UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

Study of the light's dazzling effect on the EEG signal of subjects performing tasks that require concentration

João Pedro Pires de Carvalho dos Santos

Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil em Sinais e Imagens Médicas

Dissertação orientada por:
Prof. Dr. João Miguel Pinto Coelho
Prof. Dr. Hugo Alexandre Teixeira Duarte Ferreira

2019

Abstract

The objective of this work is to study the effect of luminous glare on the electroencephalographic (EEG) signals of subjects that perform concentration-based tasks. The increasing access to high-power and directional light sources (such as laser pointers, but also some flashlights) has led to a growing concern with the potential effects of its use. More than the direct damaging of the retina, the focus has been directed at the effects related to the change in states of concentration on individuals performing tasks whose concentration is critical (such as helicopter pilots or heavy vehicles drivers). This effect is known as ”dazzling” and is typically a temporary deleterious effect on the ability to see or concentrate. However, while damage to the retina can be quantified, glare effects, being indirect (based on the effect on the execution of a given task), are typically qualitative (or at least of more subjective quantification). In this context, the use of brain-computer interfaces capable of analyzing the brain response to external stimuli, opens a door towards the creation of a new tool to evaluate the effects of dazzle. Its potential was evaluated by defining a set of strategies involving the illumination process, EEG signal recording and analysis. A continuous performance task commonly used as an assessment in cognitive neuroscience (N-back) was used to test the attention under the effect of dazzling, in parallel with EEG signals acquisition. Statistical data analysis was performed with the R programming language. ANOVA statistical significant results ($p<0.001$) for answer scores and latency were obtained for differences between the levels of difficulty, both with or without dazzling. Tukey's test further revealed that these statistical differences were on the 0-back/2-back and 1-back/2-back pairs ($p<0.005$). The differences in the pair 0-back/1-back were not significant. Peak band frequency statistical tests were not significant with or without dazzling. Statistical differences were found between dazzling conditions for the frequency band power. For the 0-back and 1-back levels, with the AF7-Fp1 electrode pair, T-student tests resulted in an alpha band frequency power increase ($p<0.003$, in both cases). The electrode pair AF8-Fp2 resulted in an alfa and beta frequency band increase for the 1-back level ($p<0.014$ and $p<0.029$, respectively). These results suggest that concentration is affected by dazzling and can be quantified by means of measuring the change in alpha and beta frequency band power. This technique holds potential and, if further researched and developed, may constitute an effective way of measuring the degree of loss of concentration under the effect of dazzling.

Keywords: dazzling, luminous glare, brain-computer interface, EEG signals.

Resumo

A maioria das pessoas já experienciou o efeito de perda de visão que acontece à noite, e por vezes durante o dia, causados por um automóvel que se dirige na nossa direção, quando os seus faróis estão acesos. Este efeito é conhecido como encandeamento ou *dazzling* e é tipicamente um efeito que afeta negativamente a capacidade de visão ou de concentração. Este efeito acontece em diversas situações, que podem ser inofensivas (e.g., ao tirar uma fotografia com flash) ou potencialmente catastróficas (e.g., queda de avião). O efeito de *dazzling* ocorre quando uma forte fonte de luz surge dentro do campo de visão de um indivíduo e depende principalmente da idade, da condição ocular (nível de visão) e da cor dos olhos, assim como da distância da fonte de luz, a sua luminância e o ângulo entre a fonte de luz e a linha de visão do indivíduo. A crescente facilidade de acesso a fontes de luz de direcionais e de alta potência (tais como apontadores laser e alguns tipos de lanterna) tem levado a uma crescente preocupação com os efeitos potenciais do seu uso. Um dos potenciais efeitos negativos do seu uso é o efeito direto danificador que este tipo de luz tem na retina mas o foco tem sido direcionado para efeitos relacionados com a mudança nos estados de concentração em indivíduos durante a execução de tarefas cuja concentração é crítica (tais como pilotos de helicóptero ou condutores de veículos pesados), porque enquanto que o dano à retina pode ser quantificado, os efeitos de encandeamento, sendo indiretos (baseados no efeito de execução de uma dada tarefa), são tipicamente qualitativos (ou, pelo menos, subjetivamente quantificáveis).

A condução de veículos é uma tarefa que requer concentração constante e está fortemente dependente da visão. O efeito de *dazzling* afeta a visão e pode potencialmente perturbar a concentração, traduzindo-se em acidentes causados pelo efeito. Parte do problema de reduzir, ou até mesmo eliminar o potencial *dazzling* dos condutores de veículos, está na forma atual de medir o *dazzling*, que é geralmente feito de forma subjetiva, através de questionários de desconforto relacionados com métricas de desempenho em tarefas que exijam concentração. Portanto, é necessário quantificar efetivamente o *dazzling* ao nível cerebral para revelar a verdadeira influência do *dazzling* na concentração e trazer atenção para este problema, abrindo caminho para que se possa efetivamente desenhar e criar medidas e/ou equipamento que previnam o perigo que este efeito pode causar tanto aos condutores de veículos, como aos seus passageiros. Neste contexto, a utilização de interfaces cérebro-computador capazes de analisar a resposta cerebral a estímulos externos, abre a porta para a criação de uma nova ferramenta para avaliar os efeitos do encandeamento. O seu potencial foi avaliado através da definição de um conjunto de estratégias envolvendo o processo de iluminação, a captura de sinais EEG e a sua análise.

Medir as consequências dos efeitos de *dazzling* utilizando tarefas de desempenho cognitivo utilizadas em neurociências e sinais EEG constitui uma nova abordagem que pode potencialmente atrair o interesse para a investigação na área de *dazzling*, levando a um futuro mais seguro na condução de veículos. O objetivo deste trabalho é de aferir a capacidade de uma interface cérebro-computador de adquirir dados de EEG em situações de encandeamento luminoso. O trabalho experimental consistiu no controlo de uma fonte luminosa para encadear um indivíduo enquanto realiza uma tarefa que exige concentração

enquanto se adquire sinais EEG.

Para a fonte luminosa foi escolhido um LED de 10 W, controlada por um Arduino UNO, que por sua vez era controlado pelo software Unity a correr num computador. Este software foi utilizado para desenvolver a tarefa a apresentar na experiência. Foi utilizada uma tarefa de medição de performance contínua que é amplamente utilizada em neurociência cognitiva (N-back) para testar a atenção sob o efeito de encandeamento luminoso, em paralelo com a aquisição de sinais EEG. Esta tarefa consiste na exposição de um indivíduo a uma sequência de curtos estímulos (neste caso, foram utilizadas letras) em que se pede a identificação do estímulo que tenha sido repetido com um intervalo n, que permite modelar o nível de dificuldade, sendo que neste trabalho foram utilizados 3 níveis de dificuldade (0-back, 1-back e 2-back). A identificação do estímulo correto é feita ao carregar no teclado, sendo que é contabilizado o número de respostas erradas. Estes 3 níveis foram apresentados de forma aleatória. Cada nível era composto por 48 letras intercaladas por uma cruz de fixação para separar os estímulos entre si. Cada letra foi apresentada durante 500 ms, período durante o qual poderia ocorrer o encandeamento luminoso. O período da subsequente cruz de fixação foi de 2000 ms, dos quais 500 ms foram utilizados como período de base para a análise do sinal EEG. Foram desenvolvidas duas versões do teste N-back, sendo que uma incluía encandeamento luminoso e a outra não, para aferir as diferenças a nível de sinais cerebrais do encandeamento luminoso, na concentração.

A captação de sinais EEG foi feita através de uma placa de medição de sinais eletrofisiológicos de nome BITalino, em que foram utilizados 4 elétrodos em modo bipolar, nas localizações frontais AF7-Fp1 e AF8-Fp2. Estes dados foram analisados utilizando a linguagem de programação Python, utilizando a técnica de Multitaper.

As variáveis eletrofisiológicas analisadas foram a potência e pico de banda e as variáveis da tarefa de concentração foram a pontuação (neste contexto, significa o número de respostas incorretas) e latência de resposta. Para além da análise aos sinais eletrofisiológicos, foi também analisado o número de respostas incorretas e a latência na resposta dos indivíduos à tarefa. Os dados estatísticos resultantes da análise dos dados foram obtidos através da utilização da linguagem de programação R e com os testes ANOVA, Tukey e T de *student*.

Um total de 6 voluntários (2 mulheres e 4 homens) participaram no estudo com uma idade média de 22,5 $\pm$ 0,9 anos de idade. Os resultados do teste ANOVA para a pontuação e latência foram estatisticamente significativos ($p<0{,}001$), para diferenças entre os diferentes níveis de dificuldade, seja com ou sem encandeamento luminoso. Em seguida foi aplicado o teste post-hoc de Tukey, que resultou em diferenças estatisticamente significantes para os pares 0-back/2-back e 1-back/2-back ($p<0{,}005$). As diferenças no par 0-back/1-back não foram estatisticamente significativas. Os testes t realizados para cada nível de dificuldade e condição de encandeamento luminoso não foram estatisticamente significativos, tanto para a pontuação, como para a latência. Os testes ANOVA e T de *student* para as frequências de pico de banda não foram significativos.

Embora não tenham sido obtido resultados estatisticamente significativos para os diferentes níveis dentro de cada condição de encandeamento para a potência de banda, foram obtidos resultados significativos entre as condições. Para a potência da banda foram, os testes T de *Student* foram estatisticamente significativos para os níveis 0-back e 1-back para a banda de frequência alfa, com o par de elétrodos AF7-Fp1, onde se verificou um aumento da potência ($p<0{,}003$, nos dois casos). O par de elétrodos AF8-Fp2 resultou também num aumento para o nível 1-back para a banda de frequência alfa e beta ($p<0{,}014$ e $p<0{,}029$, respetivamente).

A partir destes resultados preliminares, é possível concluir que existe potencial nesta técnica e que, a ser desenvolvida, poderá constituir uma forma eficaz de medir o grau de perda de concentração por

estímulos luminosos.

Palavras-chave: concentração, encandeamento luminoso, interface cérebro-computador, sinais EEG.

Acknowledgments

I would like to thank Professor João Pinto Coelho for supporting over the years and perhaps more importantly, for never giving up on this work, and Professor Hugo Alexandre Ferreira for encouraging me to keep on going with the experiment which was the subject of this thesis when I was lost in doubt as whether I would be able to complete it successfully. In the end, with the help of both professors, not only was I able to complete this thesis, but to publish a paper and a poster which were entered for the International Conference on Applications in Optics and Photonics in Lisbon, Portugal.

I would also like to thank every friend that gave up their time to participate in the experiment, without which this work would not have been possible: António Jorge Silva, Vasco Palma, Joana Lopes, Catarina Santos e Pedro Liberato.

Even though not having participated in the final experiment, the support of my friend Inês Lopes, Inês Gonçalves and Ricardo Silva has been invaluable, both for the moral aspect and the time they spent helping me with the initial stages of the experiment.

Additionally, the author would like to thank the financial support from Fundação para a Ciência e Tecnologia (FCT) under the project UID/BIO/00645/2019 and from CAMELOT Project, ID: 740736, under European grants H2020-EU.3.7.3. and H2020-EU.3.7.7.

Last but not the least, I would like to thank Hugo Plácido da Silva, co-founder of PLUX - Wireless Biosignals which produces the BiTalino board, which he himself offered to fix and tweak the BITalino board that I had acquired to assemble a low-cost EEG research device.

Contents

List of Figures

List of Acronyms

ANOVA Analysis of Variance.

AP Absolute Power.

BCI Brain-Computer Interface.

BDS Backward Digit-Span.

CIE *Comission Internationale de l'Eclairage*.

CSV Comma separated values.

DPSS Discrete Prolate Spheroidal Sequence.

ECG Electrocardiography.

ECoG Electrocorticography.

EDA Electrodermal Activity.

EEG Electroencephalography.

EMG Electromyography.

EOG Electrooculography.

ERD Event-Related Desynchronization.

ERG Electroretinogram.

ERP Event-Related Potencials.

ERS Event-Related Synchronization.

FDS Forward Digit-Span.

FFT Fast Fourier Transform.

FIR Finite Impulse Response.

HID High-Intesity Discharge.

ICA Independent Component Analysis.

IIR Infinite Impulse Response.

ISI Inter-Stimulus Interval.

LFZ Landing-Free Zone.

NTSB National Transportation Safety Board.

RP Relative Power.

TFR Time-Frequency Representation.

WAIS Wechsler Adult Intelligence Scale.

WM Working Memory.

Chapter 1

Introduction

In this chapter, the motivation that led to the development of this thesis is presented, as well as the hypothesis, the goals and a brief account on how to achieve them as well as a general outline.

1.1 Motivation

Most people have experienced the blinding effect that happens at night, and sometimes during the day, caused by an automobile coming from the opposite direction when its high, or even low beam headlights are turned on. This temporary blinding event is known as the dazzling effect. This effect happens in a myriad of situations that can be essentially harmless (e.g.: taking a photograph with a camera flash) or potentially catastrophic (e.g.: airplane crash).

Dazzling occurs when a strong source of light lies within one's field of vision and depends mainly on the age, the subject's eye health condition and iris color as well as the distance from the source of light, its luminance, its spectra and angle between the source of light and the line of sight of the subject [2, 3]. Although the dazzle we encounter in the course of our daily lives is not of appreciable peril, there are specific situations where one's life could be put in jeopardy: operating vehicles.

Operating a vehicle is a task that requires constant concentration and is heavily dependent on vision. The dazzling effect impairs vision and potentially disturbs concentration, which translates in a number of accidents caused by it. The literature is abundant with studies regarding the dazzle effect on car drivers created by High-Intensity Discharge (HID) or Halogen headlights as well as the dazzle effect on airplane pilots created by lasers or sunlight and consequent disruptive effect on vehicle operation capability [4–9]. Dazzling is a real danger and not currently adequately addressed. Part of the problem of reducing, or even eliminating potential dazzling of vehicle operators lies in the current way to access dazzle. The bulk of studies currently available are relying on subjective measures of the dazzle effect [6–8, 10–12]. Subjective measures might not reflect real concentration and vision impairment [13]. Therefore it is necessary to effectively quantify dazzle at the brain level to reveal real objective concentration impairment, bring attention to this problem and paving the way to effectively design and create measures or equipment to prevent the danger that this effect might pose to both vehicle drivers and passengers.

Measuring the degree of one's concentration is not a trivial task, not only because it depends on so many factors (e.g.: fatigue, drug influence, drowsiness, effort, etc.) but also for the difficulty in obtaining an objective measurement. To obtain an objective measurement, it is necessary to define one's maximum concentration capability in order to define a scale for which to correctly classify the individual concentration level. A proven protocol for objective concentration measurement hasn't been achieved yet. However, there have been many successful attempts at measuring the degree of relative concentration

to a baseline level [14–17]. This has been done by the use of cognitive load tasks such as the N-back task, digit span task or complex arithmetic tasks. Therefore, the basis that will enable us to achieve the goal of this dissertation has already been established and it's just a matter of connecting the dots.

Measuring the dazzle effect concentration impairment using general cognitive load tasks with EEG is a completely new approach and potentially establishing the grounds to spark a renewed interest in dazzle reduction research, leading to a future with less annoyance, fewer accidents and fewer deaths. A future where we can dazzle-proof all vehicles.

1.2 Hypothesis

It is known that concentration is affected by the light's dazzle effect in automobile drivers, resulting in an increase in reaction time and a decrease of the car's velocity, as well as airplane pilots being heavily disturbed or even prevented from landing the airplane from laser dazzle or even sunlight dazzle [9–12].

The hypothesis is that the hindering effect of dazzling on concentration must be caused by an underlying brain impairment at play which can be objectively measured and quantified via EEG.

1.3 Goals

The primary goal of this work is to find out if the dazzle effect affects a subject's concentration level by analyzing the brain waves obtained with the EEG.

The secondary objective of this work the establishment of an experimental protocol upon which future researchers could use as a basis to improve or create dazzle quantification EEG-based systems.

1.4 Proposed Solution

In order to achieve the aforementioned goals, an integrated system was created which was composed of the following essential parts: a BITalino board (Plux, Portugal) with two EEG sensors and one electrooculography (EOG) sensor, an Arduino board, a battery-powered 10 W LED light and the Unity software.

This system presented test subjects with concentration tasks with the Unity software while controlling the LED light to generate dazzle stimuli and acquiring EEG measurements with the BITalino board.

The analysis of the brain signals was done with a Python script, using the multitaper analysis approach with the extraction of a number of EEG features. The system creation was possible by combining the current paradigms of concentration EEG measurement and dazzle effect measurement.[1]

1.5 Thesis Outline

The second chapter introduces the background information necessary to understand this work, regarding the inner-workings of the EEG and respective signal analysis, as well as, the concepts of the light's dazzle effect and the current concentration measurement paradigms. The third chapter discusses, analyses and compares the most relevant and recent work found in the scientific literature with the solution proposed in this thesis. The fourth chapter presents a detailed explanation of the experimental setup and the test procedure utilized in this work. In the fifth chapter, the results from the experiment are presented

[1] To be detailed in the next chapters.

and discussed. The sixth chapter enumerates the pitfalls and future work. Finally, the seventh chapter contains a summary of the results of this work.

Chapter 2

Background Theory

We have limited brain resources to process information. Therefore, sensory overload might impair other ongoing cognitive processes, as, for example, driving a car or piloting an airplane [12, 18, 19]. Dazzle is a form of disability glare that occurs when a bright light source is imaged at an extrafoveal location, hindering one's ability to see properly [6]. This visual overload of information impairs concentration.

Concentration levels can be measured by EEG [20], thus it is possible to measure the effect that the dazzle effect has over concentration.

In the following sections, the concepts of electroencephalography, dazzle effect and concentration will be introduced, making references as to how these three concepts can be combined in order to achieve the goals that this work proposed to accomplish.

2.1 Electroencephalography

2.1.1 Introduction

The functional unit of the brain is the neuron. Neurons in the brain have the property of interconnectivity which, together with their characteristic ability to produce electricity (excitability), generate oscillations of electrical activity across the brain, giving rise to synchronous activity [21].

It is assumed that at birth, the full number of neural cells is already developed, roughly 10^{11} neurons. Adults have about 500 trillion synapses, which are connections between neurons. When a neuron is activated, local flows of sodium, potassium, calcium and chloride ions are pumped through channels in neuron membranes in the direction governed by the membrane potential, generating an electrical impulse. Large populations of active neurons can generate electrical impulses that are summed and measurable at the head surface [22].

The electrical activity reflects emotional, cognitive and sensory processes at work in the brain [23], therefore, neurological diseases (e.g.: Parkinson's, Alzheimer's) affecting the structure or function of the brain will affect the normal electrical activity in a specific way which can be measured and be used to support the diagnosis of a particular disease [17, 24, 25].

The brain's electrophysiological signals can be measured with two distinct types of approaches, according to the invasiveness of the recording device:

- **Non-invasive systems**, which record signals via electrodes places on the scalp, also called EEG. Since it is a non-invasive measuring system, it is the safest and swiftest method to employ in both research and clinical use. Since this method places the electrodes on the scalp, it's natural that this

is the method that presents the poorest signal quality since it deals with many barriers between the signal source and the electrode such as the scalp, the skull and cerebrospinal fluid. However, this method is easy to apply and can be repeatedly applied to human subjects without virtually no risk or limitation [22]. This is the method that will be used in this dissertation.

- **Invasive systems**, which record signals with electrodes placed surgically in the brain. The electrodes can be placed on the surface of the cortex, thus this method's called electrocorticography (ECoG), or deeper inside the brain, being called intracortical systems. The proximity of ECoG and intracortical electrodes to the cortical sources of EEG activity enhances their spatial resolution, as well as their sensitivity and signal-to-noise ratio, particularly for high-frequency EEG activity, as well as decreased artifacts and provides an increased performance of detection and monitoring of patients with severe cases of epilepsy, necessary for their surgical management and other neurological injuries. It's important to note that this method has several disadvantages, such as invasiveness and small coverage of brain area [26, 27].

The EEG measurement system consists of roughly 2 fundamental parts: amplifiers and electrodes.

The amplitude of electric potentials that can be measured on the scalp is on the order of the microvolt, about 100 times lower than electrocardiography (ECG) signals. Therefore, it is necessary for an effective EEG measuring system to have an amplifier with high gain, typically ranging from 100 to 100,000. However, there are other parameters to take into account, such as the signal-to-noise ratio, the range of the analog-digital converter (ADC) and the sampling rate. Each of the differential amplifiers has two inputs, one of them is connected to a measuring electrode and the other is connected to a common system reference electrode, usually placed on one of the subject's earlobes, because these sites have a sufficiently low relative electrical activity to be considered ground. Modern EEG systems are digital, therefore the amplified signal must be digitized via an ADC after being passed by an anti-aliasing filter to prevent distortion of the signal by smoothing it. ADC sampling rate must be at least double of the highest frequency component of our interest. For most applications, 256 Hz is sufficient.

The electrodes are the part of the EEG system that is in contact with the subject and the first link in signal acquisition. High fidelity signals from electrodes are of paramount importance for medical applications such as EEG, electromyography (EMG) and ECG. The most commonly used electrode is the silver/silver chloride (Ag/AgCl) type which can be found in reusable form but is most often used as a disposable electrode. These electrodes are referred to as wet electrodes, due to their use of an electrolytic gel to form a conductive path between the skin and the electrode. There are, however, other types of electrodes. Dry electrodes consist of a biocompatible metal (such as stainless steel) used with no electrolyte between electrode and skin and capacitive electrodes consist of a metal or semiconductor with a thin dielectric surface layer so that the bioelectric signal is capacitively coupled from the skin to the substrate. Searle *et al*[28] reported that the dominance of wet electrodes in clinical settings will most likely remain unchallenged by dry and insulating types because wet electrodes are simple and lightweight, are relatively cheap, available and have lower initial contact impedances and resistance to movement artifacts.

