Inês Apolinário Pancada da Silveira

Bachelor of Science in Biomedical Engineering

Validation of fNIRS System as a Technique to Monitor Cognitive Workload

Dissertation submitted in partial fulfillment
of the requirements for the degree of

Master of Science in
Biomedical Engineering

Adviser: Hugo Gamboa, Auxiliar Professor, NOVA University
of Lisbon

Examination Committee

Chair: Prof. Dr. Ricardo Nuno Pereira Verga e Afonso
Vigário
Rapporteur: Prof. Dr. Carla Maria Quintão Pereira
Member: Prof. Dr. Hugo Filipe Silveira Gamboa

FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

September , 2021

Validation of fNIRS System as a Technique to Monitor Cognitive Workload

Para a minha Mãe e para o meu Pai

Acknowledgements

Em primeiro lugar, gostaria de expressar o meu mais profundo agradecimento ao meu orientador, Prof. Dr. Hugo Gamboa, da FCT NOVA, pela oportunidade de participar neste projeto, pelo qual não podia sentir mais apreço. Pela confiança depositada, pela enorme disponibilidade, pela motivação e pela ajuda essencial.

Ao Rui Varandas, da PLUX, não chega agradecer uma e outra vez. Nem mil. O Rui foi o meu maior apoio em todas as horas e não há como o homenagear. Obrigada Rui, do fundo do coração, por teres sido sempre solidário e amigo, pelo teu profissionalismo e inteligência extremos e pela enorme aprendizagem. Não poderia ter tido mais sorte :)

Um grande obrigada aos voluntários, sem os quais este estudo não seria possível! Obrigada pela confiança, disponibilidade e inter-ajuda numa época tão atípica como a que vivemos.

Obrigada à minha familia por me segurarem quando caio. E quando não caio. Pelo apoio incondicional, pela preocupação e pelo respeito. Obrigada à minha Mãe pelo amor, pela amizade e pela luz. Obrigada ao meu Pai pela inspiração, pela força e pela motivação. Obrigada aos meus irmãos Matilde e Afonso, que só por existirem fazem do mundo um sítio melhor.

Obrigada ao meu namorado, Francisco, pela admiração e pelo carinho. Por sempre acreditar em mim e no meu valor.

Obrigada aos meus melhores amigos Freda, Con, Juca, Melo, Du, Chico, Pinho, Rui, Fred, Manel e Hugo por me acompanharem nesta jornada. Por estarem lá e por me encherem o coração.

Obrigada à minha faculdade, FCT NOVA, por estes 5 anos, pelo conhecimento incalculável que me pôs nas mãos, pelas amizades e pelas histórias. E obrigada, também, à Prof. Dr. Carla Quintão, coordenadora do mestrado de Engenharia Biomédica, pelo excelente acompanhamento ao longo destes anos, a preocupação e o carinho.

A vida é feita das amizades, do respeito pelo outro e da bondade. A todos o meu muito obrigada!

"The divine gift does not come from a higher power. But from our own minds." ★

ABSTRACT

Cognitive Workload (CW) is a key factor in the human learning context. Knowing the optimal amount of CW is essential to maximise cognitive performance, emerging as an important variable in e-learning systems and Brain-Computer Interfaces (BCI) applications. Functional Near-Infrared Spectroscopy (fNIRS) has emerged as a promising avenue of brain discovery because of its easy setup and robust results. It is, in fact, along with Electroencephalography (EEG), an encouraging technique in the context of BCI. Brain-Computer Interfaces, by tracking the user's cognitive state, are suitable for educational systems. Thus, this work sought to validate the fNIRS technique for monitoring different CW stages.

For this purpose, we acquired the fNIRS and EEG signals when performing cognitive tasks, which included a progressive increase of difficulty and simulation of the learning process. We also used the breathing sensor and the participants' facial expressions to assess their cognitive status. We found that both visual inspections of fNIRS signals and power spectral analysis of EEG bands are not sufficient for discriminating cognitive states, nor quantify CW. However, by applying machine learning (ML) algorithms, we were able to distinguish these states with mean accuracies of 79.8%, reaching a value of 100% in one specific case. Our findings provide evidence that fNIRS technique has the potential to monitor different levels of CW. Furthermore, our results suggest that this technique allied with the EEG and combined via ML algorithms is a promising tool to be used in the e-learning and BCI fields for its skill to discriminate and characterize cognitive states.

