profiling Questionnaire and Ishihara Test

The profiling questionnaire includes 12 questions – six of them related to the demographic profile of the user and the remaining as a six plate Ishihara color blindness test[1].

The demographic questions are all of short answer, being answered as choice of drop-down menus.

For the color blindness test, the user may insert the number they see in a textbox or select the button "Não é um número" when they cannot see any number.

Personality Questionnaire

This stage consists of the Portuguese version of the NEO-PI-R personality test[2,3] reduced to Extraversion. Before the test, we explain to the user that they are not being tested and there are no right or wrong answers, asking them to describe themselves honestly and state their opinions as accurately as possible.

BITalino Installation and Calibration

After completing the Profiling and Personality questionnaires, the apparatus will be set up in order to proceed with the color preference test.

Before the application of the electrodes, it is important to guarantee the skin is clean, in order to collect the best responses possible. To this end, we will give the user cotton and ethanol and ask them to clean a few specific spots, where the electrodes will be applied.

The 3-lead accessory identifies the EEG electrodes with the colors black (negative), red (positive), and white (reference). The black and red electrodes will be applied on the left and right side of the user's forehead, respectively, and the reference one behind the ear, a zone of neutral brain activity. Since BITalino measures EEG using differential amplification, it is important that both black and red electrodes are placed in the same place of the user's head, but in opposite brain hemispheres, to measure the brain activity difference between hemispheres. In order to do this, we will use the nose and the eyebrows as reference points, applying each electrode one finger above the side eyebrow, one finger from the nose line.

For EDA, the 2-lead accessory will be applied on the thenar eminence of the user's left hand, the black (negative) one closer to the user's thumb and the red (positive) one closer to the wrist, with one finger of distance.

Lastly, the pulse sensor will be clipped to the user's right earlobe.[3] When the apparatus is ready, it is important to check if the signals' visualization is according to the expected. If it is not, the electrodes' positions shall be checked and corrected.

The user will be then presented some neutral images from the International Affective Picture System (IAPS) for three minutes, in order to neutralize their emotions and record their neurophysiological signals as a baseline for posterior analysis.

The procedures for gathering the neurophysiological responses were planned according with the *WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects*[4].

Color Preference Test

With the apparatus ready and after calibrating the user's emotions, the user will be asked to perform the color preference test.

In this test, 12 will be shown, random and individually. For each color, the following steps will occur:

1. The color is shown for five seconds on full screen, then a Likert scale will appear
2. During the next ten seconds, the user must evaluate the color on the scale, according to their preference.

Before the test, a brief explanation and some recommendations will be given to the user, in order to ease the further data treatment. For instance, it is important that the user focus their attention on the circle shown in the middle of the screen, trying not to blink or move their eyes during the first five seconds of each color.

Color Preference Test Script

[user's name], neste momento está pronto a iniciar o teste de preferências de cor. Neste teste serão apresentadas 12 cores, individualmente. Primeiro, a cor vai aparecer em *full screen* por 5 segundos e depois deste período vai aparecer uma escala idêntica à utilizada no teste de personalidade, onde deverá indicar o quão gosta da cor apresentada.A avaliação de todas as

[3]This configuration may suffer changes in case the signals are not correctly recorded.
[4]https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects

cores é obrigatória e o tempo de resposta é fixo. Antes de passar à cor seguinte, aparecerá um *countdown*, de modo a ter conhecimento do tempo restante.

Enquanto a cor ocupar o ecrã inteiro, peço que se concentre no círculo presente no centro do ecrã, evitando mover os olhos e pestanejar. Peço ainda que tente manter a mesma distância do ecrã ao longo de todo o teste.

Quando estiver preparado para iniciar o teste, clique em "Iniciar Teste".

End

In the end of the color preference test, the electrodes will be removed from the user's body and the skin will be cleaned with cotton and ethanol. Then, the test developer will thank the user's participation in the study, giving them a symbolic reward.

Metrics

Since the pilot testing phase is of huge importance for a good user testing phase, a set of metrics were defined in order to appraise the system. The metrics chosen for the different tasks are *number of mistakes*, *time*, and *comments*. These are all applied to the profiling and personality tests. The color preference test is time-controlled, therefore no efficiency metric is applied to this test.

Time will also be a metric used during the BITalino installation task. In contrast to the other tasks, the installation will be performed by the developer. Nevertheless, a time estimate for this task is necessary to further tests. All of the tasks will also include a satisfaction metric, in which the user will be asked to rate the experience in a Likert scale and to leave some comments on how they would improve the task or the system.

In addition to these individual tasks, effectiveness and efficiency metrics were also established for the user test as a whole, in order to estimate the duration of each session of user testing.