In conventional scalp EEG, it is a normal preparation procedure to place the electrodes on the scalp with a conductive gel or paste (assuming the common wet electrode setting), usually after removing hair and applying light abrasion to the target scalp area in order to reduce the surface epidermal layer, which has higher impedance than the underlying tissue. The practice of skin abrasion has, however, been shown by Ferree *et al* [29] to increase the infection risk by blood-borne pathogens and that EEG recordings performed without this practice, even though having increased scalp impedance, nevertheless

yielded high-quality EEG signals, suggesting that this practice can be abandoned.

As early as the first International EEG congress, in 1947, it was recognized that a standard method of placement of electrodes used in EEG was needed in order to have consistency and reproducibility amongst studies using EEG. This need resulted in the advent of the 10-20 electrode system (Figure 2.1). The ”10” and ”20” refer to the fact that the actual distances between adjacent electrodes are either 10% or 20% of the total front-back or right-left distance of the skull. Hitherto the 10-20 electrode system has become the *de facto* standard for EEG studies, in both clinical and non-clinical settings. This system describes the location of scalp electrodes placement for EEG and is based on standard landmarks of the skull, resulting in the relative positioning of the electrodes being essentially the same for every person. The coordinates for the electrode placement are composed of letters and numbers. The letter codes Fp, F, C, P, and O represent the frontopolar, frontal, central, parietal and occipital areas, respectively. The letter system was given after the anatomical terms for the cortical areas covered with the exception of the ”C” electrodes which were termed central since they were located over the central sulcus. The ”z” is used for vertex electrodes, which are placed on the midline. The numbers differentiate between left and right homologous regions, odd numbers for the left hemisphere and even numbers for the right hemisphere. The letter codes ”Pg”, ”Cb” and ”A” correspond to additional electrodes in pharyngeal, cerebellar and auricular sites, respectively. In Figure 2.2, an illustration of the electrode positioning seen from the superior cross-section (left) and lateral (right) views can be seen.

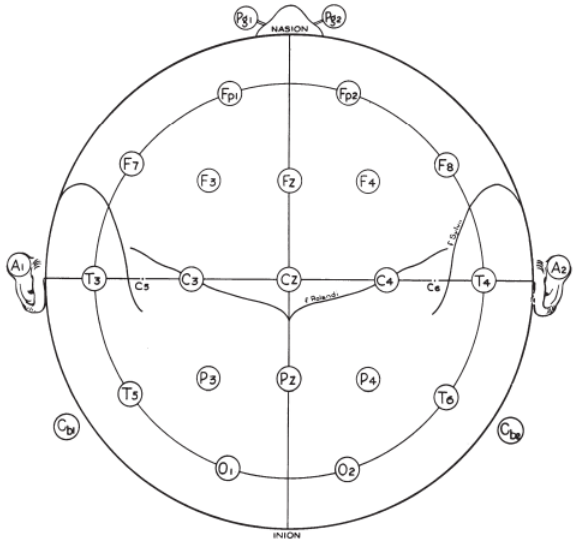

Figure 2.1: The 10-20 international standart method of electrode placement for EEG, adapted from [30].

The 10-20 system only allowed for positioning of 21 electrodes, which was more than the technology at the time could support. In modern times, scientific knowledge and technology advanced immensely and it is now common to use 128 or even 256 EEG channel systems, renewing the need to create new international standart systems to accommodate the new capabilities of modern systems. These high-density electrode systems often include an EEG cap with the electrodes already placed on specific locations for ease of use [30, 31].

Early forms of EEG were cumbersome and expensive. Nowadays technology has advanced in such a way that there are currently companies creating low-cost, portable commercial systems that record, filter, sample, digitize and store EEG data directly on the device itself or stream it wirelessly to the nearest computer or smartphone.

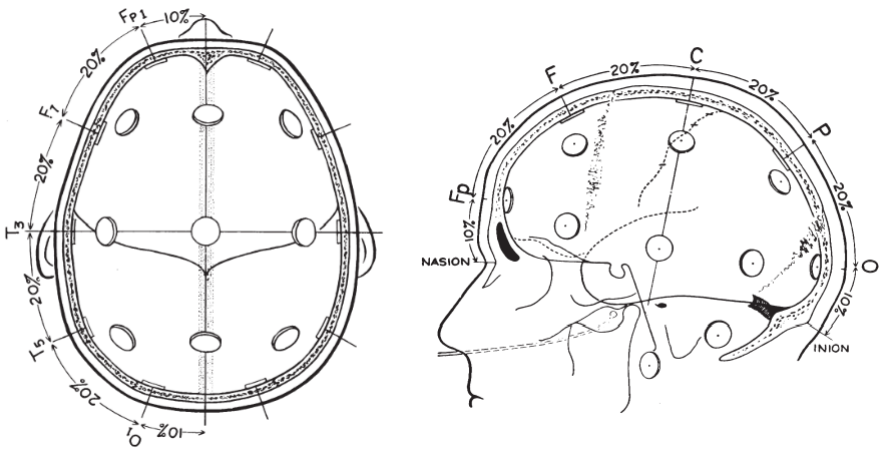

Figure 2.2: The 10-20 international standart method of electrode placement scheme for EEG, as seen from the superior cross-sectional view (left) and the lateral view (right). Adapted from [30]).

One example of these portable commercial systems is the BITalino©board, shown in Figure 2.3. BITalino is a low-cost hardware and software toolkit that has been specifically designed to deal with the requirements of body signals and can be connected to a multitude of biomedical sensors, and can be used as a brain-computer interface (BCI). A BCI, also sometimes referred to as a brain-machine interface (BMI), is a communication and/or control system that allows real-time interaction between the human brain and external devices [32]. In the case of the BITalino board, it allows the acquisition of EEG, ECG, EMG, EOG and electrodermal activity (EDA) signals. BITalino's sensors can connect either to dry or wet electrodes, however, for the purposes of this work, gold cup electrodes were used. These electrodes use dry electrode technology but it is recommended that a Ten20©conductive paste is used, to adhere the electrodes to the scalp and to reduce electric impedance. The board can have a number of electrodes, depending on the number and type of sensors connected, with one common reference. The BITalino board has 6 channels with a sampling rate of 1000 samples per second per channel, with a 10-bit resolution on 4 channels and 6-bit resolution on 2 channels. Additionally, it can be connected to a 2.0 Bluetooth transmitter for data connectivity.

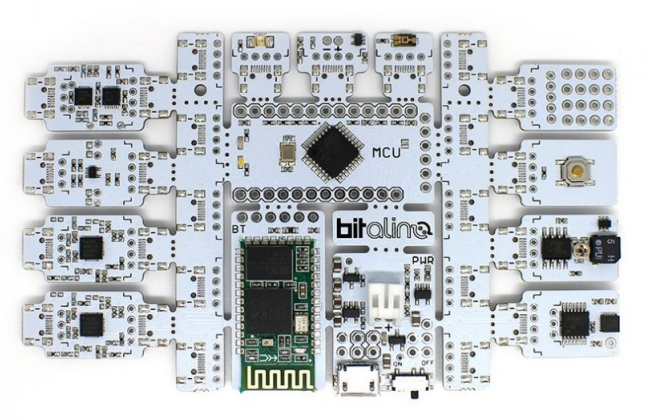

Figure 2.3: The BITalino©board.

Brain rhythmic activity contains multiple frequencies simultaneously, which can be separated through signal-processing techniques. Different cognitive processes and neural functions seem to utilize

different frequency ranges or conjunctions of frequency ranges. Brain rhythms are grouped into bands that are defined by logarithmically increasing center frequencies and frequency widths. Changes in rhythmic activity correlate with task demands, including perceptual, cognitive, motor, linguistic, social, emotional, mnemonic, and other functional processes. The main brain rhythm frequency bands are called [22, 33, 34]:

- Delta (1-4 Hz): Electromagnetic waves ranging from 100 to 200 μV in amplitude. In a conscious state, most adults exhibit almost no delta activity; instead, this activity occurs when in a deep sleep, unconscious, anesthetized or lacking oxygen.
- Theta (4-8 Hz): Electromagnetic waves with less than 30 μV in amplitude. This activity primarily occurs in the parietal and temporal regions of the brain. Such waves are produced when people experience emotional pressure, interruptions of consciousness, or deep physical relaxation. They are also linked to memory encoding processes [23].
- Alpha (8-13 Hz): Electromagnetic waves ranging from 30 to 50 μV in amplitude. This type of periodic wave is produced in the parietal and occipital regions of the brain when in a state of consciousness, quiet, or at rest. When thinking, blinking, or otherwise stimulated, alpha waves disappear. This is known as an alpha block.
- Beta (13-30 Hz): Electromagnetic waves ranging from 5 to 20 μV in amplitude. This type of activity occurs in the frontal region when people are conscious and alert. These waves are particularly apparent when a person is thinking or receiving sensory stimulation.
- Gamma (30-150 Hz): Electromagnetic waves ranging from 5 to 10 μV in amplitude. Studies have found that lower gamma activity (30 to 50 Hz) is related to selective attention and highlighted that this activity is related to cognition and perceptual activity.

These are not the only frequency bands; there are oscillations from the sub-delta up to omega (up to 600 Hz) ranges, but the frequency bands that are most typically associated with cognitive processes in the literature are between 1 Hz and 150 Hz. This grouping is not arbitrary but, rather, results from neurobiological mechanisms of brain oscillations, including synaptic decay and signal transmission dynamics. However, there are no precise boundaries defining the bands, and there are small differences between several books and studies in this regard. Furthermore, individual differences in peak frequencies have been linked to a number of individual characteristics, including brain structure, age, working memory capacity, and brain chemistry, rendering the fixation of hard limits for these frequency bands, utterly senseless [23, 34, 35].

2.1.2 Event-Related Potentials

Event-Related potentials (ERPs) are voltage changes induced within the brain in response to a variety of sensory, cognitive and motor processes. The ERP consists of a sequence of positive and negative voltage fluctuations that are labeled components.

Although the amplitudes of ERP signals are small (1-30 μV) and are embedded in EEG activity unrelated to the eliciting event, they can be recorded non-invasively from the scalp by means of signal averaging techniques in which a greater number of trials will have a greater reduction in the non-stimulus EEG activity. Good signal-to-noise ratios can be generated on a range of trials from 10 to 150, depending on the amplitude of the wave.

The resulting averaged ERP waveforms consist of a sequence of positive and negative voltage deflections, which are called peaks, waves, or components. P and N are traditionally used to indicate positive and negative peaks, respectively, and the number indicates a peak's position within the waveform (e.g.: P1 is the first major positive peak). Alternatively, the number may indicate the latency of the peak in milliseconds (e.g., N170 for a negative peak at 170 ms, see figure 2.4). If the number is greater than 5, it should be assumed that it is referring to the peak's latency. Components may also be given paradigm or function-based names, such as the error-related negativity (which is observed when the subject makes an error) or the no-go N2 (which is observed on no-go trials in go/no-go experiments, which are a type of test that is only successfully passed once two conditions are satisfied) [36].

There are three measurable aspects of the ERP waveform: amplitude, latency, and scalp distribution. While component amplitude provides an index of the extent of neural activation (i.e.: how the component responds functionally to experimental variables), component latency (i.e.: the point in time at which the peak occurs) reveals the timing of this activation. Finally, a component's scalp distribution (i.e.: the pattern of voltage gradient over the scalp at any point in time) provides information on the overall pattern of activated brain areas [36, 37].

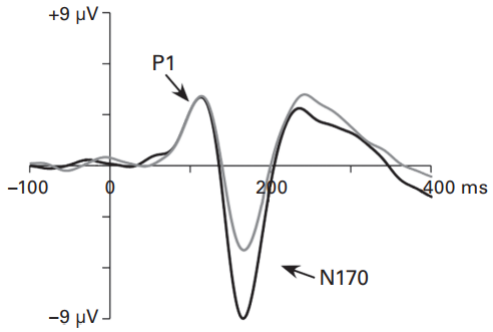

Figure 2.4: Event-Related potential of 2 visual stimuli, with the first peak of positive deflection (P1) and the negative peak at 170 ms from the stimulus onset (N170) highlighted. It is worth noting that the ERP convention is to have an inverted Y-axis, meaning that positive deflections will be pointed downwards and negative deflections will be pointed upwards. Adapted from [36]).

2.1.3 Event-Related synchronization/desynchronization

Event-Related oscillatory EEG responses can be quantified e.g., by means of the event-related desynchronization method. In this method, a relative decrease in the power of a certain frequency band during stimulus processing (as compared to a no-stimulus reference) is called event-related desynchronization (ERD), whereas the opposite, a relative increase in the power is called event-related synchronization (ERS). Synchronization occurs when the neural populations display resonant behavior (i.e., oscillate at the same frequency), conversely, desynchronization signals a loss of resonant behavior in a given frequency band. The ERD/ERS values are within-subject measures of relative changes in the EEG. As is the case with the EEG, also ERD/ERS technique is characterized by a relatively good temporal resolution and provides a suitable method to assess dynamic brain oscillatory responses during cognitive processing and memory load conditions.

The classical method, termed "power method", to compute the time course of ERD can be summarized in the following steps:

1. Bandpass filtering of all event-related trials;

2. Squaring of the amplitude samples to obtain power samples;

3. Averaging of power samples across all trials;

4. Averaging over time samples to smooth the data and reduce the variability.

To obtain percentage values for ERD/ERS, the power within the frequency band of interest in the period after the event is given by A whereas that of the preceding baseline or reference period is given by R. ERD or ERS is defined as the percentage of power decrease or increase, respectively, according to the expression:

$$ERD\% = \frac{A - R}{R} \times 100 \tag{2.1}$$

It should be noted that the power method is not recommended in the analysis of the theta and lower alpha frequency, due to possible contamination of ERP signals [17, 24, 38].

2.1.4 EEG signal analysis

The EEG signal can be interpreted as resulting from a time-series that consists of multiple sinusoidal waves added together forming a complicated waveform. Sinusoidal waves have three basic properties: amplitude, frequency, and phase. These three properties can provide information regarding brain rhythms. Therefore, the foundation of EEG signal analysis is uncovering the sine waves that compose the signal.

Spectral analysis methods allow for the decomposition of the signal into its frequency (sinusoidal) components. Traditionally, the Fast Fourier Transform (FFT) is used to perform time-frequency decomposition of EEG signals. However, studies have shown that this method is not the most suitable method for EEG signal analysis, because the EEG signal is non-stationary, that is, the characteristics of the sinusoidal waves that compose the signal change over time due to task events and due to endogenous processes. The pure FFT analysis of EEG signals leads to less defined spectral peaks, as can be seen in Figure 2.5. The nonstationary time series have a more complicated structure and therefore requires energy at a larger number of frequencies in order to represent the time series in the frequency domain. Most importantly, the Fourier transform does not show how dynamics change over time, that is, although the Fourier transform contains all the temporal information of the time series data, time-varying changes in the frequency structure cannot be observed directly in power or phase plots [34, 39].

There are several modern spectral analysis techniques which perform time-frequency decomposition of signals and are suitable for non-stationary signals, such as the Wavelet Transform, Bandpass Filtering, Hilbert transform, Short-time Fast Fourier Transform, and Multitapers. The technique that will be used in this work is the latter.

The problem with the pure FFT approach is that there is no temporal weighting, therefore, frequency changes cannot be localized in time. The temporal weighting is related to the duration of the stationarity assumption. The pure FFT assumes that the signal is stationary for the duration of the whole signal, whereas the previously mentioned techniques assume stationarity of the signal only during the time period in which the signal looks like a sine wave. EEG data often remains stationary for hundreds of milliseconds, therefore, the assumption of stationarity is likely to be valid [40].

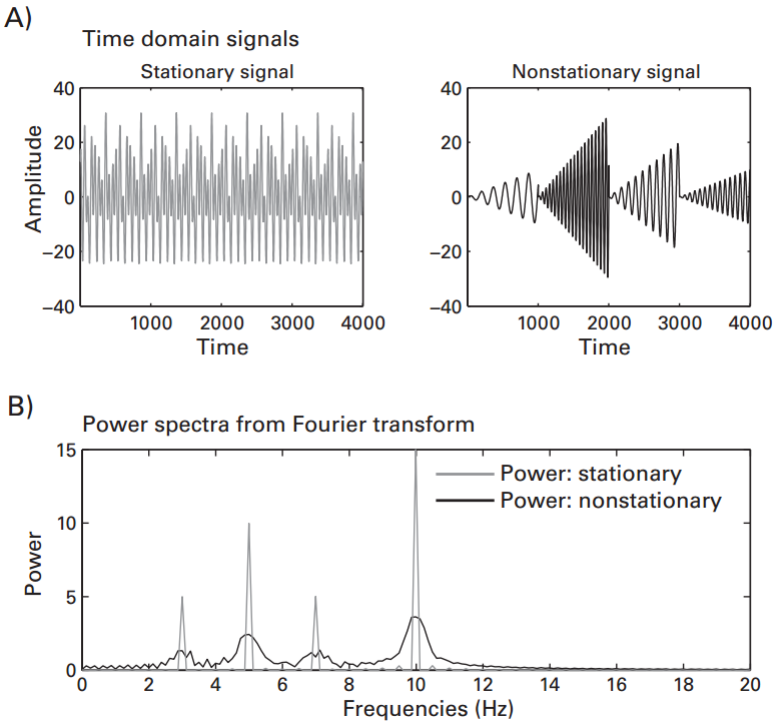

Figure 2.5: Graphical representation of a simulated stationary signal (top left), a non-stationary signal (top right) and the comparison between their FFT's (down). Adapted from [34].

A stationary periodic signal is one in which the frequency structure does not change over time. In stationary signals, it is appropriate to estimate a power spectrum, which represents the strength (spectral power) of the signal at different frequencies. When the oscillatory structure of a signal changes over time, the signal is said to be time-varying or non-stationary. If the intention is to characterize how EEG oscillatory dynamics change in time, estimating a single power spectrum for the entire signal is not appropriate. Instead, a spectrogram may be estimated for time-varying signals. The spectrogram describes the power in the signal as a function of both frequency and time. However, the prevailing techniques for EEG spectral estimation produce noisy and inaccurate estimates of the power spectrum, making it difficult to interpret the resulting spectrogram. The multitaper spectral estimation method has proven to greatly enhance the clarity of the spectral estimates over standard methods.

The simplest and most common method used for performing spectral estimation on EEG data is called the periodogram. In practice, the computation of the periodogram involves the calculation of the FFT. However, the periodogram has several suboptimal properties that cause major problems in spectral estimation. A simple periodogram is obtained by applying the mathematical operation of convolution of a rectangular window with any finite data series. Note that as displayed in Figure 2.6 a), an infinite data series is multiplied by the rectangular window (in practice, since the multiplication limits the resulting signal to the width of the rectangular window, any signal having a length greater than the length of the rectangular window can be seen as infinite even though it can have any finite length), which results in a periodogram with high spectral leakage, introducing frequency aberrations, which is not ideal in the analysis of biosignals, in which the distinction of frequency peaks and their proximity to the ideal spectrum is paramount.

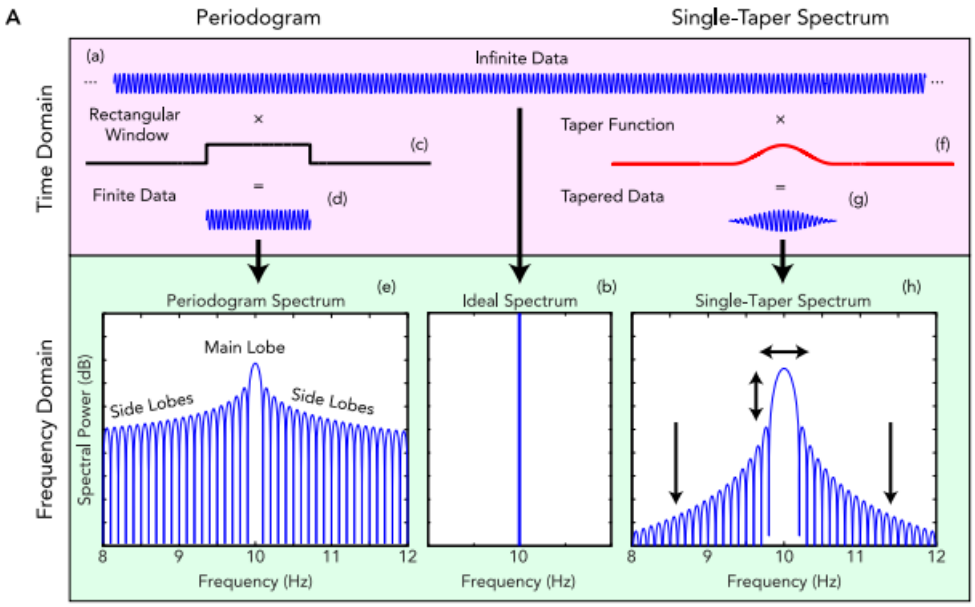

Figure 2.6: Graphical representation of an infinite sine wave data series (a), a rectangular window (c), a taper function (f), the periodogram spectrum resulting from the convolution of the data series with the rectangular window (e) and the power spectrum resulting from the convolution of the data series with a single taper function (h). Adapted from [41].

Given the strong bias, the periodogram is an exceptionally inaccurate spectral estimation method. However, there are ways to greatly reduce periodogram bias. The most common method for improving periodogram bias is to apply a function called a taper or window function to the finite data before performing spectral estimation. The result of this process can be called a single taper spectrum.

The main and side lobes of the periodogram for finite data arise due to the sharp changes in the rectangular window at the start and end of the data. To reduce the bias caused by the side lobes, the abruptness of these transitions must be smoothed out. By taking the product of the raw data and a taper function that has gradual transitions between 0 and 1, the processed data, as can be seen in Figure 2.6 f), becomes tapered at both ends, decreasing the magnitude of the discontinuities. As a result, the power in the side lobes of the single taper spectrum is greatly reduced compared with the periodogram, and the difference in power between the main lobe and the highest of the side lobes is increased. This means that there will be less power from surrounding frequencies leaking into the spectral estimate, and thus the estimate will be less biased.

However, the periodogram produces estimates with high variance across all frequencies. This is illustrated in Figure 2.6 b), where the ideal spectrum has a single vertical peak, whereas the periodogram has visible and highly variable side lobes around each peak. Thus the periodogram has a much higher variance than the ideal spectrogram. As the length of the data set becomes larger, the variance of the periodogram remains constant. Thus, no matter how much data is collected, the periodogram estimate will not improve. Consequently, we call the periodogram an inconsistent estimator of the spectrum. The variance of the spectral estimate is an especially important concern in EEG data analysis, requiring a high temporal resolution, which requires the use of relatively short data windows. Furthermore, while tapering reduces bias, it can actually increase the variance in short data sets. This is because tapering the data forces the time points near the ends to converge to zero, effectively reducing the amount of data available to make the estimate and increasing the variance. Thus the analyses of short segments of noisy

EEG data are prone to high variance, particularly when we use tapering to reduce the bias [41].

Multitapers

A technique called multitaper spectral estimation has been shown to have superior statistical properties compared with single-taper spectral estimates. The multitaper method works by averaging together multiple independent spectra estimated from a single segment of data. Instead of using a single-taper function to compute the spectrum, it uses multiple taper functions to compute single-taper spectra, which are averaged together. These tapers come from a particular class of functions called the discrete prolate spheroidal sequence (DPSS). DPSS tapers are special because they are not only optimized to reduce bias, but they also have a special mathematical property called orthogonality, which enables them to extract uncorrelated single-taper spectral estimates from the same data. Because these single-taper estimates are uncorrelated with each other, they can be averaged together as if they were independent trials of the same condition, producing a spectrum with reduced variance. Another useful property of DPSS tapers is that they make it particularly easy to define the spectral frequency resolution and smoothness of the resulting spectrum. Estimating the multitaper spectrum is therefore no more complicated than taking the average of several single-taper spectra. Figure 2.7 shows a schematic diagram for multitaper spectral estimation.

In practice, multitaper spectral estimates are defined by several parameters, which control the number of DPSS tapers and their properties. These parameters are N, the size of the data segment in seconds; TW, the time-half-bandwidth product; and L, the number of tapers. By understanding these three parameters, it is possible to explicitly control features underlying the multitaper estimate. To choose the values for these parameters, we must only know two things: the time period over which the data are thought to be stationary and the desired spectral resolution. First, the data segment size N should be defined as the maximum length of time (in seconds) at which the data are thought to be stationary. Practically, N should also reflect the scale of the dynamics that can be adequately observed at the time scale of the visualization. Next, the spectral resolution (Δf) is the bandwidth (in Hz) of the main lobe in the spectral estimate, which controls the minimum distance between peaks that we can resolve. In practice, a large Δf will produce smooth, low-resolution peaks, whereas a small Δf will produce higher resolution peaks with greater detail. Typically, it is better to err on the side of a smaller Δf, so that potentially important spectral features are not overlooked. Given these Δf and N, we can compute TW, the time-half-bandwidth product, as

$$TW = \frac{N\Delta f}{2} \tag{2.2}$$

where TW is greater or equal than 1. The time-half-bandwidth product (dimensionless) is a parameter used in computing the DPSS tapers that relate the frequency resolution to the data window size and is simply the product of the window duration (N) and half the bandwidth of the main lobe (Δf/2). Finally, we must determine L, the number of tapers used in the estimate.

$$L = [2TW] - 1 \tag{2.3}$$

The number of tapers is important since it turns out that the variance of the spectral estimate is improved by a factor of L compared with a single-taper estimate. However, increasing the number of tapers means that the time-half-bandwidth product must be larger, which implies that the spectral resolution of the estimate will be reduced if N is fixed. It is thus important to account for the underlying assumptions about data stationarity and spectral resolution when adjusting L to minimize the variance.

There are available software libraries such as Python's neuroscience library MNE, which implement multitaper power spectra calculation automatically, with the minimum input being the data itself and the

data's sampling frequency.

The low variance of the multitaper spectral estimate greatly facilitates the identification of sensitive features of neural activity in EEG [41].

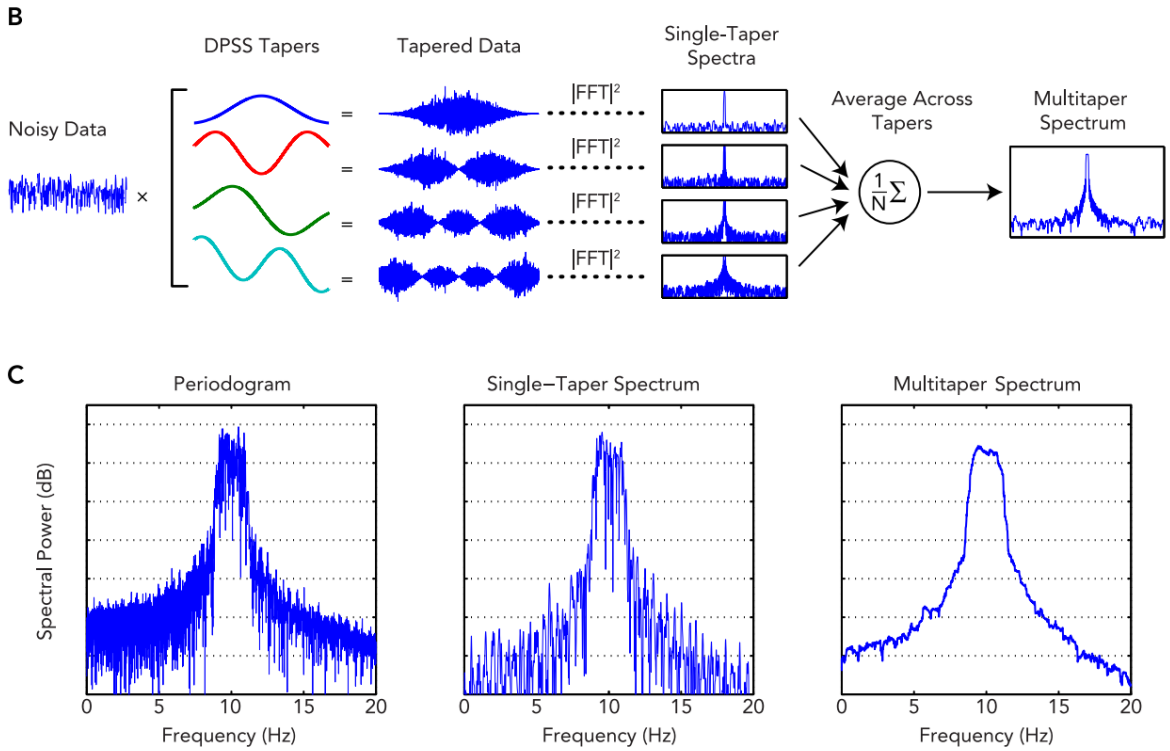

Figure 2.7: Graphical representation of the process of applying convolution of multiple DPSS Taper functions to noisy data and the following power spectrum result b) and the comparison between a periodogram, single-taper and multitaper power spectra c). Adapted from [41].

EEG Artifacts

The EEG measuring system detects and greatly amplifies small electrical signals, so it is bound to capture unwanted signals. EEG artifacts can essentially be categorized as internal and external to an individual.

These artifacts can be attenuated or altogether eliminated by the use of filters or by the setting of epoch-removal criteria, respectively [34]. Signals can be contaminated by external or physiological (internal) artifacts. The external artifacts are generated from the recording system and can be significantly reduced by a careful design of the system and by following appropriate signal recording procedures. The greatest sources of internal artifacts are EMG, ECG, electrooculography (EOG) signals and sweating. Muscle activity is controlled by motor neurons. These neurons elicit muscle contraction via electrical signals, which can have amplitudes of up to 10 mV, depending on the distance to the muscle locus [42], thus being of greater magnitude than the EEG signal. The recognition, identification, and elimination of artifacts is an important process to minimize the chance of misinterpretation of EEG signals.

Internal artifacts can be summarized as follows:

- Artifacts from the eyes and eyelids - The movement of the eyeballs causes a change of potential in the electrodes near the eyes at Fp1-Fp2 (frontoparietal). The fluttering of the eyelids appears as a signal with a frequency between 3 and 10 Hz.

- Eye movement artifacts - The ERG or electroretinogram is a signal regarding the potential difference between the retina and the cornea of the eye which changes with incident light, caused by the contraction of the iris sphincter muscle which affects the size of the pupil, causing artifacts in the EEG signals. Voltage amplitude is proportional to the angle of gawk.

- Eye blink - Eye blinks produce high amplitude signals that can be many times greater than the amplitude of EEG signals of interest. Repetitive blinks produce slow waves, with a frequency that overlap with the delta frequency range

- EMG artifacts - Produced by muscles in the body, the amplitude of these artifacts depends on the degree of muscle contraction and can be caused, for example, by tongue movement, swallowing, grimacing or chewing. The frequency range overlaps with the beta band and these outputs most commonly appear in the frontal and temporal electrodes.

- ECG artifacts - Electrical activity produced by the heart can be detected in the scalp producing ECG artifacts. These artifacts are more common in short neck individuals. Most of the cardiac artifact frequencies are near 1 Hz and the amplitude can reach several millivolts.

External artifacts can be categorized as follows:

- Transmission-line artifacts - AC devices in the recording environment introduce 50 Hz (or 60 Hz in the Americas and parts of Asia) noise into the data, also called powerline interference. This occurs because electric and magnetic fields incident on the electrode leads and body generate potentials which add linearly to the signal [29]. As the bandwidth of most of the EEG signal acquisition ranges from 0.5 to 60 Hz and the frequency of transmission lines is 50 Hz, the signal easily affects most EEG signal acquisitions. This artifact affects all channels with poor impedance matching. This artifact can be easily removed by using a notch filter of 50 Hz.

- Mobile phone artifacts - Mobile phones generate electromagnetic signals. This manifests as a high-frequency signal captured by EEG, therefore, phones should not be used or be around when performing EEG acquisition experiments.

- Electrode artifacts - Poor electrode contact gives rise to low-frequency artifacts. They are usually limited to one electrode and are synchronized with respiration due to the motion of the electrode. Electrode movement is another form of electrode artifacts that is very different from the EEG activity of any form. Perspiration can also generate electrode artifacts with low amplitude with waves that typically have durations greater than 2 seconds, and so they are beyond the frequency range of brain generated EEG signals.

- Physical movement artifacts - These artifacts appear because of a loss of contact of the electrode due to an abrupt physical movement of the subjects. Its morphology is very different from EEG signals.

Most of the artifacts can be prevented while recording by making a good recording protocol, which includes giving instructions to the subject to avoid physical and eye movements and not allowing mobile phones in the room. Transmission line artifacts can be removed by notch filters. Information regarding memory and concentration can be found in high-frequency neuronal activity, which overlap with the spectral bandwidth of muscle activity which typically ranges from 20 to 300 Hz, leading to EMG artifact contamination. Muscle electrical activity differs between muscles, direction, and force of contraction

as well as the participant's sex, but it is reported that, for instance, frontal muscles elicit frequencies around 20 - 30 Hz and 40 - 80 Hz for temporal muscles. The extraocular muscles that control saccadic eye movements produce activity that peaks around 65 Hz. Contamination from EMG artifacts starts at 20 Hz, reaching high intensities at 40 Hz and even higher at 80 Hz. However, for this work, we are focusing on the neuronal activity below 30 Hz, and most of EMG artifacts can be removed by using low pass filters at around 30 Hz, which are traditionally used for this purpose. It is worth noting that the levels of EMG contamination increase when subjects are asked to perform tasks that induce cognitive load. Finally, EOG artifacts can be removed with a high-pass filter with a cutoff frequency of 3 to 4 Hz [43, 44].

The previously mentioned information regarding EEG artifacts is summarized in Table 2.1.

Table 2.1: Table summarizing the main types of EEG artifacts.

Artifact Type	**Source**	**Frequency Range (Hz)**	**Amplitude (mV)**
ECG	Heart	>1	1 - 10
Transmission Line Noise	Transmission Line	50 - 60	Low
EMG	Body Muscle	20 - 300	Low
EOG	Eye	0.5 - 3	100
Phone	Mobile and landline phones	High	High
Eletrode	Eletrode and sweating	Very low	High
Physical Movement	Physical Movement	Very low	Very High

Baseline Normalization

The frequency spectrum of EEG data tends to show decreasing power at increasing frequencies. This is not specific to EEG data but also characterizes the relationship between power and frequency of many signals. This decrease in power as a function of an increase in frequency follows the power law[1] (which is a function of the form c/f^x, where c is a constant and x is an exponent). Because EEG time-frequency power follows this law, the power at higher frequencies (e.g., gamma) has a much smaller amplitude than the power at lower frequencies (e.g., delta). This fact presents five important limitations to interpreting and working with time-frequency power data:

- It is difficult to visualize power across a large range of frequency bands.
- It is difficult to make quantitative comparisons of power across frequency bands.
- Aggregating effects across subjects can be difficult with raw power values. This is because individual differences in raw power are influenced by skull thickness, sulcal anatomy, cortical surface area recruited, recording environment and other factors that are independent of the neurocognitive process under investigation.
- Task-related changes in power can be difficult to disentangle from background activity. This is particularly the case for frequencies that generally tend to have higher power or frequencies that tend to have higher power during baseline periods.

[1]Note that here the power law refers to raising a number to a power of another number, not the squared magnitude of a complex number.

- Raw power values are not normally distributed because they cannot be negative and they are strongly positively skewed. This limits the ability to apply parametric statistical analyses to time-frequency power data.

To solve all these limitations, a baseline normalization should be performed. There are several techniques of baseline normalization, such as decibel conversion, percentage conversion, and the z-transform. However, the decibel conversion technique was chosen for this work because it is robust to many limitations of the power law and subject-specific and electrode-specific idiosyncratic characteristics.

The decibel (dB) is a ratio between the strength of one signal (frequency-band-specific power) and the strength of another signal (a baseline level of power in that same frequency band). The base unit is called a bel and is the logarithm of a ratio of numbers. Typically, tens of bels are used, hence decibels. The conversion is defined by:

$$dB_{tf} = 10 \log_{10} \left(\frac{activity_{tf}}{\bar{baseline}_f} \right) \tag{2.4}$$

Where the horizontal bar over *baseline* indicates the mean across the baseline time period, and t and f are time and frequency points (note that the baseline has no t subscript, indicating that all time points within a frequency band use the baseline period).

The decibel conversion enables the scale and the interpretation of frequency-band-specific power to become the change in power relative to baseline rather than the absolute level of power. Because both activity and the baseline within each frequency band are equally affected by the power law scaling, any frequency-band-specific activity or amplitude level that is constant over time, including background activity, will be removed. The baseline is a period of time, typically a few hundred milliseconds before the start of the trial, when little or no task-related processing is expected. Alternatively, the baseline period can be chosen as a separate time interval from the trials (e.g., a period before the tasks where the subjects are told to relax) [34].

2.2 The Dazzling Effect

Across the literature, the very concept of glare is not well defined, at least not in a generally accepted way. For the purposes of this thesis, the terms and concepts of the majority of the most recent literature about this subject were be adopted [45].

Glare is the sensation produced by luminance within the visual field that is sufficiently greater than the luminance to which the eyes are adapted to cause annoyance, discomfort or loss in visual performance and visibility, caused by the interaction of a bright source of light with the eye media and can be essentially be dichotomized into discomfort glare and disability glare.

Discomfort glare does not impair vision (i.e., even though causing the contraction of the pupil, it does not significantly reduce the ability to perceive information), but it can be startling and distracting, cause blinking, squinting, ocular aversion and fatigue. It has always been associated with a strong flinch reflex in the extra-ocular (facial) muscles surrounding the eye. The physiological and psychophysical origins of discomfort glare remain uncertain. As the vision remains unaltered, the phenomenon is subjective. It is, therefore, most often measured by means of subjective rating scales, in which participants, while looking in the direction of a light source, indicate how annoying it is. Although there is no consensus about which rating scale should be used, the nine-point DeBoer scale is most widely used in the field of automotive and public lighting scale. This scale includes the following ratings: unbearable (1), disturbing (3), just

admissible (5), satisfactory (7) and unnoticeable (9). This type of discomfort varies at different times and different individuals, being worse on older people [1, 6, 10, 46].

Disability glare does indeed impair visual performance and if often referred to as veiling glare or dazzling glare. It negatively affects visual detection and resolution which, in terms of detection, represents the case where the contrast of a particular object has been reduced to below the eye's threshold, essentially masking it and rendering it no longer visible. Disability glare, which will be referred to as the dazzling effect, invokes a reaction of avoidance in order to protect against possible overexposure, which might lead to either temporary blinding or even permanent retinal damage. However, the retina does not possess the required pain sensors to prevent any such situations, being the rapidly contracting iris, which drives the pupil to over-constriction and pupillary spasm, that originates the discomfort leading to the avoidance reaction. The physical mechanism that originates the dazzle disability glare is light scattering [1, 3, 47, 48].

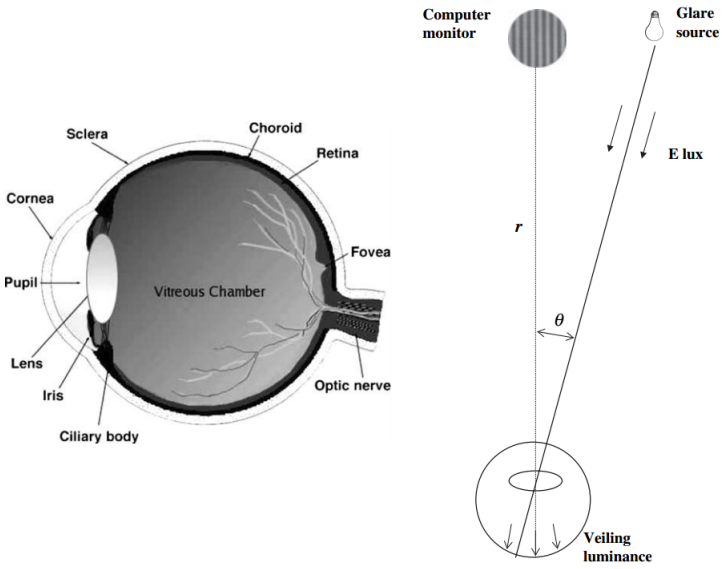

Figure 2.8: Diagram of the structure of the eye (left, adapted from [49]) and veiling luminance caused by a glare source (right, adapted from [2]).

If the human eyes (Fig. 2.8) were perfect optical devices, light from a bright light source would perfectly propagate from the cornea through the retina, and one would only experience a small saturated spot in his vision that would only obscure the appearance of the source of light itself, thus, vision around this spot would be completely unaffected. Nonetheless, he human eye is an imperfect optical device. Light from a visual environment enters the eye through the pupil and is imaged on the retina. Additional light enters the eye by transillumination through the iris and sclera. Part of this light is absorbed in photoreceptor photopigments or other pigments such as melanin. The other part, before it is detected by the retina, it encounters a number of scattering sources internal to the eye (i.e., the cornea, lens, and the fundus of the retina), which collectively produce intraocular scatter. The scattering spatially spreads out the illumination on the retina, directing photons across angles beyond the retinal image of the light source itself.

Intraocular scatter is caused by spatial variations in the refractive index within the optical media and occurs primarily during transmission through the eye's cornea and lens, and via reflections from the retina, which can be absorbed by different retinal areas. There is also a smaller contribution from the

ocular wall, comprising the sclera and the iris. These internal scatter sources give rise to unfocused light, manifesting as glare across the scene, is termed "stray" light. Light scattering, reflection, and absorption determine the spectrum of intraocular stray light. The directionality and spectrum of scattered light depends on the density and size of the particles that scatter the light. Particle density determines the intensity of light scattering. Particle size determines its directionality and wavelength dependence.

Light scattering theory may be characterized in terms of the following two theoretical models: Rayleigh scattering and Mie scattering. Rayleigh scattering, or small particle scattering, has no preferential direction, but small particles scatter shorter wavelengths more efficiently than longer wavelengths. Mie scattering, or large particle scattering, is not wavelength dependent, but the light is preferentially scattered in the forward direction (i.e., the direction of the light). Mie scattering by retinal pigment epithelial melanin granules, roughly 1 μ in diameter, improves retinal image contrast by suppressing side scattered light. Stray light from the cornea and lens decreases with increasing wavelength, showing the influence of small particle scattering. Stray light from fundus reflectance or transillumination increases with increasing wavelength, showing the influence of the decreased optical absorption of melanin and hemoglobin in the red end of the visible spectrum. The net effect is that stray light reaching the fovea has little wavelength dependence. Thus, blue-white sources of light should produce no more foveal stray light or disability glare than white headlights of the same luminance, which was indeed verified by Flannagan et al [50]. The levels of stray light are variable for individuals, with some common factors such as increased scatter with age, primarily because of yellowing of the lens and possible appearance of cataracts, and reduced scatter experienced by individuals with dark-colored eyes, owing to greater pigment absorption and reduced ocular wall transmission [2, 6, 47].

Human intraocular eye scattering of light is modeled by the standard CIE (*Comission Internationale de l'Eclairage*, or International Commission on Illumination) equation, referred to as CIE General Disability Glare equation:

$$\frac{L_{\text{veil}}}{E_{\text{glare}}} = \frac{10}{\theta^3} + \left[1 + \left(\frac{\text{A}}{62.5}\right)^4\right] \times \left[\frac{5}{\theta^2 + \frac{0.1\times \text{p}}{\theta}}\right] + 0.025 \times \text{p} \tag{2.5}$$

where θ is the angle between the glare source and the line of sight with a full range validity of $0.1^\circ < \theta < 100^\circ$, A is the age of the individual (years) and p represents the eye pigmentation factor (which ranges from 0 for black eyes, 0.5 for brown eyes,1 for light blue eyes and 1.2 for very light eyes). E_{glare} is the illuminance (lux) incident on the cornea and L_{veil} (cd/m^2) is the veiling luminance, or equivalent veiling luminance, defined as the effective luminance that reached the retina after intraocular scattering. It is worth noting that from (2.5), the veiling luminance increases rapidly as the incident angle of the glare source approaches the visual axis [2, 3].