Table of Metrics

Task	Type	Metric	Minimum	Target	Optimal
Profiling	Effectiveness	Num. of mistakes (Observation)	2	0	0
	Efficiency	Time (Observation)	90 sec	60 sec	45 sec
	Satisfaction	Comments (Interview)	-	-	-
Personality	Effectiveness	Num. of mistakes (Observation)	4	2	0
	Efficiency	Time (Observation)	7 min	5 min	3 min
	Satisfaction	Comments (Interview)	-	-	-
BITalino Installation	Efficiency	Time (Observation)	10 min	8 min	6 min
	Satisfaction	Comments (Interview)	-	-	-
Color Preference	Effectiveness	Num. of mistakes (Observation)	4	2	0
	Satisfaction	Comments (Interview)	-	-	-
Total (incl. Introduction)	Effectiveness	Num. of mistakes (Observation)	5	2	0
	Efficiency	Time (Observation)	45 min	30 min	20 min
	Satisfaction	Comments (Interview)	-	-	-

Material

To record the current values for these metrics, only a stopwatch, a notebook and a pen will be needed. The developer shall use the stopwatch to measure *time* and write the results on the notebook, as well as the number of mistakes performed during the tasks and the comments made at the end of each one.

References

1. Alwis, D. V. & Kon, C. H. A new way to use the Ishihara test. *J. neurology* **239**, 451–454 (1992).

2. Costa, P. T. & MacCrae, R. R. *Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO-FFI): Professional manual* (Psychological Assessment Resources, Incorporated, 1992).

3. Lima, M. & Simões, A. Inventário de personalidade neo revisto (neo-pi-r). *Avaliação psicológica. Instrumentos validados para a população portuguesa* **1**, 15–32 (2003).

Appendix B

Colorcode - Study Protocol

Joana Condeço[1], Sandra Gama[2], and Daniel Gonçalves[3]

[1]joana.condeco@tecnico.ulisboa.pt
[2]sandra.gama@tecnico.ulisboa.pt
[3]daniel.goncalves@inesc-id.pt
[+]Instituto Superior Técnico – Taguspark, 2744-016 Porto Salvo, Portugal

The following document is the study protocol used for the *ColorCode* project, which purpose is to explore the influence of personality (Extraversion) in color preferences. The document was created after the performance of a pilot study, in which effectiveness, efficiency and satisfaction metrics were defined and tested. In order to replicate the study, the described steps must be followed.

Room Conditions

Since the study is based on human perception and reactions, controlled room conditions are a crucial aspect to guarantee its consistency. It is important that, during a user test, the person feels comfortable, in order to produce more accurate test results.

In order to achieve this, the tests will occur in a room without daylight, with fully controlled illumination and temperature. During each test, only the developer and the user shall be in the room. In case this is not possible, the presence of others must be unnoticeable, not disturbing the test. During the test, the room shall be quiet, except when the user has questions to ask.

All the tests will be performed using the same computer and monitor, in order to provide the same resolution and color calibration of the monitor, which is a crucial factor of the user tests. The calibration step shall be made before the tests, using, for instance, the color calibration settings available in some Operating Systems.

Material

Computers

- **Neurophysiological data collection:** Any computer with an Operating System compatible with the software OpenSignals (r)evolution (e.g., Win 32-bit v.2017), specialized for data acquisition and visualization.
- **Questionnaire data collection:** Any computer with an Operating System compatible with Google Chrome (e.g., version 60.0.3112.101) to access the questionnaire[1].

Neurophysiological responses

Toolkit BITalino[2], using:

- **General:**
  - Ethanol
  - Cotton discs
  - Ten20 conductive - neurodiagnostic electrode paste
- **Electroencephalography:**
  - Electroencephalography (EEG) sensor
  - 3-lead accessory
  - Gelled self-adhesive disposable Ag/AgCl electrodes (x3)

[1]http://web.ist.utl.pt/ist168624/colorcode.html
[2]http://bitalino.com/en/

- Sensor cable

- **Heart Rate: Pulse sensor composed by:**
  - Sensor connection cable (1m)
  - Ear clip

Introduction

The user test will start with a brief introduction to make the user more comfortable In this step, the test developer will introduce themselves and any other observer in the room. Next, the test purpose and goal will be explained to the user, clarifying that the user is not being tested (existing no right or wrong answers) and has the chance to stop any time.

The duration of the test will also be mentioned and, lastly, the user is asked to sign a consent form.

Greeting Script

Olá, [user's name].

O meu nome é [developer's name] e este teste faz parte da recolha de dados para a minha tese de mestrado, Colorcode. Com este trabalho, vamos explorar a influência que a personalidade tem nas preferências de cor das pessoas. Para tal, o teste estará dividido em quatro fases que permitem a recolha dos dados necessários para relacionar estes dois aspectos. Entre fases, haverá um momento de pausa, em que darei uma pequena explicação sobre a fase seguinte.

Não existem respostas certas ou erradas às questões colocadas ao longo do teste, e por isso pedimos que responda às questões da forma mais fiel possível. Em caso de dúvida, pode chamar-me a qualquer momento. Não será possível alterar respostas dadas em páginas anteriores, por isso pedimos que verifique as suas respostas antes de avançar entre testes.

O teste tem uma duração estimada de 30 minutos, que podem variar consoante a duração de cada fase. Caso deseje, tem também a opção de desistir a qualquer momento.