A source of light that is sufficiently bright to cause disability glare will also produce photostress and afterimages. Photostress is a psychophysical process which persists after light exposure because dark adaptation takes time to restore visual sensitivity to its pre-exposure level. An afterimage is the visual impression which appears more or less immediately after the decay of the stimulation at the site where the irradiation took place. Areas on the retina which have been exposed by a source of light at a sufficient level lose their sensitivity. If an individual looks onto a homogeneously illuminated area in the visual field, with a higher or lower luminance, after the exposure from a temporary blinding light source has ceased local adaptation processes, will result in various interpretations of the new visual field as far as brightness, hue, and saturation are regarded: i.e., an afterimage will be generated. The afterimage can be considered as a high degree of local retinal photostress. The appearance of an afterimage is a temporary

phenomenon and does not result in a permanent scotoma[2] [6, 51].

2.3 Concentration

Mental effort is not always the same: it depends on the task and the subjective difficulty of the task, which is determined by the individual's mental capacity. EEG is considered the most effective and objective indicator of attention level and it has been shown to be able to be correlated with varying degrees of mental effort [20]. In order to elicit varying degrees of mental effort, one must carefully consider what task should be used, in order to simply and easily modify the degree of difficulty. Two such tasks are the widely used mental arithmetic and N-back [14, 16, 52, 53]. Mental effort is related to attention, concentration and mental workload, which are terms widely used in the literature. These terms generate confusion due to their meaning being similar or even the same across the literature.

Attention is the ability to focus on certain aspects of the environment that one considers important or interesting and to flexibly manipulate this information. Arousal and alertness are prerequisites for attention. Attention, in turn, is a prerequisite for memory, communication, and executive brain functions, although these cognitive processes will, in turn, determine to what we attend.

Attention can be viewed as an organizing force for all behavior. As such, it encompasses the abilities to receive information, to select what is relevant from the incoming stimulation, to persevere at and complete an activity, and to change course when appropriate. Many aspects of frontal lobe function are increasingly being interpreted in attentional terms. Attention thus performs a pervasive role in enabling a person to successfully complete diverse cognitive, personal care, social, and educational/vocational activities [54].

There are five main categories of attention:[54]

- Focused attention is the ability to perceive individual items of information.
- Sustained attention is commonly called concentration, which predominantly involves vigilance.
- Selective attention is the ability to avoid distractions from both external (e.g., noise) and internal (e.g, worries) stimuli.
- Alternating attention is the ability to shift the focus of attention and to alter it between tasks.
- Divided attention is the ability to respond to multiple tasks at the same time, and as one would expect, this is much more difficult to accomplish within the same modality (e.g., vision) as it is between modalities (e.g., vision and audition).

These five categories of attention are mainly originated in the frontal-temporal regions of the brain, and can function in either a conscious or automatic mode, although most tasks require a combination of both. Conscious attention is slow and effortful, requires focus and concentration, and involves serial processing (i.e., when learning a new skill or solving a problem). On the other hand, automatic attention is rapid and involves parallel processing (i.e., when performing a learned skill) [54].

Mental workload is of primary interest as it has a direct impact on users' performance in executing tasks. Even though there is no agreement upon its definition, mental workload can be seen in terms of resources or mental energy expended, including memory effort, decision making or alertness. Mental

[2] A scotoma is a blind spot in the visual field.

workload is usually defined as the ratio between task demands and a person's capacity where the workload is high when task demands are close to exceeding capacity [55].

For the purpose of this work, the term "concentration" shall be an amalgam of the previously defined focused attention, sustained attention and selective attention. Mental workload can be seen in terms of mental energy expended, and concentration requires mental effort, so it could be said that concentration is a form of mental workload. However, henceforth concentration will be considered as a state required to engage in mental workload, that is, concentration is a state of mental readiness that not only enables but must be proportional to mental workload (e.g., a high level of mental workload requires an equally high level of concentration and a low level of mental workload requires an equally low level of concentration). Therefore, by measuring the changes in mental workload, one could make inferences about the state of concentration of an individual.

The working memory (WM) system has been defined as a system for the temporary holding and manipulation of information during the performance of a range of cognitive tasks such as comprehension, learning, and reasoning. Memory load is considered to be a major component and reasonable approximation of mental workload [15]. Tasks that evaluate working memory are, therefore, good indicators of mental workload levels. Two of the most commonly used tasks in neuroscience to study working memory are the N-back and digit span tasks.

The N-back task is one of the most widely used experimental paradigms in cognitive neuroscience in studies of working memory, in which subjects are instructed to monitor a sequence of stimuli and to respond whether a stimulus presented is the same as the one presented n trials previously, where n is a pre-specified integer, varying usually from 0 to 3, as it can be seen in Figure 2.9 (e.g., in the 0-back condition, subjects are asked to respond when the target letter "X" appears on the screen, whereas in the other N-back conditions a letter is a target if it is identical to the letter presented n trials before). During the performance of this working memory task the stimuli are sequentially registered and stored, and the task performance requires continuous updating of stimulus information. The increase of memory load in the N-back task is typically witnessed on the behavioral level as increased reaction times and as an increased number of incorrect responses [17, 25].

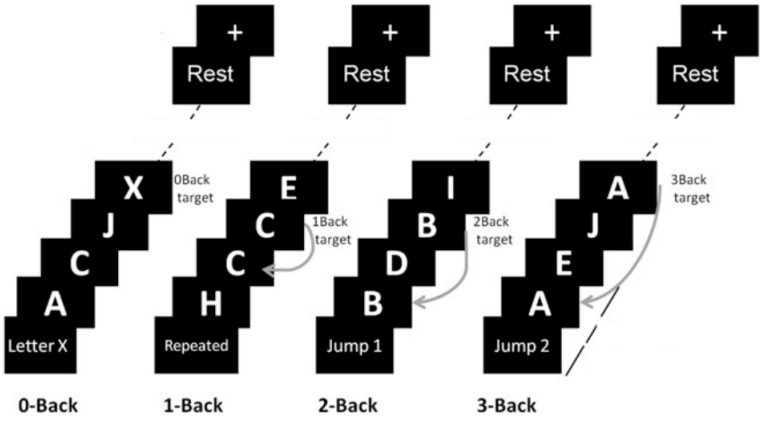

Figure 2.9: Diagram of the N-back digit task with 3 difficulty levels. From left to right: the 0-back, 1-back, 2-back and 3-back digit tasks. Adapted from [25].

Its application has enabled researchers to clearly identify the anatomical substrates of the WM system, including the following six cortical regions: (1) the bilateral posterior parietal cortex, (2) the

bilateral premotor cortex, (3) the dorsal cingulate/medial premotor cortex, including the supplementary motor area, (4) the bilateral rostral prefrontal cortex or frontal pole, (5) the bilateral dorsolateral prefrontal cortex and (6) the bilateral-mid-ventrolateral prefrontal cortex. Moreover, the nature of the N-back paradigm is such that studies can be designed to examine the effects of WM load variations [17, 25].

2.4 Summary

In this chapter, we introduced the concept of electroencephalography, a non-invasive method for measuring brain activity on a subject's scalp and approached a commercially available solution for measuring brain activity, the BITalino board. We also described the fundamental concepts of one of the main methods of EEG analysis, the time-frequency transformation using Multitapers, which is used in this work.

The concepts regarding glare were introduced, namely the dazzling glare, which is a type of glare that impairs vision. Concentration is the ability to perceive individual items of information in a vigilance state, while at the same time being able to avoid distractions from both external and internal stimuli. Finally, two of the most widely used tasks for working memory testing were introduced: the N-back task and the digit span task.

Chapter 3

State of the Art

In this chapter, we present and analyze the most relevant and recent work that was developed on the two paradigms of dazzle effect evaluation and concentration level measurement, as well as the intersection of both. This chapter sets the foundation for the methodology used in this work, which is discussed in the next chapter.

3.1 Dazzling Effect Evaluation

In this section we discuss the work that has been done on the topic of the dazzling effect evaluation, focusing on the type of light source used, the methods and findings.

The work on dazzling effect quantification using EEG measurement is rather scarce in the literature. In fact, there is, to the author's knowledge, only one paper that effectively tried to approach it. Most of the literature that can be found on the subject, refer to subjective rating scales of performance or discomfort and accuracy on the performance of a given task. Therefore, the work done by Lee *et al* [1], will be presented in detail given its relevance for being the study that more closely resembles the work done on this thesis.

Lee *et al* [1] studied the effect of glare source and illumination levels using a quantitative methodology. The main features measured were EEG signals, pupil diameter and a subjective discomfort glare scale, during a simulation of an oncoming vehicle on typical night time driving environment. To achieve this, they used a set of distinct tools, namely: a pocket lux-2 metre, a small portable illumination measurement tool; three light sources: Halogen, HID and LED; FaceLab-5, which is a eye tracking system; and the Biopac system, which is a system for biosignal measurement.

Three types of headlamps were used: Halogen, HID and LED. Each of these light sources had its spectral distribution measured and tabulated. The surface area of the halogen lamp is 50.24 cm^2, the HID lamp is 56.25 cm^2 and the LED lamp was 36.75 cm^2. The variable conditions in this study were the type of headlight and the three pupil-plane illumination levels: 0.7, 2 and 5 lux.

EEG measurements were collected from the F_z and O_1 sites. The pupil diameter was measured with the eye-tracker, which has a 60 Hz sampling rate. The data collected from the experiment concerning pupil size and EEG data were separated according to the type of headlamp and illumination condition. An analysis of variance (ANOVA) was performed to test statistical significance, and then a *post hoc* test was conducted. Ocular artifacts were present, which were removed from the EEG signal using the Biopac software. The EEG power spectrum was obtained through pure FFT analysis and a specific EEG metric, the EEG theta band ratio was estimated using the power of all frequency bands, as defined in the

following equation:

$$\theta \; band \; power \; ratio = \frac{\theta \; power}{(\theta + \alpha + \delta + \gamma \; power)} \tag{3.1}$$

The participants were seated in a darkroom designed to simulate typical night time driving conditions at background luminance of 1 cd/m^2, according to the schematic view of the experimental layout on Figure 3.1. To simulate an oncoming vehicle at 30 m away, the headlamp was installed 0.66 m from the floor, and the eye level of the seated subject was 1.15 m from the floor. The eye-tracker was installed 50 cm in front of the participants. The vertical eccentricity was 4.4 m, and the horizontal eccentricity was 5°. The participants were seated 4.4 m away from the glare source and a source aperture of 2.1 cm was used in this study. At the start of the experiment, the study procedure was explained to the participants. A chin rest was used to fix the participant's seated eye-level height and an EEG cap was used to position the electrodes. Each trial took about three seconds, with a 20-minute rest between each trial. At the end of each trial, the subjects were asked to rate discomfort, according to the 10-point Borg scale.

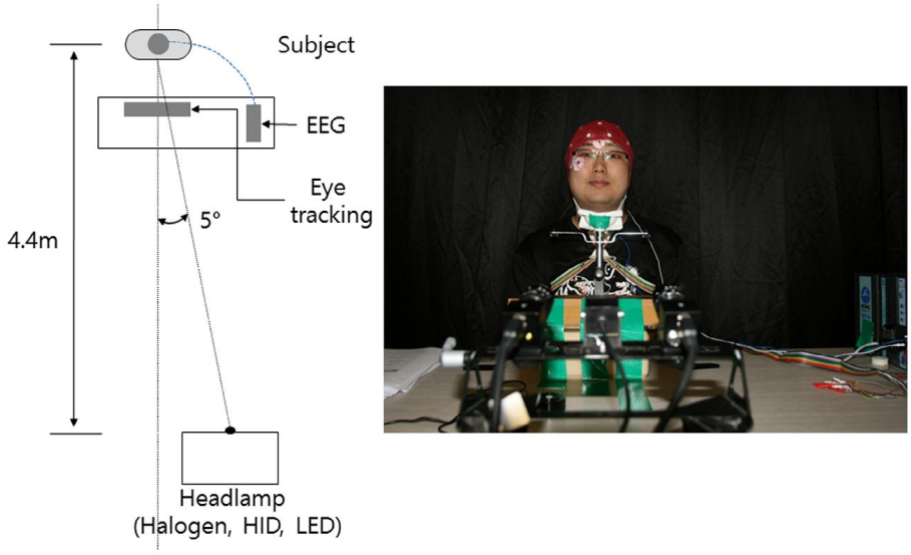

Figure 3.1: Schematic view of the experimental layout on the left and the darkroom that was designed to simulate typical night-time driving conditions on the right are shown. Adapted from [1]).

The results showed that pupil sizes differed significantly between the headlamp type ($p = 0.01$) and between illumination conditions ($p = 0.01$). Pupil size was bigger with the Halogen lamp type and smaller with the LED type, according to the following relation: halogen $>$ HID $>$ LED. In addition, when illumination was increased, pupil size decreased according to the relation: 0.7 lx $>$ 2 lx $>$ 5 lx. The theta waves from the F_z electrode differed significantly between the illumination conditions ($p = 0.1$), but not between headlamp types. When illumination increased, theta waves also increased according to the relation: 0.7 lx $>$ 2 lx $>$ 5 lx. There was no significant main effect of headlamp type and illumination condition on the O_1 electrode. Subjective discomfort differed significantly between the headlamp types ($p = 0.01$) and between the illumination conditions ($p = 0.01$). These results showed that the subjective discomfort with a LED headlamp was greater than other lamps following the order: Halogen $<$ HID $<$ LED. When the illumination increased, the subjective discomfort increased with the increase in illumination level.

3.1.1 Impact of glare on vehicle operation tasks

The occurrence of the dazzle effect during the operation of a vehicle may potentially put the lives of both drivers and passengers at risk. The dazzle effect tends to happen at night, because the low background illuminance increases the contrast of light sources, resulting in the transformation of, for example, car's headlights in immensely bright lights which would otherwise be unnoticeable during the daytime. The potential danger of night-time driving has lead to the development of scientific literature exploring the topic.

Car driving

Theeuwes *et al* [10] conducted a study in order to find if there was a relation between glare and driving performance. They prepared a stretch of road with gray plywood boards, which is common in pedestrian visibility studies, and had participants drive on this road at night, with actual traffic. The vehicle that the participants drove was modified with an apparatus that allowed the researchers to simulate glare equivalent to a continuous stream of oncoming cars with four light levels corresponding to glare illuminances of 0 (control), 0.28, 0.55, and 1.1 lx at the observer's eye. Besides measuring driving performance, the deBoer rating scale was given to the participants to assess the discomfort level. The color temperature of the light of the lightning rig was about 3100 K, which matches very closely with the headlamp colors on the road. They concluded that the relatively low glare source caused a significant drop in detecting simulated pedestrians along the roadside and made participants drive significantly slower on dark and winding roads. This effect increases with age, as older participants showed the largest drop in pedestrian detection performance and reduced their driving speed the most. They also concluded that the deBoer rating scale, which is of common use for discomfort glare, is useless as a predictor of driving performance, possibly due to each individual's adaptive driving mechanisms which may be able to reduce glare discomfort while affecting driving skill. The increase of glare susceptibility with age, is reported by many other studies [56–58]. Driving speed reduction caused by visual distractors was verified later by Engstrom *et al* [59], which increases the weight on the hypothesis of the deteriorating effect of glare on the performance of concentration-demanding tasks.

Laser dazzling of car drivers is another area of focus on the scientific literature, mainly due to its importance to the military and law enforcement officers. Lasers can be effectively used to provoke a strong dazzle effect on car drivers, making them unable to continue moving in the direction of the laser or stopping the car, compromising their ability to perform offensive actions. By having participants drive through an obstacle road track, Steinvall *et al* [5] concluded that laser dazzling does indeed reduce the driver's ability to make visual judgments and to maneuver the car. The effect of laser dazzling is highly dependent on the general illumination (i.e., during the day the influence on driving is limited but at dusk and darkness it increases with decreasing background illuminance). It is worth noting that, at night time driving, some drivers even tended to completely stop due to the dazzling. Although the effect of laser dazzling during daylight is minimal, it can still give an obvious warning effect in daylight. The laser used had a power of 5 W/m^2, which is half of the Maximum Permissible Exposure (MPE), resulting in no damage to the vision of the participants.

Aircraft operation

Glare is an important effect to an aircraft pilot. Besides occasionally suffering from natural sunlight glare, airplane pilots have been targets of laser attacks, which may put at risk the lives of everyone aboard the

aircraft. There have been several studies on this matter, to assess the impact of this type of attack.

Nakagawara *et al* [11] queried the National Transportation Safety Board (NTSB) Aviation Accident/Incident Database from 1988 to 1998 for incidents related to sunlight glare. They found that 130 accidents had occurred in which direct or reflected glare from the sun was found to be a contributing factor in the event. These accidents represent 0.5% of the total number of accidents in that period. Temporary visual impairment due to glare can have serious consequences for pilots during critical maneuvers performed at low altitude. The study revealed that 55% of the accidents involving glare occurred during the approach/landing and takeoff/departure phases of the flight, indicating the importance of perfect pilot vision during these phases. Rogers *et al* [12] further revealed that the presence of glare causes the most important impairment in the pilot's ability to see their instruments and to fly their airplane when the glare was straight ahead, as well as slightly to the side. The more forward the glare is and the longer the glare duration, the greater the impairment to the pilot's ability to see their instruments and to fly the aircraft. The procedure consisted in participants using a flight simulator that was illuminated by a halogen lamp with such a power that the luminance measured at the eye of the participants was between 1000 to 2000 lux, which was sufficient to simulate glare from reflected sunlight and produce after-images.

Several hundred incidents involving the illumination of aircrew members by laser light have been reported in recent years, including several that could have had serious consequences. Another study by Nakagawara *et al* [9] explored the effects of laser attacks on pilots undergoing flight simulations with lasers of power ranging from 0.5 to 50 $\mu W/cm^2$ (class 1 lasers). They concluded, by means of a subjective survey, that 75% of the subjects experienced adverse visual effects resulting in some degree of operational difficulty when illuminated by low-level laser radiation. This difficulty is caused by the appearance of effects such as glare, flash-blindness, and afterimages which resulted in missed approaches and even relinquishment of command to the co-pilot. Furthermore, even the lowest laser power tested is unacceptable within the immediate proximity of the airport which is defined as Laser-Free Zone (LFZ), extending 2 nautical miles in all directions from the runway centerline and around 600 m of airspace height.

The aforementioned studies where lasers were used in subjects, were performed by the Federal Aviation Administration of the United States of America. However, there are issues with exposing humans to lasers, even at irradiance levels below the safety limits for potential eye damage. Human testing can be costly and time-consuming, requiring ethics approvals which can also limit the ability to conduct repetitive testing and fully explore the matter. There are several models available on the literature that attempt to simulate the laser eye dazzle effect in order to overcome these limitations [48, 60]. However, these models have yet to be applied in tasks such as a flight simulation in order to further explore the effects of laser eye dazzle on aircraft pilot performance.

3.2 Concentration Measurement

In this section we discuss the work that has been done on the topic of concentration measurement, focusing on the EEG setup, the EEG analysis, the tasks used and findings.

3.2.1 Overview

There have been many attempts at measuring concentration with EEG in the literature.

It is known for a long time that concentration tasks block alpha activity [61]. As a matter of fact,

a game has been developed which the player controlled with his alpha activity: relaxation increased alpha activity and concentration blocked alpha activity. 70% of the participants were able to successfully control the game with only one electrode placed at P_z location [62]. Not only the alpha activity is blocked, but, for arithmetic and visuospatial tasks, the peak alpha frequency tends to shift toward higher frequency within the alpha band as the task difficulty increases, suggesting that peak alpha frequency shift is a reliable index of cognitive activity. Alpha absolute power (AP) also tends to decrease with increasing task difficulty. Moreover, the hemispheric asymmetry in the peak alpha frequency is larger in the difficult tasks and is more pronounced in the frontal area, even though for easy tasks, processing is bilateral. In other words, hemispheric lateralization is likely to increase as task demands increase [52]. Mental calculation tasks affect other frequency bands as well: delta AP and relative power (RP) increases while alpha AP and RP decrease during tasks, as well as an increased beta AP in frontal leads is observed [53].

There is now ample evidence that the alpha and theta frequency band oscillations reflect cognitive and memory performance. Gevins *et al* [63] demonstrated that theta power increased with increasing difficulty in the N-back task, while the alpha power decreased. Alpha signal was most prominent over posterior regions and it was highly sensitive to task difficulty. They further advanced than the theta power increase was also correlated with the practice of the task over an extended period, suggesting that it is related to increased effort to maintain a task-related attentional set. Jensen *et al* [64] provided further insight on the theta frequency band relation to memory, conducting an EEG measurement during a Sternberg task (subjects are required to retain a list of 1, 3, 5 or 7 visually presented digits during a 3 seconds retention period and afterward numbers are presented to the subject, which has to indicate if each number belongs to the memorized list). It was found that the frontal activity in the theta band increased parametrically with the number of items retained in working memory. Following the memory task, the theta activity reduced, indicating that theta oscillations generated in frontal brain regions play an active role in memory maintenance.

The delta frequency band activity might be related to performance during mental tasks. Again, with the use of the Sternberg task, Harmony *et al* reported that there was a greater increase of delta activity during difficult than during the easy version of the task. The same occurred for the beta frequency band [65].

3.2.2 N-back

Overview

The N-back task is used extensively in the literature as a working memory paradigm and it is increasingly used as a measure of individual differences. It is a very useful tool for the experimental investigation of working memory processes because it allows the load to be manipulated in a very simple, straightforward way. It is good at predicting inter-individual differences in other higher cognitive functions, such as fluid intelligence, especially if used at higher levels of load [16].

Design

Pesonen *et al* [17] studied the brain oscillatory 4-30 Hz responses during a visual N-back memory task with varying memory load.

The experiment was performed on thirty-six participants. The task presented to the participants was the classic N-back task, with 4 levels of difficulty (0-back, 1-back, 2-back, and 3-back), corresponding

to working memory load varying from 0 to 3 items.

In order to reduce muscle artifacts in the EEG signal, the participant was instructed to assume a comfortable position and to avoid movement. The participants were instructed to look at a TV screen placed in front of them at a distance of 1.6m and to avoid unnecessary movements. The participants were instructed to respond as fast and as accurately as possible to each stimulus by pressing one button for targets and another for non-targets. The participants' responses were monitored (reaction times and correct responses) and only trials with correct responses were included in the EEG analysis. The length of a single memory load experimental condition was 10 min (plus response time) and there were 120 trials in each memory load experimental condition. Each task with different memory load conditions was performed twice. The order of the four memory load conditions was counterbalanced between the participants. The stimulus was presented on the TV screen until the participant responded, and after the response with 2500 ms ISI (inter-stimulus interval), the next stimulus was presented. The visual stimuli were generated and presented using a NeuroScan STIM software (Compumedics, USA).

The EEG data were gathered from 19 electrodes by using Electro-Cap (Electro-Cap International, Inc., USA.) with the international 10/20 system of electrode placement. Additionally, two EOG electrodes were placed on the outer sides of the eyes and all electrodes were referred to the right mastoid electrode. The EEG data were recorded using Neuroscan 386 Scan 4.1 data acquisition system (Compumedics, USA) with a frequency band of 0.1-50 GHz and sampled at 250 Hz. The impedance of recording electrodes was monitored and kept always below 5 kΩ.

The digitalized EEG data were processed in MATLAB©(MathWorks, Inc., USA) environment. The data were processed to time-frequency representations (TFR's), which display the power of the EEG signal as a function of time and frequency in the same matrix. The calculations were performed for each EEG channel separately. Artifact rejection was set to $\pm$ 100μV. Poor quality signal channels were excluded from further analysis by means of visual inspection of the TFR's displaying the absolute power values.

TFR's for each participant and for each event (presentation of target/non-target stimuli in each memory load condition) were calculated using a Morlet wavelet (width 8) for EEG frequencies 4-30 Hz. The EEG data were epoched using a time window of 2600 ms for the stimuli and the reference was epoched using a time window of 1000 ms prior to each stimulus presentation and was averaged over time. The TFR's for the events and the reference were averaged separately for each participant. Thereafter, the relative difference in the power of EEG between the reference and the stimulus presentation (target/non-target) was calculated as a function of time and frequency. The difference was expressed as ERD/ERS percentage in which negative values indicate relative power decrease (ERD) and positive values indicate relative power increase (ERS). The averaged ERD/ERS TFR's were calculated for both stimulus types, targets and non-targets, in all four memory load conditions. The ERD/ERS TFR's were displayed for five electrode locations (F_{p1}, F_{p2}, F_3, F_z, F_4, F_7, T_3, C_3, T_5, F_8, T_4, C_4, T_6, C_z, P_3, P_z, P_4, O_1, O_2) as a function of time (0-1800ms) and frequency (4-30 Hz).

The non-parametric Quade test (statistical test to compare groups of samples) was used to assess the statistical significance of the ERD/ERS responses *per se*. Statistically significant differences between memory load conditions for target and non-target stimuli were assessed pair-wise (0- vs. 1-back, 1- vs. 2-back, 2- vs. 3-back) using the same statistical test. Additionally, the significance of any differences between the ERD/ERS responses elicited by targets and non-targets within each memory load condition was assessed.

Reaction times and the number of correct responses to the task were recorded for each memory load condition. The statistical significance of the differences in the response times and the percentages of

correct answers between the memory load conditions were analyzed using a repeated measures analysis of variance (ANOVA) with one within subjects for memory load factor (with the respective 4 levels of difficulty). Factors with three or more levels carry an assumption known as sphericity or circularity. Violation of this assumption results in positively biased testes. One method to compensate for this bias is to adjust the degrees of freedom using the Greenhouse-Geisser procedure (used to assess the change in a continuous outcome with three or more observations across time or within-subjects). When appropriate, this procedure was used to compensate for non-sphericity. Corrected degrees of freedom and corrected p-values are displayed for the behavioral results.

Reaction times increased and the number of correct responses decreased with increasing memory load. The presentation of all visual stimuli elicited long-lasting (0-1800ms) theta (~4-6 Hz) ERS responses. The magnitude of these responses was statistically significant and of greater magnitude for the target stimuli than for the non-target stimuli, especially in the parietal recording sites.

Statistically significant alpha frequency range (~8-12 Hz) ERD responses were observed in all memory load conditions for both targets and non-targets beginning at 100 ms after visual stimulus onset. The duration of these alpha ERD responses increased with increasing memory load, being ~100-500 ms in the 0-back and ~100-1600 ms in the 3-back memory load condition. The magnitude of these alpha ERD responses in the 2 and 3-back memory load conditions was statistically significant and of greater magnitude for the targets compared to the non-targets.

Finally, in all memory load conditions, early appearing beta ERD responses were elicited with an onset at 100 ms. With increasing memory load, these beta rhythm ERD responses became longer in duration. Significant differences between the magnitude of these beta ERD responses between the targets and non-targets arose in the 2 and 3-back memory load conditions, such that greater beta ERD was observed for the targets than for the non-targets. Only in the 0 and 1-back memory load conditions, frontal ERS responses were observed approximately at ~500-1800 ms after visual stimulus onset. These ERS responses were of greater magnitude for the target stimuli as compared to non-targets, which effect was witnessed especially in the frontal electrodes.

Frequency band power values are not the only features affected by varying N-back task load. Watter *et al* hypothesized the N-back task to be a dual task, permitting the imposition of parametrically increasing attentional and working memory demands, while keeping constant the demands of an embedded matching subtask. They analyzed the visual ERP's during the performance of varying levels of difficulty of the task and reported that the P300 latency was constant across all levels, reflecting constant perceptual and cognitive demands of the matching subtask. However, P300 peak amplitude decreased with increasing memory load, reflecting the reallocation of attention and processing capacity away from the matching subtask to working memory activity, which supported their initial hypothesis [66].

Distinguishing various workload levels of the N-back task can be performed by the use of classification models. Brouwer *et al* developed, trained and tested different classification models using ERP features, frequency power features or a combination of both. All three types of models achieved classification accuracies between 80% and 90% when distinguishing between the highest and the lowest workload condition after 2 minutes of data. Power in the alpha (alpha power decreases with increasing memory load) and less so in the theta band generally seemed to contain the most relevant features for distinguishing the highest and lowest N-back level in the models based on spectral power. Results from the classification models using ERP's indicated that the bulk of the information was present in a time interval consistent with the P300, confirming the results of Watter *et al* [15, 66].

3.3 Discussion

Introduction

This section is dedicated to the summarization and discussion of the works referenced in the former sections in this chapter. A general overview of methods and findings is presented.

Discussion

The reader should now be aware of the existent gap in the literature. There has been only one paper attempting to quantitatively measure the effects of task impairment caused by the dazzle effect with EEG. This thesis will focus on closing this gap in knowledge.

Table 3.1 presents a summary of studies that shed light in different aspects of concentration measurement. From the inspection of the table, it can be seen that absolute power and relative power, particularly in the alpha and the theta frequency bands, are of great importance for the detection of workload levels. The analysis of ERP's is also useful, especially the P300 peak, which was also proved to be important on the measurement of mental workload levels.

The literature is abundant with the FFT method for time-frequency decomposition of EEG signals [63, 65, 66], but the Multitaper spectral estimation method stands out for delivering a superior analytical performance [41].

The amount of electrodes used in the selected studies is highly variable, ranging from 115 electrodes mesh [63] to one electrode [62]. However, most studies seem to use electrode numbers ranging from 15 to 30 electrodes. The electrode placement is made such as every area of the brain is covered. The great majority of studies use monopolar electrode setups and all studies followed the 10/20 international system for electrode placement.

In the selected studies, only one study tried to correct artifacts [14], the others rejected epochs in which artifacts were present based either on visual inspection and/or an established rule (such as rejection if the root mean square amplitude of a given epoch was greater than 30 μV [66]). Most studies had EOG electrodes placed in order to easily identify ocular artifacts [14, 65, 67].

In the analysis of the EEG signal, averaging of the signal or its power is common practice. Averaging of the signal epochs was used for obtaining ERP's and performed on every artifact-free epoch of each condition [15, 63, 66, 67]. In all studies, epochs contained a period of pre-stimulus signal and a period of the post-stimulus signal. The most commonly used time interval for epochs was 2000 ms, which was then divided in many distinct ways for pre-stimulus and post-stimulus (e.g., 500 ms pre-stimulus and 1500 ms post-stimulus) [15, 66, 67].

Even though not all studies present in table 3.1 are based on the N-back task, all of the studies concern mental workload tasks and therefore their findings can be potentially identified in analyzing the results of the N-back tasks conducted on this thesis.

3.4 Summary

In this chapter, we presented two paradigms: dazzle effect evaluation and concentration measurement by EEG. Furthermore, we presented works pertaining to important information to the design of this thesis's experiment, how to analyze the data and what to expect.

Table 3.1: Selected works associated with concentration measurement by EEG.

Year, author	Type of Task	Major Findings
1984 Osaka *et al.*[52]	Arithmetic and Visuo-spatial	Peak alpha frequency shifts toward higher frequency with increased task difficulty. Hemispheric asymmetry increases and alpha power decreases with increased task difficulty.
1995 Fernandez *et al.*[53]	Arithmetic	Delta AP and RP increased while Alpha AP and RP decreased during tasks. Beta AP increased in frontal leads.
1996 Harmony *et al.*[65]	Sternberg (working memory)	Greater increase of delta and beta AP during difficult than during the easy version of the task.
1997 Gevins *et al.*[63]	N-back	Increase in theta power and decrease in alpha power with increased task difficulty (memory load). Effort to maintain attentional over long periods of time increases theta power.
2001 Watter *et al.*[66]	N-back	P300 peak latency is constant across varying levels of N-back task difficulty. P300 peak amplitude decreases with increasing memory load.
2002 Jensen *et al.*[64]	Sternberg (working memory)	Theta band power increases parametrically with the number of items retained in working memory.
2003 Stipacek *et al.*	Digit-span	Desynchronization increases linearly in the upper alpha band with ascending cognitive load.
2004 Clark *et al.*[68]	Digit-span	Digit-span performance decreases with age. Positive relationship between alpha peak frequency and working memory performance.
2007 Pesonen *et al.*[17]	N-back	Alpha band and beta ERD responses during task, with ERD duration increasing with increasing memory load and reaction time.

Chapter 4

Methods

This chapter contains a detailed description of the experiment and data analysis performed in this thesis.

4.1 Introduction

The objective of the experiment is to elicit a concentration state in participants using an N-back task and to discover what EEG features are affected by dazzling the participant while performing the task.

We propose an experiment where EEG data will be acquired, through a BITalino board, from a participant which sits in front of a laptop. The laptop running Unity software will present N-back tasks which the participant will have to perform, answering with the arrow key ”up” when a stimulus is a target. The Unity software that the laptop is running controls an Arduino UNO board via serial connection that is linked by USB to the laptop. The Arduino board, via an auxiliary circuit, in turn is able to turn on and off a 10 W white LED, which is placed in front of the participant near the keyboard. The LED light source will trigger the dazzling effect upon the participant, affecting his performance on the N-back task, which should be reflected on the acquired EEG data. Figure 4.1 is an illustration of all the fundamental components in the experimental setup, which provides a graphical overview for the reader to ease the comprehension of the interplay of all the components.

It is necessary to careful design and to orchestrate all the components in order to effectively create an experimental setup capable of delivering the data that we require to perform a clear analysis. In Figure 4.2, the aforementioned fundamental components are laid out in an illustrative diagram. The experimental setup is composed of 5 fundamental components: a light source, Arduino, Unity, BITalino, and the computer. These components form two functional blocks. The first block is composed of the Light Source, Arduino, Unity software and the computer, which serves the purpose of the N-back implementation and dazzle stimuli control. The second block is composed of the BITalino board and the computer, which serves the purpose of data acquisition. The first block is connected via wire and the second block is connected wirelessly via Bluetooth.

A low-cost approach was taken on the selection of components and experimental set-up, which aims to seek an answer as to whether this kind of equipment can provide quality results for a fraction of the price, being a cheap alternative and empowering biomedical researchers.

Each of these components serves a very specific purpose and will be thoroughly detailed in the subsequent subsections.

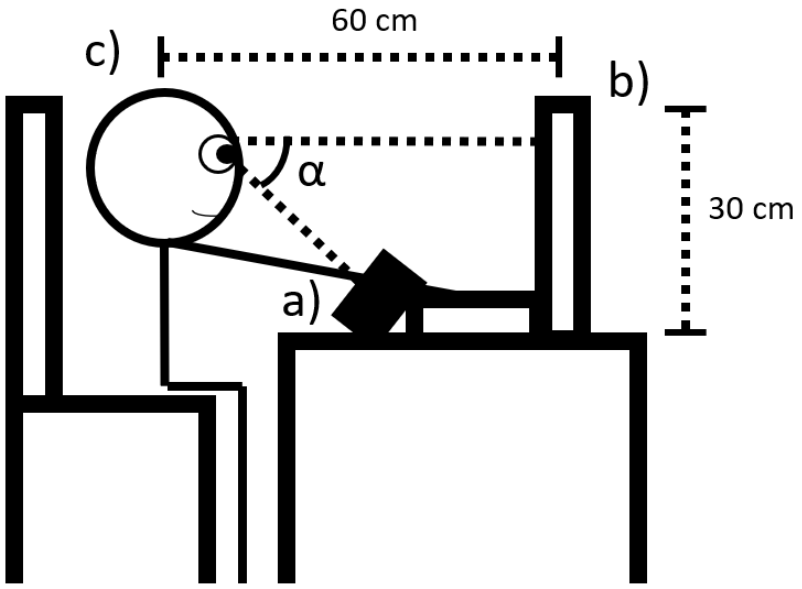

Figure 4.1: Illustration of the experimental setup. a) represents the source of light, b) represents the laptop screen where the task will be presented, c) represents the subject participating in the experiment and α, having a value of 45 °, represents the angle between the line of vision of the subject and the direction of the light source.

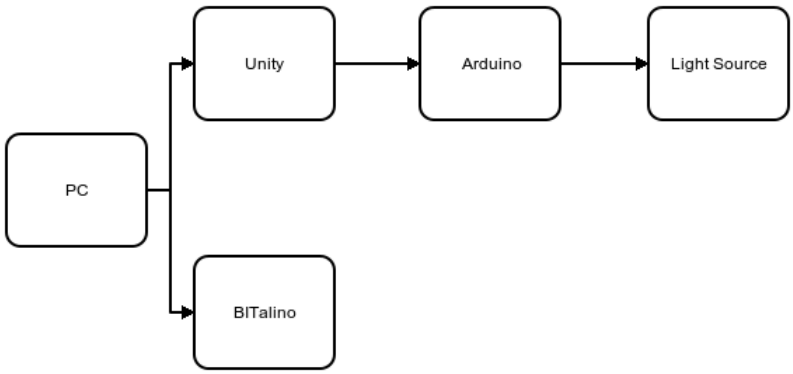

Figure 4.2: Diagram of the fundamental components of the experimental setup. The arrows indicate the flow of information through the component's chain of control. The computer receives Bluetooth signals from the BITalino board and runs the N-back tasks on Unity software which controls an Arduino via USB serial connection. The Arduino then controls a LED light source via a transistor-based circuit.

4.2 Experimental Setup

4.2.1 Unity

Unity is a cross-platform game engine that is primarily used for the creation of computer video games. It is a framework that allows for relatively easy content production due to its programming simplification. Through scripting, one can easily create simple N-back tasks. Unity's ample community, easy-to-learn production environment and the availability of a free Unity package for personal use are the three reasons for which this framework was chosen for test creation in this thesis.

The Unity component of the experimental setup has the following key responsibilities: record input information from the test subject, present the cognitive load tasks, communicate with the Arduino to control the source of light and receive biosignal data from the BITalino board.

The task itself consists in presenting a centered black letter stimulus alternated with a black fixation cross on a white screen.

The dazzle effect is dependent on age. For this reason, the age gap of the experiment's participants is between 20 to 25 years old. The application will also record the incorrect answers to each of the tasks, providing an objective measure of the subject's performance. All the information recorded by the application will be written in a text file that is created by the application itself.

The Unity application is scripted in such a way that it can communicate with the Arduino via a serial connection, through a USB port. This enables the application to indirectly control the source of light, allowing the scripting of the specific times at which the light will turn on and off and thus, the dazzling effect.

In order to acquire data from the BITalino, we used an open source library for Unity made available at the board fabricator's website, which handles the connection with great simplicity.

Cognitive Load Task and Dazzle Stimuli

The cognitive load task used in this work is the N-back task. It was chosen because of its extensive use in combination with EEG in the scientific literature, providing a solid ground of which EEG features to search for in the data analysis and which effects to expect. Furthermore, this task allows for easy manipulation of difficulty in which we can define a gradient of mental effort required for task undertaking that can be detected in specific EEG features. The implementation of a standard cognitive task with different difficulty levels that are known to generate differentiated cognitive load, allow for both the comparison of the results with those present on the literature and the experiment's direct comparison between the dazzle and no-dazzle conditions.

Before the task takes place, instructions and an initial training period are presented to the participants to familiarize them with the proceedings of the task. This training phase is performed until the participants feel that they have fully comprehended the task.

The task is organized in conditions, trials and epochs. Condition refers to a dazzle or no-dazzle scenario, with each having three trials which are specific levels of difficulty (e.g., 2-back represents one trial, 1-back represents another trial). Epochs represent the basic element of each trial and are composed of a fixation cross period, stimuli and response period. The latter contains a small period of time which we consider to be the baseline period, which we assume to be devoid of task related brain activity. In the N-back task, one epoch is composed of a single letter stimulus and subsequent fixation cross period, during which the participant has to determine if the stimulus is a target by pressing ”up” arrow key. The objective of the fixation cross is twofold: to avoid eye wandering, thus eliminating possible eye movement artifacts, and to further focus the participant's attention to the task at hand. The structure of each condition for the task is presented in Figure 4.3.

In a study involving the N-back task, Brouwer *et al* noted that the fourth difficulty level of the N-back task, the 3-back, was too difficult and participants tended to give up or performing very poorly on the task, leading to poor data quality. For this reason, there will only be 3 difficulty levels in the task presented in this work [15].

As it can be seen in Figure 4.3, the N-back task will be composed of instructions, followed by 2000 ms of a fixation cross period. This initial period ensures that a smooth transition is made from the

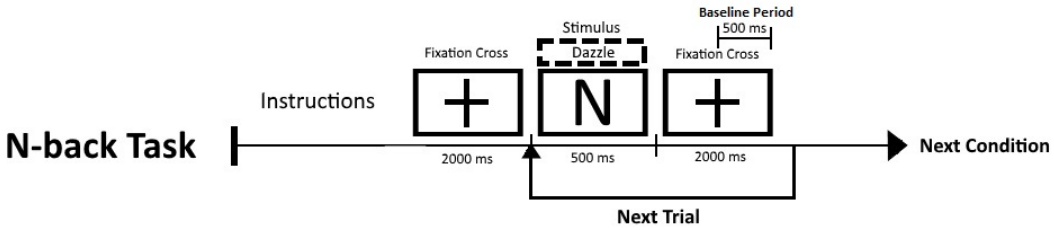

Figure 4.3: Schematic view of the N-back cognitive load task.

instruction screen and the actual task so that the participant is not overwhelmed by the task. After this initial period, the epochs begin. Epochs are composed of a letter stimuli, which lasts for 500 ms, and a fixation cross period, which lasts for 2000 ms. We consider the last 500 ms of the fixation cross period to be the baseline period, as is well beyond the expected normal response time and where we assume that task-related brain activity should be lower. Each individual epoch lasts 2500 ms in the N-back task. The duration of the letter stimulus is set to 500 ms to induce fast-thinking. During the epoch, the participant can answer if the stimulus is a target or non-target. The answer time and score is recorded by the Unity software. The scoring system implemented counts one point per wrong answer, thus, the final score of each trial is equal to the number of wrong answers given in that trial. The higher the score, the poorer the participant performed in that given trial.

Similarly to a study performed by Brouwer *et al* [15], the letters used for the N-back task are generated from the 21 consonants in the Portuguese alphabet. Vowels are excluded to reduce the likeliness of participants developing chunking strategies which reduce mental effort. Furthermore, we also applied the proportion of targets used in the study, which is of 33% from the 48 letters presented in each trial.

The N-back tasks are to be presented in random order for all the participants to eliminate expectation bias in the results. Within each task, each difficulty level is also randomized for the same reason, as it can be seen in Figure 4.4.

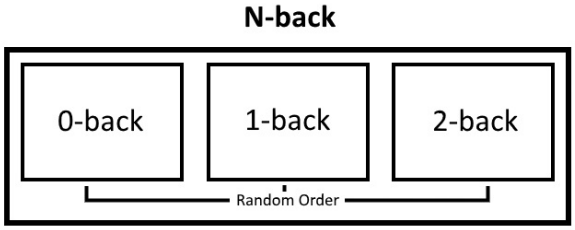

Figure 4.4: Diagram of the task order. Each level of the N-back task is to be presented randomly.