Para participar no teste, pedimos que assine esta declaração de consentimento.

[Give the user the consent form]

Consent Form

Eu, ____, declaro que aceito participar nos testes com utilizadores do projecto *ColorCode*, tendo conhecimento dos seguintes aspectos:

- Serão guardados os meus dados demográficos, de personalidade e de preferências de cor (auto-avaliação e respostas neurofisiológicas)

- Os meus dados serão tratados e utilizados de forma anónima, apenas para fins académicos.

________________ ____

Data Assinatura

Profiling Questionnaire and Ishihara Test

The profiling questionnaire includes 12 questions – six of them related to the demographic profile of the user and the remaining as a six plate Ishihara color blindness test[1].

The demographic questions are all of short answer, being answered as choice of drop-down menus.

For the color blindness test, the user may insert the number they see in a textbox or select the button "Não é um número" when they cannot see any number.

Personality Questionnaire

This stage consists of the Portuguese version of the NEO-PI-R personality test[2,3] reduced to Extraversion. Before the test, we explain to the user that they are not being tested and there are no right or wrong answers, asking them to describe themselves honestly and state their opinions as accurately as possible.

BITalino Installation and Calibration

After completing the Profiling and Personality questionnaires, the apparatus will be set up in order to proceed with the color preference test.

Before the application of the electrodes, it is important to guarantee the skin is clean, in order to collect the best responses possible. To this end, we will give the user cotton and ethanol and ask them to clean a few specific spots, where the electrodes will be applied.

The 3-lead accessory identifies the EEG electrodes with the colors black (negative), red (positive), and white (reference). The black and red electrodes will be applied on the left and right side of the user's forehead, respectively, and the reference one behind the ear, a zone of neutral brain activity. Since BITalino measures EEG using differential amplification, it is important that both black and red electrodes are placed in the same place of the user's head, but in opposite brain hemispheres, to measure the brain activity difference between hemispheres. In order to do this, we will use the nose and the eyebrows as reference points, applying each electrode one finger above the side eyebrow, one finger from the nose line.

Lastly, the pulse sensor will be clipped to the user's right earlobe[3]. When the apparatus is ready, it is important to check if the signals' visualization is according to the expected. If it is not, the electrodes' positions shall be checked and corrected.

The user will be then presented some landscape and floral images from the International Affective Picture System (IAPS) for four minutes, in order to neutralize their emotions and record their neurophysiological signals as a baseline for posterior analysis.

The procedures for gathering the neurophysiological responses were planned according with the *WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects*[4].

Color Preference Test

With the apparatus ready and after calibrating the user's emotions, the user will be asked to perform the color preference test.

In this test, 12 will be shown, random and individually. For each color, the following steps will occur:

1. The color is shown for five seconds on full screen, then a Likert scale will appear
2. During the next ten seconds, the user must evaluate the color on the scale (including a countdown timer during the last five seconds), according to their preference.

Between colors, the a black full screen will be showed for five seconds (countdown timer included), in order to neutralize the user's eyes, preparing them for the next color.

Before the test, a brief explanation and some recommendations will be given to the user, in order to ease the further data treatment. For instance, it is important that the user focus their attention on the circle shown in the middle of the screen, trying not to blink or move their eyes during the first five seconds of each color.

Color Test Script

[user's name], neste momento está pronto a iniciar o teste de preferências de cor. Neste teste serão apresentadas 12 cores, individualmente. Primeiro, a cor vai aparecer em *full screen* e depois vai aparecer uma escala idêntica à utilizada no teste de personalidade, onde deverá indicar o quão gosta da cor apresentada. A avaliação de todas as cores é obrigatória e o tempo de resposta é fixo. Para controlar este aspecto, aparecerá um *countdown* do tempo restante para responder.

Enquanto a cor ocupar o ecrã inteiro, peço que se concentre no círculo presente no centro do ecrã, evitando mover os olhos e pestanejar. Peço ainda que tente manter a mesma distância do ecrã ao longo de todo o teste.

Entre cores será apresentada a cor preta por cinco segundos, tempo que deverá servir para se preparar para a cor seguinte.

Quando estiver preparado para iniciar o teste, clique em "Iniciar Teste".

End

In the end of the color preference test, the electrodes will be removed from the user's body and the skin will be cleaned with cotton and ethanol. Then, the test developer will thank the user's participation in the study, giving them a symbolic reward.

[3]This configuration may suffer changes in case the signals are not correctly recorded.

[4]https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects

References

1. Alwis, D. V. & Kon, C. H. A new way to use the Ishihara test. *J. neurology* **239**, 451–454 (1992).

2. Costa, P. T. & MacCrae, R. R. *Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO-FFI): Professional manual* (Psychological Assessment Resources, Incorporated, 1992).

3. Lima, M. & Simões, A. Inventário de personalidade neo revisto (neo-pi-r). *Avaliação psicológica. Instrumentos validados para a população portuguesa* **1**, 15–32 (2003).