The no-dazzle condition version of the task is to be presented first and the dazzle version to be presented afterward. The structure of the two conditions is exactly the same, the only difference being the existence of dazzle stimuli at specific periods in time. The dazzle stimulus is presented during the presentation of a letter. Thus, the duration of the dazzle stimuli is always of 500 ms. The duration of the dazzle effect should not be continuous or long to avoid the participant's adaptative mechanisms from taking action and reducing the potential dazzle effect. Due to the dazzle stimulus taking place exactly during the presentation of the presentation of task-relevant stimuli, it is expected that the participant suffers a loss in the degree of concentration, due to potential confusion and difficulty to see and process

the task stimuli generated by the dazzle effect.

The estimated duration of the experiment is about 30 minutes, which we consider to be the maximum limit in order to avoid fatigue and subsequent contamination of the results (this estimate does not take into account the time spent on the placement of the electrodes on the participant, the reading of instructions, the initial training, and answers to the tasks). The experiment's structure consists in two conditions, no-dazzle and dazzle, composed of three trials (the three N-back levels), which themselves are composed of 48 epochs each (letter stimulus). This structure is ran two times, having a short interval in between of no longer than 5 minutes and are considered as independent from each other for the statistical analysis.

4.2.2 Arduino

The Arduino is an open-source electronic prototyping platform. The Arduino model used in this work is the Arduino UNO, and it's programmed to communicate with the Unity software via USB serial connection.

The Arduino board tackles the challenge of controlling a source of light via a computer. It controls a 10W LED via the circuit in Figure 4.5. This circuit is composed of a 1K Ω resistor, a TIP120 NPN Bipolar transistor and a 12V battery composed of 8 AA batteries of 1.5 V each, connected in series, which powers the circuit. The transistor receives the Arduino's signal on its gate and acts as a switch to turn on or off the relatively high voltage needed to power the LED.

Both Arduino and the auxiliary LED circuit can be seen in Figure 4.6, which represents a photograph of the experimental setup.

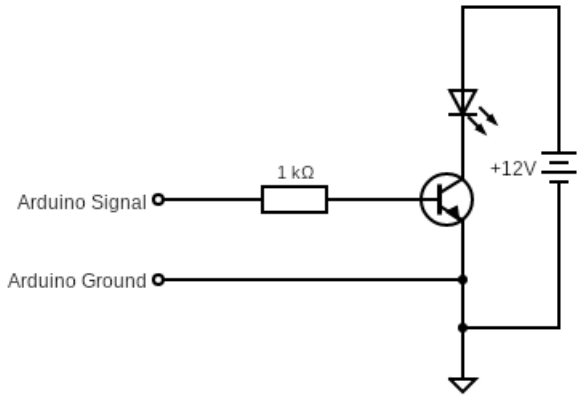

Figure 4.5: Auxiliary circuit intended to support the Arduino board in controlling the 10 W LED light source.

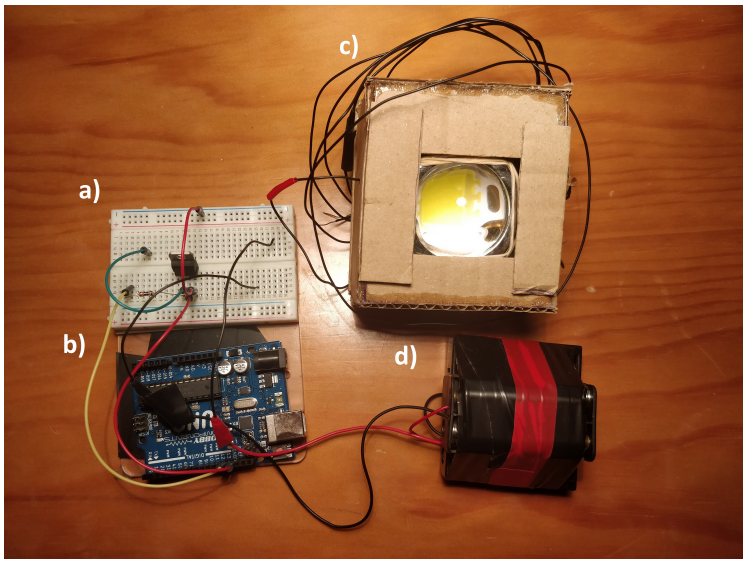

Figure 4.6: Photograph of part of the experimental setup featuring: a) the auxiliary LED circuit, b) the Arduino UNO board, c) the 10 W LED light source along with its lens encased in a cardboard protective case and d) a 12 V battery pack made of 8 AA 1.5 V batteries connected in series to power the LED.

4.2.3 Source of Light

The source of light is a 10 W white LED. The 800 Lm LED requires a 9-12V DC input voltage and up to 1A of input current. In order to obtain the maximum luminous power that the LED could provide, we supplied 12V with sufficient current draw.

The light specifications for the computer monitor used in this work are as of 11.2 ± 0.2 Lm for a white background and of 0.24 ± 0.022 Lm for a black background. The luminance was measured with a TES 1330 Digital Lux Meter.

The selection of the LED type was done based on the affordability and ease of control as well as for its luminous power. Furthermore, LEDs turn on at faster speeds than incandescent light bulbs, which was required for this experiment in order to keep latency to a minimum. The LED was enclosed in a holder which featured a glass lens to collimate the LED's light and to give it physical protection.

Figure 4.7 represents the electromagnetic spectrum of the LED light used in this work.

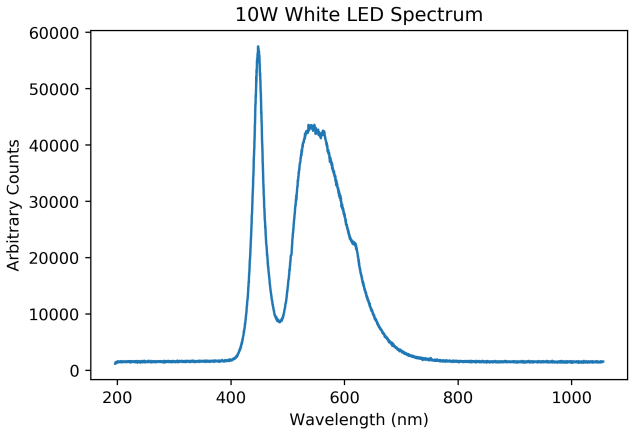

Figure 4.7: Spectrum of the 10W White LED light source used in this experiment.

4.2.4 Bitalino

The BITalino Board is an Arduino-compatible, hardware and software toolkit that has been specifically designed to deal with the requirements of body signals. It can be connected to an array of proprietary sensors that can be used to sample many signals such as EEG, EMG, heart activity (ECG) and EDA. Any combination of sensors can be used. Data is sampled at 1000 Hz. The board communicates wirelessly to a computer via Bluetooth 2.0.

The setup used in this case used a Freestyle disposition of the BITalino board, which provides flexibility as how to arrange the individual parts of the board together so as to optimize space and fit the requirements of biomedical applications. The individual Freestyle boards were: main BITalino board, wifi module, power module, battery, two EEG sensors and one EOG sensor. Due to not having enough space on the breadboard where the BITalino is connected (Figure 4.9 d)) and the necessity of having a long headband reach, the sensors were placed separately in a cardboard box and connected through long wires. The connections soldered in the Freestyle boards are represented in Figure 4.8.

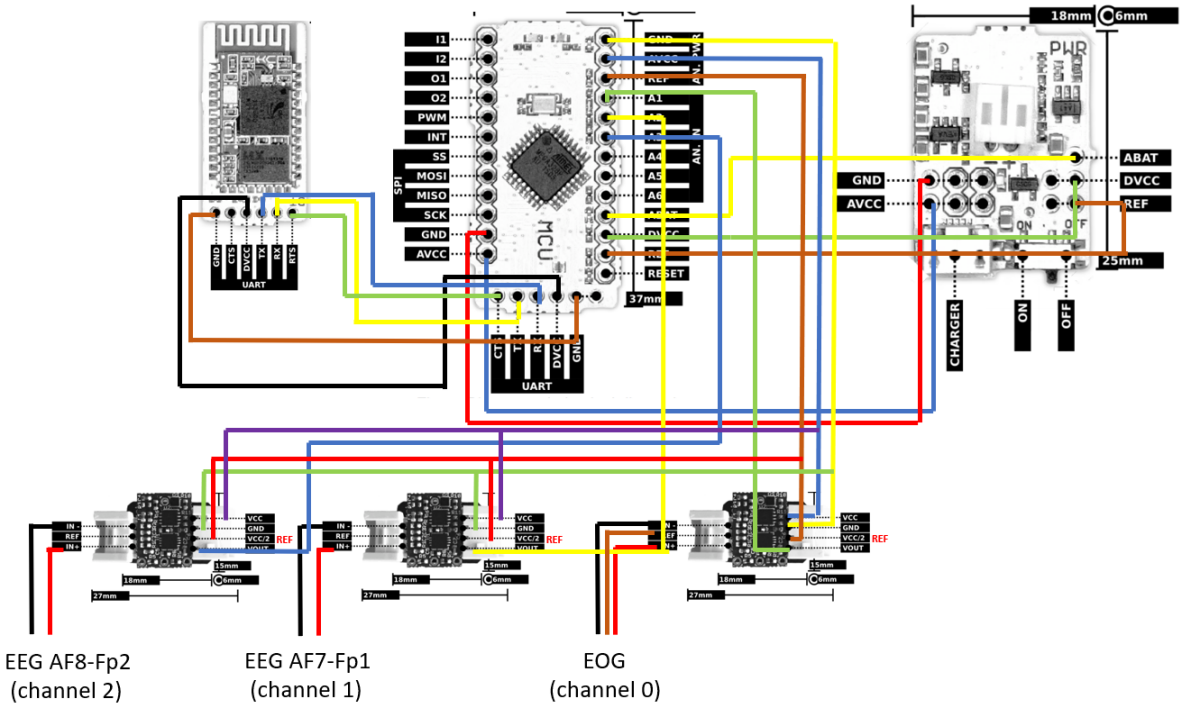

Figure 4.8: Graphical representation of the soldered wired connections between the BITalinos's wifi module (left), BITalino's main module (middle), BITalino's power module (right) and BITalino's EEG and EOG sensors (down). The colors used for wiring were selected to facilitate wire differentiation and do not correspond to a specific color code.

Signal Recording

A total of 7 gold cup electrodes were used in this experiment. Each electrode was embedded with Ten20 conductive paste to achieve better contact with the skin. The electrodes used for EEG were sewn into an elastic EEG headband and the EOG and reference electrodes were firmly fixed by electric isolating tape.

The electrodes were not intended for this board and had to be adapted. The electrode adaptation involved cutting the electrode sensible thin wires and soldering them to a more robust wire to be connected to the sensor itself. A photograph of the BITalino and the headband with the electrodes is

presented in Figure 4.9.

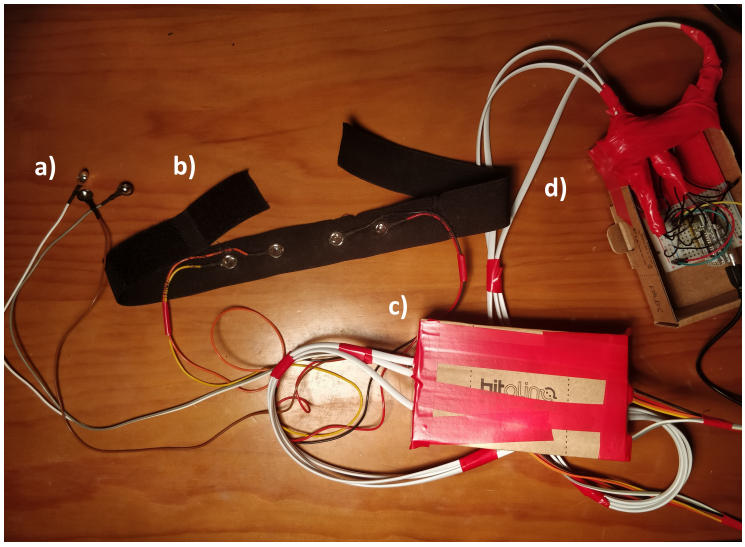

Figure 4.9: Photograph of part of the experimental setup featuring the reference and EOG electrodes a), the 4 EEG electrodes embedded in a headband b), a cardboard case containing two EEG sensors and one EOG sensor c) and the BITalino board, encased in a protective cardboard case d).

The number of electrodes used for EEG measurement were 4. The electrodes were used in dipolar mode, put in AF7 and Fp1 and in AF8 and Fp2 to measure frontal brain activity, which is attributed to concentration and attention states or the lack of them. For the EOG sensor, 2 electrodes were used in a dipolar setting and placed above and under the left eye (only 2 eletrodes were available, hence only one axis could be measured). Both the EOG and EEG sensors have a common reference which is an additional electrode that was placed in the left earlobe.

4.2.5 Test Subject

A total of 6 participants were selected for this experiment. All of them are aged between 20 to 25 years old. Participants were instructed to remain calm and relaxed, avoiding movements except those absolutely necessary in order to prevent motion artifacts. Even though this experimental setup does include EOG electrodes, it is suggested to the participants to avoid blinking in a subtle manner. This may prevent excessive EOG artifacts accompanied by muscle artifacts. It is advised against stressing this suggestion because it may involuntarily affect mental workload brain waves as the participants actively focus on not blinking, contaminating task data. The latter situation is highly undesirable so the suggestion must be made so as not to seem very important to the participant.

4.3 Data Analysis

The data output of the Unity software is an amalgam of several elements of information: biosignals, time and several event markers. To be able to extract useful information from these data, a number of steps of data processing must be done. The following sections provide a detailed view of what features we wanted to extract from the data and how the extraction was performed.

The data analysis performed in this work was performed identically for each electrode. As there were only two electrodes, there were insufficient data to create brain activity maps. Averaging the data

from the two electrodes was a possible alternative, but it was discarded as it would result in a loss of information and little gain in insight as the individual analysis of only two electrodes is not cumbersome and more information is retained.

4.3.1 Selected Features

The following features were selected for their ability to measure task difficulty and concentration despite not being biosignals:

- **Score**: This feature indicates the number of incorrect answers given by the subjects. The higher the score, the higher the number of incorrect N-back answers. The python script calculates this feature in the data analysis;

- **Latency**: This feature indicates the time in seconds that the subject takes to provide an answer to the N-back letter. The Unity script calculates this feature in real time. Each time that a subject presses the ”up” arrow key, independently of the answer being correct or incorrect, Unity records the reaction time by subtracting the keyboard press time, by the time in which the last letter stimulus appeared.

Based on the information collected in Table 1, the following EEG features have been selected for analysis:

- **Absolute power**: This feature represents the total power of the measured signal, which in this work was divided by frequency bands. The python script calculates this feature in the data analysis;

- **Peak frequency**: This feature represents the frequency in Hertz which has the highest absolute power value within a given frequency band. The python script calculates this feature in the data analysis.

4.3.2 EEG Analysis

The processing of the EEG data entails many steps in order to obtain significant practical information. This processing will be done by means of a python script, which will follow a data pipeline structure, comprised of the following key elements: data import and cleaning; epoch labelling; data filtering; score and latency analysis; multitaper baseline extraction; multitaper signal analysis and baseline normalization; feature extraction and plotting; report generation. These elements will be detailed in the following subsections and schematized in figure 4.10.

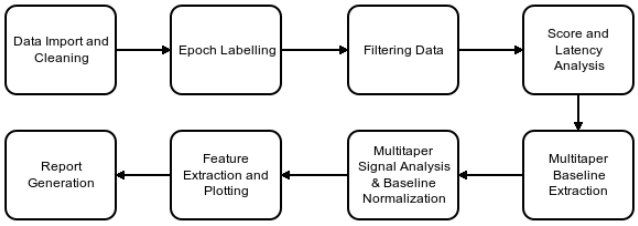

Figure 4.10: Diagram of the data processing pipeline applied in a Python script.

Data Import and Cleaning

The first stage of the data pipeline is importing and cleaning the data. Unity produces a comma-separated values (CSV) file which has to be assigned to proper python data structures, with correct headers and having trimmed off excess data acquired before and after the test itself is performed on the subject.

The output data is organized according to the following table header: trial, type, channel 0, channel 1, channel 2, time, stimulus time, latency; points and finally, dazzle time. Each of the previous data elements is detailed below:

- Trial: Indicates which of the two trials (independent experiment runs) the other data elements are referring to;
- Type: Indicates which phase of the trials the other data elements are referring to, according to a predefined code: different values were attributed to the two baseline periods and N-back levels according to their dazzle/no-dazzle conditions;
- Channel 0, Channel 1 and Channel 2: These data elements correspond to the biosignal data itself - Channel 0 refers to the EOG electrodes data, Channel 1 refers to the AF7-Fp1 electrodes data and Channel 2 refers to the AF8-Fp2 electrodes data;
- Time: Times all the other data elements;
- Stimulus Time: Marks the points in time where the N-back letter stimuli appear;
- Latency: When a subject answers to a letter stimulus, the time he took to answer is recorded in this data element;
- Points: When a subject answers to a letter stimulus, a point is recorded in this data element if the subject gave the wrong answer;
- Dazzle Time: Marks the point in time where a dazzle light stimulus has occurred.

Epoch Labeling

Epochs are fragments of the EEG signal of equal duration which result of the division of the EEG signal. They are the basic elements of the EEG analysis.

In the epoch labeling stage, epochs are classified according to a number that uniquely identifies each epoch. This is done by looping through the input data structure, targeting each stimulus time and attributing it a linearly increasing number. The two big baseline periods are given a negative number to differentiate them from the rest.

Data Filtering

In this work, we are only interested in specific frequencies in the EEG signals. To remove linear trending and artifacts, digital filters must be applied.

The filters used for this work were Finite Impulse Response (FIR) filters with a Kaiser window, with 3627 taps. The FIR filter type was chosen over the Infinite Impulse Response (IIR) type for its Linear phase property, which ensures that signals of all frequencies are delayed by the same amount of time, thereby eliminating the possibility of phase distortion.

High-pass filtering of 4 Hz was applied to the signal, removing DC noise and blink artifacts. The filter response can be found in Figure 4.11.

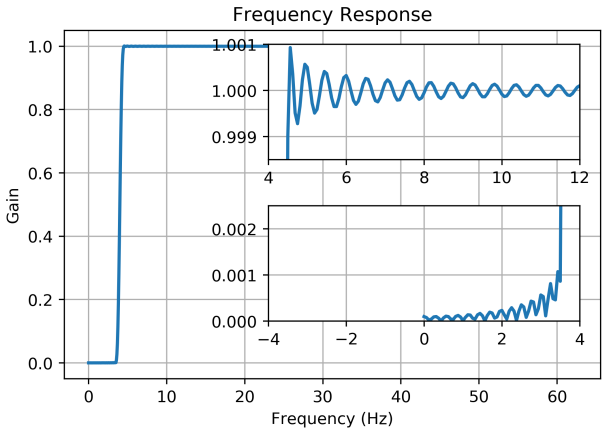

Figure 4.11: Plot of the 4 Hz high-pass FIR filter frequency response with 3627 taps and a Kaiser window.

Low-pass filtering of 31 Hz was applied to the signal, removing all frequencies above the beta frequency band, which were not required for this work as they are regarded as not having important information for concentration assessment. Furthermore, several muscle artifacts can be found in the higher frequencies, further increasing the need to remove these frequencies from the signal. The filter response can be found in Figure 4.12.

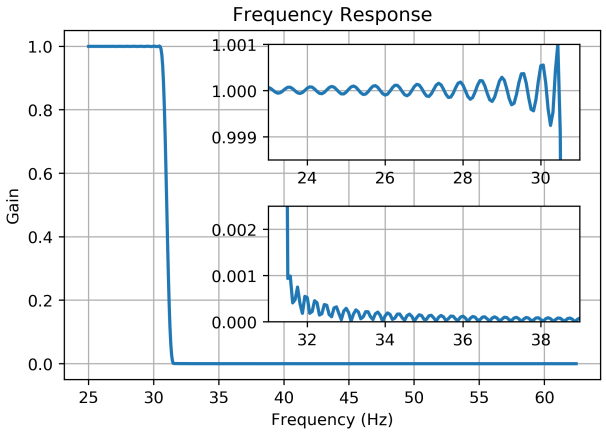

Figure 4.12: Plot of the 31 Hz low-pass FIR filter frequency response with 3627 taps and a Kaiser window.

It is important to note that by applying these filters, especially the high-pass filter, we are removing the blink artifacts from the signal, removing the need to use information from the EOG signal. Although the use of this electrode was initially included in the experiment design to remove the blink artifacts, it was discarded in place of digital filtering. The reason for this was that due to the small number of electrodes, ICA techniques could not be properly applied and linear subtraction techniques required a training period which was not included in the design of this experiment. Therefore, digital filtering came as an easy and effective way of removing these artifacts, by sacrificing the delta frequency band of the signal, which was not of great impact, since the available literature mostly linked the activity of this band to sleep brain activity and not to concentration or attention.

Score and Latency Analysis

The score analysis was based on simple addition of score for each condition, based on the previously made labeling. The more wrong answers were given by the subject, the higher the score, being that each wrong answer had the value of 1 point. The resulting score value for each condition was, therefore, the total number of wrong answers.

The latency analysis was performed similarly to the score analysis, being the only difference that the sum of latencies for a given condition was then averaged by all stimuli in a condition in order to obtain an average latency time for each condition.

The score and latency outputs were organized in python dictionaries in a structured way, where they could readily be accessed.

Multitaper Baseline Extraction

The baseline period must reflect activity with little or no task-related processing. A baseline period of a few hundreds of milliseconds prior to trial onset is the most common time period. For this experiment, a period of 500 ms prior to trial onset was chosen.

When calculating the baseline average for the baseline normalization, one could use a condition-specific or a condition-average baseline. The former means that each condition is baseline-normalized independently of all other conditions, and the latter means that the conditions are first averaged together, then the baseline activity is computed, and then that average baseline activity is applied to each condition. There are two advantages to the condition-average baseline. First, the signal-to-noise ratio of the baseline power will increase because more trials will be included in the average. Second, any condition differences in baseline activity will be retained, which is advantageous if you have hypotheses about condition differences in pretrial activity. We chose to perform the condition-average baseline.

Even though two runs of this experiment ran in succession (the dazzle and no-dazzle conditions were both repeated once per participant), due to even one run being exhaustive and the electrode and headband apparatus being cumbersome, we considered doing two condition-average baselines - one for each run. In doing this, we can minimize the effects of tiredness leaking into the signals.

The multitaper function was applied using the function $psd\_array\_multitaper$ of the $mne.time\_frequency$ Python library (with the default parameter of half-window bandwidth of 4 Hz) to each of the corresponding pre-stimulus baseline for all 6 levels contained in a run, in a total of 288 baseline periods. The power spectrum obtained was then summed and averaged for each frequency, yielding an averaged baseline power spectrum. The same process was repeated for the levels contained in the second run.

Multitaper Signal Analysis and Baseline Normalization

The same multitaper python function is now applied to the rest of the signal, that is, a 2000 ms period, starting from the start of the N-back letter stimulus. After being summed and averaged for each N-back level and organized and stored in python dictionaries, we performed decibel (dB) baseline normalization. A Decibel baseline normalization is one of the most commonly used normalization methods in cognitive electrophysiology.

A decibel baseline normalization will be applied to the data so that the resulting plots and values reflect only the change relative to the baseline period. Comparing changes across subjects is much more

reliable than comparing absolute values because the level of brain activity is highly variable among subjects. In order to apply a normalization technique to the data, one must define a baseline period.

The baseline normalization yields important effects on the biosignals regarding signal quality and statistical analysis, which we detail below:[34]

1. Power data is transformed to the same scale, allowing visual and statistical comparisons from different frequency bands, electrodes, conditions, and subjects;
2. Baseline normalization is performed to a pretrial baseline, which has the effect of disentangling task-related time-frequency dynamics from background or task-unrelated dynamics;
3. Resulting data puts all power results into a common and easily numerically interpretable metric;
4. Baseline-normalized power data are normally distributed, allowing the use of parametric statistical analyses. This facilitates quantitative group-level analyses and also facilitates integration with other kinds of data, such as score and latency analysis.

To calculate this normalization, equation (2.1.4) was applied to each of the power spectra of the N-back levels with respect to the respective trial baseline average. The result of this process was power spectra in decibels that reflected a change in brain activity with respect to a pre-stimulus baseline.

Feature Extraction

The final output of the previous data processing steps resulted in several python dictionaries filled with organized information. While at this point the score and latency information are ready for feature extraction, the frequency band power, and peak frequencies information are in dictionaries filled with matrices containing relative power values for each frequency. Therefore, a final transformation has to be applied.

We now have the relative power in dB per frequency per condition. To extract a single value that assesses the overall relative power for a given frequency band, we perform an integration of the relative power values with respect to the frequencies in a given band. This was done using the scipy.integrate.simps function, which applies the Simpson's rule for numerical integration. The lower and upper limits of the integration were the lower and upper limits of each frequency band. The result was a value that represented the total dB relative power of a given frequency band.

In order to find the peak frequency of a given frequency band, a search for the maximum relative power value was performed for each frequency band using the function numpy.argmax.

These values were again organized in python dictionaries with all features readily accessible for table generation and further analysis.

Report Generation

The final stage of the data pipeline consists in the output of a PDF format report containing the subject name, age, score and latency values, and the selected feature's value tables and plots, to facilitate further analysis. This report is made by previously defining an HTML template to be filled with python variables.

Python Libraries

To be able to effectively analyze the data with python, a number of libraries were used. A list of the most important libraries is provided below:

- pandas, numpy, and matplotlib: These libraries provided access to mathematical functions, data processing functions, data structures, and plotting functions;
- scipy.signal and scipy.integrate: These libraries provided access to signal processing functions such as filter functions and mathematical integration functions;
- mne.time_frequency: MNE is an open-source python library for exploring, visualizing, and analyzing human neurophysiological data. It provided access to multitaper functions;
- jinja2 and xhtml2pdf: Libraries required to be able to create an HTML template, embed it with data, and export as PDF format.

4.3.3 Statistical Analysis

Statistical analysis enables one to infer information of a given population from a given sample of the same population. The two main branches of statistical tests are called parametric and non-parametric tests. The first assumes that the sample's population is approximately normal or Gaussian, while the latter doesn't rely on assumptions that the data are drawn from a given probability distribution. If the assumptions are met, a parametric test will be more powerful than a non-parametric test (the power of a statistical test is the probability that the test will reject the null hypothesis when the alternative hypothesis is true), although non-parametric tests are usually more robust than parametric tests (a larger sample size can be required to draw conclusions with the same degree of confidence) [69].

By performing the baseline normalization, normality is assumed.

Three types of statistical tests will be used in this work: ANOVA, Tukey's HSD (honestly significant difference) and the T-Test.

Most of the hypotheses that will be tested are based on the comparison of samples of three groups: 0-back, 1-back, and 2-back. This type of testing is meant to provide insights as to the effect of the progressive mental workload on the selected biosignal features. For this purpose, a combination of ANOVA and Tukey's HSD tests will be used.

ANOVA stands for "analysis of variance" and it accounts for the rapidly expanding degrees of freedom in a sample as variables are added. Ultimately, it provides a statistical test of whether the population means of several groups are equal, and therefore generalizes the t-test for more than two groups. It is conceptually similar to multiple two-sample t-tests, but is more conservative, resulting in fewer type I errors (when the null hypothesis is true but is rejected). Thus, ANOVA allows one to infer if there are differences between the means, but it doesn't clarify which of the groups differ and how they differ (e.g.: which group has the biggest mean). If a given ANOVA test yields a significant result for the difference in the group's means, it is still necessary to determine which specific groups differed. To tackle this problem, a *post-hoc* test is used, the Tukey's HSD test. The Tukey test is a *post-hoc* test in that the comparisons between variables are made after the data has already been collected. This differs from an *a priori* test, in which these comparisons are made in advance. This test compares the differences between the means of values rather than comparing pairs of values and relies on the output of the ANOVA test. Tukey's test is equivalent to multiple t-tests, but it corrects for family-wise errors as it pools the estimate from all groups, yielding information as to which pair of variables differ in their means and how they differ.

Besides testing for the effect of the progressive mental workload, it is necessary to test for the differences between the Dazzle/No-Dazzle conditions. This will be achieved by applying a standard t-test, directly comparing pairs of the same difficulty of the N-back test for the two conditions.

The t-test is commonly used when directly comparing the means of two groups of samples that follow a normal distribution, allowing to determine if the two means differ or one is bigger (or smaller) than the other.

All of the statistical analysis was performed with the R programming language.

4.4 Summary

In this chapter, we presented the detailed aspects of the experiment performed in this work. A system composed of a computer running the Unity software, which was scripted to be able to receive data from the BITalino board and communicate via serial USB with the Arduino whilst presenting the cognitive task, the N-back task. The Arduino, in turn, controlled a battery-powered LED light which triggered the dazzle stimuli.

We also presented details regarding the data analysis processes that will be used in this work, such as the data pipeline structure written in a Python script and the subsequent statistical analysis using the R programming language.

In the next chapter, we will present the results obtained from this experiment.

Chapter 5

Results and Discussion

5.1 Introduction

A total of 6 subjects consented to participate in the study with a mean age of 22.5 +/- 0.85 years old. ANOVA was used to analyze differences within conditions (e.g.: 3 levels of N-back within a No Dazzle condition) and, if significative, a Tukey's test was performed to identify the levels which contributed to the difference. Between conditions, standard t-tests were applied to compare each N-back level (e.g.: 0-back No Dazzle to 0-back Dazzle conditions).

During the course of the experiment, all participants experienced discomfort both for the duration of the experiment (30 minutes) and the headband, which was considered to be painful to wear. Thus, there might be contamination in the EEG signal due to the participant's discomfort.

5.2 Score and Latency

Table 5.1 displays the results for the ANOVA test for score and latency. All conditions showed a highly significant difference within the conditions ($p < 0.001$). By applying Tukey's test, it was further shown that these differences were highly significant for the 0-back/2-back and the 1-back/2-back pairs. The 0-back/1-back pair difference was not significant. Score in this context means the number of incorrect answers. The higher the score, the higher the number of incorrect answers. This confirms the results present in the literature where the number of correct answers decreases and the latency times increase with increasingly higher N-back levels.

Table 5.2 summarizes the t-tests performed for each level and condition. There were no statistically significant differences between the levels for no-dazzle/dazzle conditions, both for score and latency.

Table 5.1: Table summarizing the ANOVA results for the task's score and latency results. One asterisk (*) means $p \leq 0.05$, two asterisks (**) means $p \leq 0.01$ and three asterisks (***) mean $p \leq 0.001$.

p-values		Score		Latency	
		No Dazzle	Dazzle	No Dazzle	Dazzle
	ANOVA	3.11×10^{-5}***	2.16×10^{-5} ***	4.87×10^{-4}***	3.61×10^{-4}***
Tukey	0-Back — 1-Back	6.25×10^{-1}	1.50×10^{-1}	9.98×10^{-1}	9.77×10^{-1}
	0-Back — 2-Back	4.63×10^{-5}***	1.54×10^{-5}***	2.03×10^{-3}***	3.54×10^{-3}***
	1-Back — 2-Back	6.82×10^{-4}***	3.67×10^{-3}***	1.37×10^{-3}***	8.40×10^{-4}***

Figure 5.1 shows the box-and-whiskers plots obtained from Table 5.2. From visual inspection, although there are no statistical differences between conditions, it can be seen that the average score (wrong answers) increases with increasing N-back levels and that this effect is more predominant in the score than the latency times.

Table 5.2: Table summarizing the T-Test results for the N-back task's score and latency.

p-values		Score			Latency		
		No Dazzle (Mean)	Dazzle (Mean)	p-value	No Dazzle (Mean)	Dazzle (Mean)	p-value
No Dazzle/Dazzle	0-Back	0.750	0.583	0.628	0.497	0.512	0.433
	1-Back	1.33	2.08	0.25	0.495	0.524	0.234
	2-Back	3.92	4.83	0.37	0.611	0.722	0.069

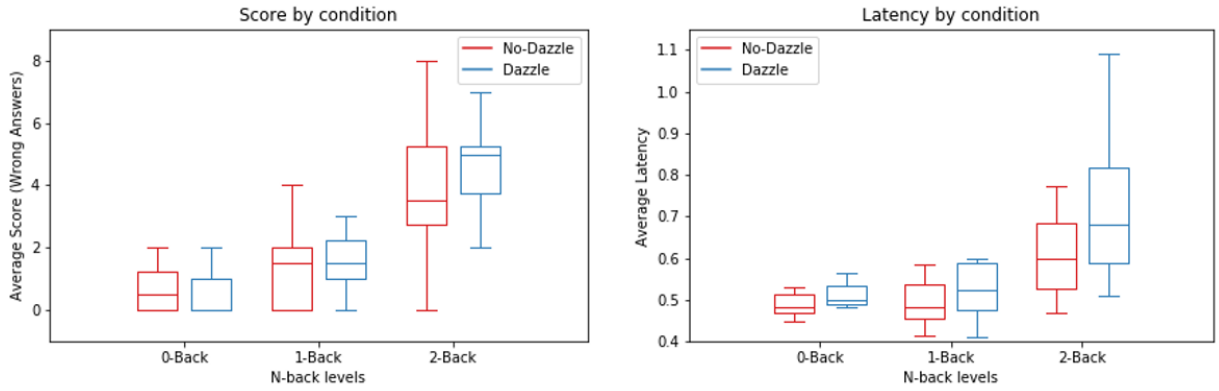

Figure 5.1: Score and latency bar plots comparing the no-dazzle/dazzle conditions by different levels.

5.3 Peak frequency and Absolute Power

5.3.1 Peak Frequency

From table 5.3 and table 5.5 no statistically significant differences can be drawn. Even though there are several sources in the literature that indicate that the alpha peak frequency tends to increase with the increasing difficulty of the N-back levels, we could not find such evidence in the results. One possible culprit could be the low quality wire connection of the electrodes to the sensors. This could have rendered the peak frequency inaccurate. The small number of subjects may, in turn, have greatly increased the uncertainty associated with the statistical tests. Another possibility is that by using adapted electrodes with the BITalino's sensors, the data quality could have been impaired, ultimately producing highly impaired frequency band spetra, which even not having greatly affected the absolute power, could have led to the appearance of several and highly varying peaks.

5.3.2 Absolute Power

The results obtained for the absolute power of the frequency bands went through a baseline normalization process. As the resulting power spectra reflect differences in brain activity regarding a baseline activity,

Table 5.3: Table summarizing the ANOVA results for the frequency band's peak frequency (Hz).

p-values		Peak Alpha		Peak Beta		Peak Theta	
		No Dazzle	Dazzle	No Dazzle	Dazzle	No Dazzle	Dazzle
Electrode 1	ANOVA	0.852	0.652	0.826	0.792	0.823	0.603
Tukey	0-Back	-	-	-	-	-	-
	1-Back	-	-	-	-	-	-
	2-Back	-	-	-	-	-	-
Electrode 2	ANOVA	0.510	0.386	0.963	0.826	0.604	0.839
Tukey	0-Back	-	-	-	-	-	-
	1-Back	-	-	-	-	-	-
	2-Back	-	-	-	-	-	-

Table 5.4: Table summarizing the ANOVA results for the frequency band's absolute power (dB).

p-values		Alpha AP		Beta AP		Theta AP	
		No Dazzle	Dazzle	No Dazzle	Dazzle	No Dazzle	Dazzle
Electrode 1	ANOVA	0.582	0.135	0.910	0.736	0.861	0.718
Tukey	0-Back	-	-	-	-	-	-
	1-Back	-	-	-	-	-	-
	2-Back	-	-	-	-	-	-
Electrode 2	ANOVA	0.926	0.206	0.616	0.120	0.369	0.305
Tukey	0-Back	-	-	-	-	-	-
	1-Back	-	-	-	-	-	-
	2-Back	-	-	-	-	-	-

a positive value indicates therefore that there was an increase in brain activity, a zero value indicated no change in brain activity regarding the baseline and a negative value indicates a decrease in brain activity.

Although no statistically significant differences can be drawn within each condition for every frequency band analyzed (Table 5.4), from Table 5.6 it can be seen that there are highly statistically significant differences between conditions for the 0-back and 1-back levels for the Alpha frequency band with the electrode pair AF7-Fp1. The electrode pair AF8-Fp2 displayed similar results for the 1-back level on the Alpha frequency band and for the Beta frequency band.

Figures 5.2 and 5.3 contain the plots that mirror the information in Table 5.6. These results support the literature evidence that alpha and beta frequency band's absolute power are the strongest EEG features in order to assess concentration. It is well documented that the alpha frequency band absolute power decreases and the beta frequency band absolute power increases with increasing task difficulty. From Figure 5.2 it can readily be seen that the alpha frequency band power increases when a participant is affected by the dazzle effect. This implicates that the participants are generally less concentrated during the dazzle condition, which supports our hypothesis. As for the beta frequency band's absolute power, there was an increase in the 1-back level in the dazzle condition. While appearing contradictory, it might indicate that the dazzle effect, or the knowledge that it might happen, might lead the participants to increase their concentration to compensate for the effect, in the effort to prevail and maintain performance even suffering from the dazzle effect. Another possibility is that it was an anomaly due to low statistical power. Although there are some evidence that the theta frequency band's absolute power increases with increasing task difficulty level, we could not find such evidence.

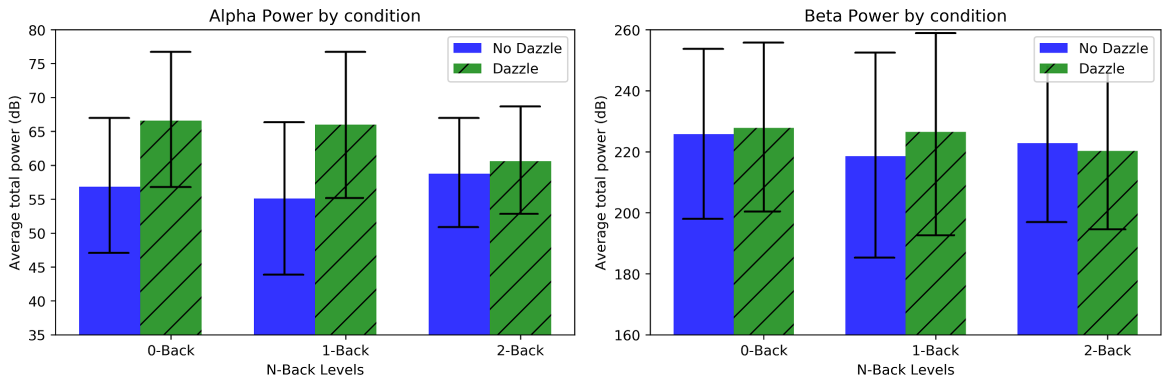

Figure 5.2: Bar plots of the frequency band's absolute power by N-back level and condition: alpha frequency band (left), beta frequency band (right). The error bars correspond to the 95% confidence interval.

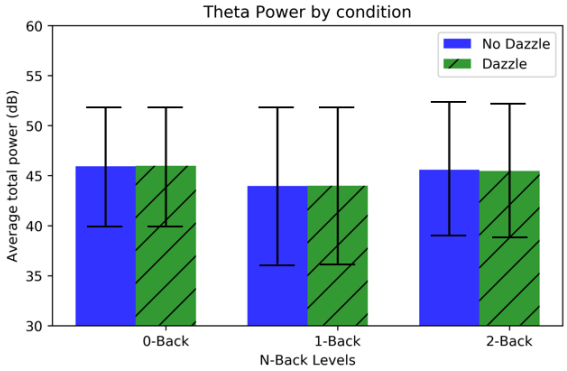

Figure 5.3: Bar plots of the frequency band's absolute power by N-back level and condition for the theta frequency band. The error bars correspond to the 95% confidence interval.

Table 5.5: Table summarizing the T-Test results for the frequency band's peak frequency.

p-values		Peak Alpha			Peak Beta			Peak Theta		
		No Dazzle (Mean)	Dazzle (Mean)	p-value	No Dazzle (Mean)	Dazzle (Mean)	p-value	No Dazzle (Mean)	Dazzle (Mean)	p-value
No Dazzle/Dazzle (Electrode 1)	0-Back	11.08	10.50	0.36	22.50	21.30	0.49	5.96	5.88	0.85
	1-Back	10.70	10.90	0.77	21.20	20.50	0.77	6.08	6.33	0.58
	2-Back	10.80	10.30	0.56	21.90	21.90	1.00	5.79	6.08	0.56
No Dazzle/Dazzle (Electrode 2)	0-Back	10.60	11.20	0.47	21.80	20.90	0.74	6.21	7.88	0.32
	1-Back	10.20	11.60	0.06	21.10	19.60	0.55	6.08	7.13	0.19
	2-Back	11.04	10.70	0.58	21.30	19.50	0.48	5.67	7.08	0.09

Table 5.6: Table summarizing the T-Test results for the frequency band's absolute power. One asterisk (*) means $p \leq 0.05$, two asterisks (**) means $p \leq 0.01$ and three asterisks (***) mean $p \leq 0.001$.

p-values		Alpha AP			Beta AP			Theta AP		
		No Dazzle (Mean)	Dazzle (Mean)	p-value	No Dazzle (Mean)	Dazzle (Mean)	p-value	No Dazzle (Mean)	Dazzle (Mean)	p-value
No Dazzle/Dazzle (Electrode 1)	0-Back	56.80	66.60	0.00***	225.00	228.00	0.87	45.90	46.00	0.99
	1-Back	55.10	66.02	0.00***	219.00	227.00	0.63	44.00	44.00	1.00
	2-Back	58.80	60.60	0.64	223.00	220.00	0.84	45.60	45.50	0.97
No Dazzle/Dazzle (Electrode 2)	0-Back	61.00	68.10	0.15	227.00	232.00	0.76	45.30	44.50	0.83
	1-Back	59.20	71.20	0.01**	212.00	250.00	0.03**	40.80	47.10	0.14
	2-Back	60.50	63.60	0.42	215.00	218.00	0.81	76.40	42.30	0.32

Chapter 6

Conclusion

Concentration while doing a task is temporarily impaired if one is affected by light dazzling. In this thesis, we sought to obtain insights as to the effect of this impairment in concentration at the brain activity level. We proposed the use of a low-cost EEG acquisition setup to investigate whether the dazzle caused by a 10 W white LED affected the performance of a subject in an N-back task.

The N-back test effectively differentiates levels of concentration. Results of score and latency show that difficulty increases highly in the 2-back level, as opposed to the 0-back and 1-back levels, which are similar in difficulty. This suggests that the 0-back level can be left out altogether as it does not add new information for being virtually of the same difficulty of the 1-back level.

Dazzling of a subject with a bright LED light while doing a task that requires concentration (N-back) affects brain activity in a significant manner. Results show that the absolute band power of the alpha and beta frequency ranges are affected by dazzle. This effect was more pronounced in the 0-back and 1-back levels and was expressed as a significant increase in power. Interestingly, the dazzle effect was more pronounced in those levels where concentration was lower.

Both the theta frequency band absolute power and the peak frequencies of all frequency bands were found not to significantly change with dazzling.

A low-cost EEG acquisition set-up can be used to investigate the effect of dazzling on brain activity during tasks that require concentration. With as few as two electrodes, the absolute power of the alpha and beta frequency bands was shown to increase with the dazzle effect. Additionally, the N-back is an effective task to measure concentration, as previous works have shown.

Chapter 7

Pitfalls and Future Work

7.1 Pitfalls and Future Work

Naturally, during the course of this work, points of improvement were readily identified. In the event that the object of this work will be further investigated by other colleagues, we will provide recommendations as to where this work could be improved and what courses of investigation could be followed in the future.

7.1.1 Pitfalls

Below are summarized the main points upon which this work was lacking. Improving or removing these problems altogether could greatly improve the quality of the results and subsequent insight into the research topic.

- Low-quality wire connections resulting from the adaption of electrodes not meant to be used with the BITalino's sensors;
- Low sample size resulted in low power statistics;
- Uncomfortable headband might potentially have led to a partial distraction of the participants;
- Subjects were uncomfortable with the extended experiment time of 30 minutes;
- The use of the EOG electrode proved cumbersome and unnecessary for this setup.

Besides the aforementioned pitfalls, there is an aspect that although not being a pitfall *per se*, it made part of the process of the research for this work, despite ending up not being used. In the initial experimental stages of this work, different types of concentration task were tested, namely a car driving game and a drone flying simulator on the Steam gaming platform. The former was a task based on a simple two-dimensional car driving game, where a subject would have to avoid certain objects. The subject would be dazzled during the course of the game and it was expected that it would impair concentration. However, the task proved to be too complex, having many variables that could affect the measurement of clean EEG signals and at the same time it lacked a proven track of scientific research to back up the effectiveness in measuring concentration. For this reason, this task was abandoned and the work focused solely on the N-back task. The latter was a game based on the simulation of drone racing. The subject would play the game with a gaming pad and would be subject to dazzling during the course of the game. This task was discarded as well because it had the same issues that the two-dimensional car

driving game had, plus the difficulty of integrating the gaming pad input with both the Unity software and the gaming interface on Steam. The need for a simple task with easily controlled variables and easy integration, as well as a proven scientific background which could corroborate our results, lead to the selection and use of the N-back task.

7.1.2 Future Work

Much can be done in researching this subject as this work showed that even with low-cost equipment, one can achieve significant results, albeit not numerous in this work, due to previously mentioned restrictions. Although the focus of future research should be the absolute power of the alpha and beta frequency bands, more features can be researched if one could have access more equipment. Due to having a low amount of electrodes, for instance, it wasn't possible to perform a proper investigation of asymmetric brain activity. The following are other proposed aspects that could bring important insights into the effect of light dazzling on brain activity:

- Repeat the experiment with a different source of light - a stronger LED or even a laser, there is a considerable body of research that indicated that the source of light has great influence on the dazzle effect;
- Test for different angles of incident light when applying the dazzle stimuli;
- Use different concentration tests, such as the number span or even esports games;
- Automation of data processing which could be used in combination with artificial intelligence to detect or predict dazzle according to brain activity.

Automation

Automation of the data analyses performed in this work could be done without a significant effort, which opens the possibility for both lines of research and other applications, such as:

- Offline standard automation for ease of processing and research to simplify the amount of time spent;
- A graphical tool could be produced for this kind of low-cost solutions;
- Online automation of processing could be used for dazzle detection applications, such as a system that aircraft cabins could have to detect when a pilot is dazzled.

7.2 Summary

In this section we described, in hindsight, what could be avoided during the course of this work and what could be improved. Furthermore, we also provided ideas as to future courses of action.

Bibliography

[1] Ho-Sang Lee, Jung-Yong Kim, Murali Subramaniyam, Sangho Park, and Seung-Nam Min. Evaluation of quantitative glare technique based on the analysis of bio-signals. *Ergonomics*, 0139 (April 2017):1–8, 2016. ISSN 0014-0139. doi: 10.1080/00140139.2016.1251620.

[2] Tariq M. Aslam, David Haider, and Ian J. Murray. Principles of disability glare measurement: An ophthalmological perspective. *Acta Ophthalmologica Scandinavica*, 85(4):354–360, 2007. ISSN 13953907. doi: 10.1111/j.1600-0420.2006.00860.x.

[3] Johannes J. Vos. On the cause of disability glare and its dependence on glare angle, age and ocular pigmentation. *Clinical and Experimental Optometry*, 86(6):363–370, 2003. ISSN 08164622. doi: 10.1111/j.1444-0938.2003.tb03080.x.

[4] R Emdad, K Belkic, T Theorell, S Cizinsky, C Savic, and K Olsson. Psychophysiologic sensitization to headlight glare among professional drivers with and without cardiovascular disease., 1998. ISSN 1076-8998.

[5] Ove Steinvall, Stig Sandberg, Ulf Hörberg, Rolf Persson, Folke Berglund, Kjell Karslsson, Johan Öhgren, Zhaohua Yu, and Per Söderberg. Laser dazzling impacts on car driver performance. *Proceedings of SPIE - The International Society for Optical Engineering*, 8898:88980H, 2013. ISSN 0277786X. doi: 10.1117/12.2028505.

[6] M. A. Mainster and G. T. Timberlake. Why HID headlights bother older drivers. *The British journal of ophthalmology*, 87(1):113–7, 2003. ISSN 0007-1161. doi: 10.1136/bjo.87.1.113.

[7] Alexander Toet and Johan W. A. M. Alferdinck. Effects of high power illuminators on vision through windscreens and driving behavior. 8898:1–18, 2013. ISSN 0277786X. doi: 10.1117/12.2028224.

[8] Alexander Toet. Optical countermeasures against human operators. *Proc. SPIE*, 9251(May): 92510G, 2014. ISSN 1996756X. doi: 10.1117/12.2066125.

[9] V. B. Nakagarawa, R. W. Montgomery, A. E. Dillard, L. N. McLin, and C. W. Connor. The effects of laser illumination on operational and visual performance of pilots during final approach. *Proceedings of the 75 th AsMA Annual Scientific Meeting: Frontiers in Aerospace Medicine*, na (June 2004):1–11, 2004.

[10] Jan Theeuwes, Johan W. A. M. Alferdinck, and Michael Perel. Relation Between Glare and Driving Performance. *Human factors*, 44(1):95–107, 2002. ISSN 00187208. doi: 10.1518/0018720024494775.

[11] Van B. Nakagawara, Kathryn J. Wood, and Ron W. Montgomery. Natural sunlight and its association to civil aviation accidents. *Optometry - Journal of the American Optometric Association*, 75(8):517–522, 2004. ISSN 15291839. doi: 10.1016/S1529-1839(04)70177-3.

[12] Federal Aviation Administration (USA). Evaluation of Glare as a Hazard for General Aviation Pilots on Final Approach. (July 2015), 2015. URL `https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2010s/media/201512.pdf`.

[13] J. Kruger and D. Dunning. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, 77 (6):1121–34, 1999. ISSN 0022-3514. doi: 10.1037/0022-3514.77.6.1121.

[14] Chris Berka, Daniel J. Levendowski, Michelle N. Lumicao, Alan Yau, Gene Davis, Vladimir T. Zivkovic, Richard E. Olmstead, Patrice D. Tremoulet, and Patrick L. Craven. EEG Correlates of Task Engagement and Mental\nWorkload in Vigilance, Learning, and Memory Tasks. *Aviation, Space, and Environmental Medicine*, 78(5):B231–B244, 2007. ISSN 0095-6562. doi: 10.1016/j.biopsycho.2011.03.003.

[15] Anne-Marie Brouwer, Maarten A. Hogervorst, Jan B. F. van Erp, Tobias Heffelaar, Patrick H. Zimmerman, and Robert Oostenveld. Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of Neural Engineering*, 9(April 2016):1–14, 2012. ISSN 1741-2560. doi: 10.1088/1741-2560/9/4/045008.

[16] Susanne M. Jaeggi, Martin Buschkuehl, Walter J. Perrig, and Beat Meier. The concurrent validity of the N -back task as a working memory measure. *Memory*, 18(4):394–412, 2010. ISSN 0965-8211. doi: 10.1080/09658211003702171.

[17] Mirka Pesonen, Heikki Hämäläinen, and Christina M. Krause. Brain oscillatory 4-30??Hz responses during a visual n-back memory task with varying memory load. *Brain Research*, 1138 (1):171–177, 2007. ISSN 00068993. doi: 10.1016/j.brainres.2006.12.076.

[18] Christopher D. Wickens. Multiple Resources and Mental Workload. *Human Factors*, 50(3): 449–455, 2008. ISSN 0018-7208. doi: 10.1518/001872008X288394.

[19] Dave Lamble, Tatu Kauranen, Matti Laakso, and Heikki Summala. Cognitive load and detection thresholds in car following situations: Safety implications for using mobile (cellular) telephones while driving. *Accident Analysis and Prevention*, 31(6):617–623, 1999. ISSN 00014575. doi: 10.1016/S0001-4575(99)00018-4.

[20] Yongchang Li, Xiaowei Li, Martyn Ratcliffe, Li Liu, Yanbing Qi, and Quanying Liu. A real-time EEG-based BCI system for attention recognition in ubiquitous environment. *Ubicomp '11 The 2011 ACM Conference on Ubiquitous Computing*, pages 33–40, 2011. doi: 10.1145/2030092.2030099.

[21] Marieke K. van Vugt, Per B. Sederberg, and Michael J. Kahana. Comparison of spectral analysis methods for characterizing brain oscillations. *Journal of Neuroscience Methods*, 162(1-2):49–63, 2007. ISSN 01650270. doi: 10.1016/j.jneumeth.2006.12.004.

[22] Michal Teplan. Fundamentals of EEG measurement. *Measurement Science Review*, 2(2):1–11, 2002. ISSN 15353893. doi: 10.1021/pr070350l.

[23] Wolfgang Klimesch. EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, 29(2-3):169–195, 1999. ISSN 01650173. doi: 10.1016/S0165-0173(98)00056-3.

[24] G. Pfurtscheller and F. H. Lopes. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, 110:1842–1857, 1999. ISSN 1388-2457. doi: 10.1016/S1388-2457(99)00141-8.

[25] Roser Sala-Llonch, Cleofé Peña-Gómez, Eider M. Arenaza-Urquijo, Dídac Vidal-Piñeiro, Nuria Bargalló, Carme Junqué, and David Bartrés-Faz. Brain connectivity during resting state and subsequent working memory task predicts behavioural performance. *Cortex*, 48(9):1187–1196, 2012. ISSN 00109452. doi: 10.1016/j.cortex.2011.07.006.

[26] Allen Waziri, Jan Claassen, R. Morgan Stuart, Hiba Arif, J. Michael Schmidt, Stephan A. Mayer, Neeraj Badjatia, Lewis L. Kull, E. Sander Connolly, Ronald G. Emerson, and Lawrence J. Hirsch. Intracortical electroencephalography in acute brain injury. *Annals of Neurology*, 66(3):366–377, 2009. ISSN 03645134. doi: 10.1002/ana.21721.

[27] Nathan E. Crone, Alon Sinai, and Anna Korzeniewska. High-frequency gamma oscillations and human brain mapping with electrocorticography. *Event-Related Dynamics of Brain Oscillations*, Volume 159(06):275–295, 2006. ISSN 0079-6123. doi: http://dx.doi.org/10.1016/S0079-6123(06)59019-3.

[28] A. Searle and L. Kirkup. A direct comparison of wet, dry and insulating bioelectric recording electrodes. *Physiological measurement*, 21(2):271–83, 2000. ISSN 0967-3334. doi: 10.1088/0967-3334/21/2/307.

[29] Thomas C. Ferree, Phan Luu, Gerald S. Russell, and Don M. Tucker. Scalp electrode impedance, infection risk, and EEG data quality. *Clinical Neurophysiology*, 112(3):536–544, 2001. ISSN 13882457. doi: 10.1016/S1388-2457(00)00533-2.

[30] G. H. Klem, H. O. Lüders, H. H. Jasper, and C Elger. The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. *Electroencephalography and clinical neurophysiology. Supplement*, 52(3):3–6, 1999. ISSN 0424-8155. doi: http://dx.doi.org/10.1016/0013-4694(58)90053-1.

[31] Robert Oostenveld and Peter Praamstra. The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, 112(4):713–719, 2001. ISSN 13882457. doi: 10.1016/S1388-2457(00)00527-7.

[32] J. N. Mak and J. R. Wolpaw. Clinical Applications of Brain-Computer Interfaces: Current State and Future Prospects. *IEEE reviews in biomedical engineering*, 2:187–199, 2009. ISSN 19373333. doi: 10.1109/RBME.2009.2035356.Clinical.

[33] Ning Han Liu, Cheng Yu Chiang, and Hsuan Chin Chu. Recognizing the degree of human attention using EEG signals from mobile sensors. *Sensors (Switzerland)*, 13(8):10273–10286, 2013. ISSN 14248220. doi: 10.3390/s130810273.

[34] Mike X. Cohen. *Analyzing Neural Time Series Data: Theory and Practice*, pages 3–239;447–515. 2014. ISBN 0262019876. doi: 10.1017/CBO9781107415324.004.

[35] Ernst Niedermeyer and F. H. Lopes Da Silva. *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, volume 1. 2004. ISBN 0781751268.

[36] Steven J. Luck. An Introduction to the Event-Related Potential Technique. *Monographs of the Society for Research in Child Development*, 78(3):388, 2005. ISSN 1540-5834. doi: 10.1118/1.4736938.

[37] David Friedman and R. A. Y. Johnson. Event-Related Potential (ERP) Studies of Memory Encoding and Retrieval: A Selective Review. *Microscopy Research and Technique*, 28(January):6–28, 2000. doi: 10.1.1.473.8095.

[38] J. Kalcher and G. Pfurtscheller. Discrimination between phase-locked and non-phase-locked event-related EEG activity. *Electroencephalography and Clinical Neurophysiology*, 94(5):381–384, 1995. ISSN 00134694. doi: 10.1016/0013-4694(95)00040-6.

[39] M. Akin. Comparison of wavelet transform and FFT methods in the analysis of EEG signals. *Journal of Medical Systems*, 26(3):241–247, 2002. ISSN 01485598. doi: 10.1023/A:1015075101937.

[40] Gernot Florian and Gert Pfurtscheller. Dynamic spectral analysis of event-related EEG data. *Electroencephalography and Clinical Neurophysiology*, 95(5):393–396, 1995. ISSN 00134694. doi: 10.1016/0013-4694(95)00198-8.

[41] Michael J. Prerau, Ritchie E. Brown, Matt T. Bianchi, Jeffrey M. Ellenbogen, and Patrick L. Purdon. Sleep Neurophysiological Dynamics Through the Lens of Multitaper Spectral Analysis. *Physiology*, 2017. ISSN 1548-9213. doi: 10.1152/physiol.00062.2015.

[42] M. B. I. Raez, M. S. Hussain, and F. Mohd-Yasin. Techniques of EMG signal analysis: detection, processing, classification and applications. *Biological procedures online*, 8(1):11–35, 2006. ISSN 1480-9222. doi: 10.1251/bpo115.

[43] Avinash Tandle. Classification of Artefacts in EEG Signal Recordings and Overview of Removing Techniques. *International Journal of Computer Applications*, 2015. ISSN 23944714. doi: 10.5120/cae2016651997.

[44] Suresh D. Muthukumaraswamy. High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations. *Frontiers in Human Neuroscience*, 2013. ISSN 1662-5161. doi: 10.3389/fnhum.2013.00138.

[45] T. J. T. P. Van Den Berg. On the relation between glare and straylight. *Documenta Ophthalmologica*, 78(3-4):177–181, 1991. ISSN 00124486. doi: 10.1007/BF00165678.

[46] John Bullough, Zengwei Fu, and John. Van Derlofske. Discomfort and disability glare from halogen and HID headlamp systems. *SAE Technical Paper Series*, 2002-01-00(724):0, 2002. doi: 10.4271/2002-01-0010.

[47] Craig A. Williamson and Leon N. McLin. Nominal ocular dazzle distance (NODD). *Applied Optics*, 54(7):1564–1572, 2015. ISSN 0003-6935. doi: 10.1364/AO.54.001564.

[48] Craig A Williamson. Simple computer visualization of laser eye dazzle. *Journal of Laser Applications*, 28(1):012003, 2016. ISSN 1938-1387. doi: 10.2351/1.4932620.

[49] Rodolfo Repetto. An analytical model of the dynamics of the liquefied vitreous induced by saccadic eye movements. *Meccanica*, 41(1):101–117, 2006. ISSN 00256455. doi: 10.1007/s11012-005-0782-5.

[50] Michael J Flannagan. Subjective and objective aspects of headlamp glare: effects of size and spectral power distribution (Report No. UMTRI 99-36. *Ann Arbor: The University of Michigan Transportation Research Institute*, (November):17.

[51] Hans-dieter Reidenbach. Local susceptibility of the retina , formation and duration of afterimages in the case of Class 1 iaser products , and disability glare arising from high-brightness light emitting diodes. *Journal of Laser Applications*, 21(1):46–56, 2009. ISSN 1042346X. doi: 10.2351/1.3071565.

[52] M. Osaka. Peak alpha frequency of EEG during a mental task: task difficulty and hemispheric differences. 21(1):101–105, 1984. ISSN 0048-5772. doi: 10.1111/j.1469-8986.1984.tb02325.x.

[53] Thalía Fernández, Thalía Harmony, Mario Rodríguez, Jorge Bernal, Juan Silva, Alfonso Reyes, and Erzsébet Marosi. EEG activation patterns during the performance of tasks involving different components of mental calculation. *Electroencephalography and Clinical Neurophysiology*, 94(3):175–182, 1995. ISSN 00134694. doi: 10.1016/0013-4694(94)00262-J.

[54] T. Bennett, K. Malia, B. Linton, M. Raymod, and K. Brewick. Rehabilitation of attention and concentration deficits following brain injury. *Journal of Cognitive Rehabilitation*, March/Apra (16(2)):8–13, 1998.

[55] Maher Chaouachi, Imène Jraidi, and Claude Frasson. Modeling mental workload using EEG features for intelligent systems. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 6787 LNCS, pages 50–61, 2011. ISBN 9783642223617. doi: 10.1007/978-3-642-22362-4_5.

[56] Nathaniel H. Pulling, Samuel P. Sturgis, Donald R. Vaillancourt, and James J. Dolliver. Headlight Glare Resistance and Driver Age. *HUMAN FACTORS*, 22(1):103–112, 1980. ISSN 0018-7208. doi: 10.1177/001872088002200111.

[57] Stephen J. Anderson and Ian E. Holliday. Night driving: effects of glare from vehicle headlights on motion perception. *Ophthalmic and Physiological Optics*, 15(6):545–551, 1995. ISSN 14751313. doi: 10.1016/0275-5408(95)00070-T.

[58] S. P. Sturgis and D. J. Osgood. Effects of glare and background luminance on visual acuity and contrast sensitivity: implications for driver night vision testing. *Human factors*, 24(3):347–360, 1982. ISSN 00036870. doi: 10.1016/0003-6870(83)90019-4.

[59] Johan Engström, Emma Johansson, and Joakim Östlund. Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(2 SPEC. ISS.):97–120, 2005. ISSN 13698478. doi: 10.1016/j.trf.2005.04.012.

[60] João M. P. Coelho, José Freitas, and Craig A. Williamson. Optical eye simulator for laser dazzle events. *Applied Optics*, 55(9):2240, 2016. ISSN 0003-6935. doi: 10.1364/AO.55.002240.

[61] Eric Klinger, Kenneth C. Gregoire, and Steven G. Barta. Physiological Correlates of Mental Activity: Eye Movements, Alpha, and Heart Rate During Imagining, Suppression, Concentration, Search, and Choice, 1973. ISSN 14698986.

[62] Laurent George, Fabien Lotte, Raquel Viciana Abad, and Anatole Lécuyer. Using scalp electrical biosignals to control an object by concentration and relaxation tasks: Design and evaluation. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pages 6299–6302, 2011. ISSN 1557170X. doi: 10.1109/IEMBS.2011.6091554.

[63] Alan Gevins, Michael E. Smith, Linda McEvoy, and Daphne Yu. High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, 7(4):374–385, 1997. ISSN 10473211. doi: 10.1093/cercor/7.4.374.

[64] Ole Jensen and Claudia D Tesche. Frontal theta activity in humans increases with memory load in a working memory task. *Neuroscience*, 15(8):1395–1399, 2002. ISSN 0953816X. doi: 10.1046/j.1460-9568.2002.01975.x.

[65] Thalía Harmony, Thalía Fernández, Juan Silva, Jorge Bernal, Lourdes Díaz-Comas, Alfonso Reyes, Erzsébet Marosi, Mario Rodríguez, and Miguel Rodríguez. EEG delta activity: An indicator of attention to internal processing during performance of mental tasks. *International Journal of Psychophysiology*, 24(1-2):161–171, 1996. ISSN 01678760. doi: 10.1016/S0167-8760(96)00053-0.

[66] Scott Watter, Gina M. Geffen, and Laurie B. Geffen. The n-back as a dual-task: P300 morphology under divided attention. *Psychophysiology*, 38(6):998–1003, 2001. ISSN 0048-5772. doi: 10.1111/1469-8986.3860998.

[67] Wolfgang Klimesch, Bärbel Schack, and Paul Sauseng. The functional significance of theta and upper alpha oscillations. *Experimental Psychology*, 2005. ISSN 16183169. doi: 10.1027/1618-3169.52.2.99.

[68] C. Richard Clark, Melinda D. Veltmeyer, Rebecca J. Hamilton, Elena Simms, Robert Paul, Daniel Hermens, and Evian Gordon. Spontaneous alpha peak frequency predicts working memory performance across the age span. *International Journal of Psychophysiology*, 53(1):1–9, 2004. ISSN 01678760. doi: 10.1016/j.ijpsycho.2003.12.011.

[69] Marius Marusteri and Vladimir Bacarea. Comparing groups for statistical differences: how to choose the right statistical test? *Biochemia Medica*, 2010. ISSN 18467482. doi: 10.11613/BM.2010.004.