Keywords: cognitive workload, fNIRS, cognitive states, EEG, machine learning

Resumo

O esforço cognitivo (CW) é um factor relevante no contexto da aprendizagem humana. Conhecer a quantidade óptima de CW é essencial para maximizar o desempenho cognitivo, surgindo como uma variável importante em sistemas de e-learning e aplicações de Interfaces Cérebro-Computador (BCI). A Espectroscopia Funcional de Infravermelho Próximo (fNIRS) emergiu como uma via de descoberta do cérebro devido à sua fácil configuração e resultados robustos. É, de facto, juntamente com a Electroencefalografia (EEG), uma técnica encorajadora no contexto de BCI. As interfaces cérebro-computador, ao monitorizar o estado cognitivo do utilizador, são adequadas para sistemas educativos.

Assim, este trabalho procurou validar o sistema de fNIRS como uma técnica de monitorização de CW. Para este efeito, adquirimos os sinais fNIRS e EEG aquando da execução de tarefas cognitivas, que incluiram um aumento progressivo de dificuldade e simulação do processo de aprendizagem. Utilizámos, ainda, o sensor de respiração e as expressões faciais dos participantes para avaliar o seu estado cognitivo. Verificámos que tanto a inspeção visual dos sinais de fNIRS como a análise espectral dos sinais de EEG não são suficientes para discriminar estados cognitivos, nem para quantificar o CW. No entanto, aplicando algoritmos de *machine learning* (ML), fomos capazes de distinguir estes estados com exatidões médias de 79.8%, chegando a atingir o valor de 100% num caso específico. Os nossos resultados fornecem provas da prospecção da técnica fNIRS para supervisionar diferentes níveis de CW. Além disso, os nossos resultados sugerem que esta técnica aliada à de EEG e combinada via algoritmos ML é uma ferramenta promissora a ser utilizada nos campos do e-learning e de BCI, pela sua capacidade de discriminar e caracterizar estados cognitivos.

Palavras-chave: carga de trabalho cognitiva, fNIRS, estados cognitivos, EEG, machine learning

Contents

List of Figures

LIST OF TABLES

Abbreviations

BCI   Brain-Computer Interfaces

CBF   Cerebral Blood Flow
CLT   Cognitive Load Theory
CW   Cognitive Workload

deoxyHb   Deoxyhaemoglobin
DPF   Differential Pathlength Factor

EEG   Electroencephalography
ET   Extra Trees

fNIRS   functional Near-Infrared Spectroscopy

LTM   Long-Term Memory

MAE   Mean Absolute Error
MBLL   Modified Beer-Lambert Law
ML   Machine Learning
MRI   Magnetic Resonance Imaging
MSE   Mean Square Error
MTLI   Modified Task Load Index

NIR   Near-Infrared
NVC   Neurovascular Coupling

oxyHb   Oxyhaemoglobin

PFC   Prefrontal Cortex

RF   Random Forest

STM   Short-Term Memory

TLI   Task Load Index
totalHb   Total Haemoglobin

WM   Working Memory

INTRODUCTION

Cognitive Workload (CW) is an important indicator of mental activity and one of the most important variables in psychology and ergonomics [1, 2]. In fact, people in their professions face increased CW and decreased attention, which may injure labor productivity due to impaired cognitive functioning [3, 4]. Thereby, it appears as a valuable factor concerning safety and productivity issues in critical decision-making applications and high-risk fields [2, 4–6]. In the United States alone, it is estimated that 44 - 98 thousand people die in hospitals every year due to medical errors caused by, among other factors, CW [7, 8].

However, CW is also a very complex concept, with numerous definitions and interpretations [1]. One possible way to define it is the amount of mental effort required to complete a task within a limited time period [9]. Thus, one believes that an increased workload may lead to degraded performance, memory deficits, stress, irritability, mental exhaustion, and impaired learning [10–12].

When it comes to education, the evaluation of CW is often used as a means of improving student learning process [2]. Along with CW level, cognitive states also affect cognitive performance, appearing as a crucial parameter for the learning process understanding [13–16]. In truth, boredom can prevent the completion of a task, while interest might motivate it [13, 14]. Paradoxically, frustration seems to do both [15].

Learning concerns several cognitive processes, such as attention maintenance, motivation, work memory update, and long-term memory integration. Furthermore, it may comprehend an emotional component, possibly monitored by the autonomous nervous system, once it triggers different responses to different emotional conditions [17]. The

presence of these processes involves different cognitive effort levels, wherefore monitoring CW might provide information about the learning process as a whole. In this sense, standard functional techniques for mapping brain activations as Electroencephalography (EEG) and Magnetic Resonance Imaging (MRI) are very complex or expensive and cannot always be applied in real-world settings [18].

A brain-computer interface (BCI) is a communication system that uses brain activity to control computers or prostheses without using peripheral nerves and muscles, appearing as a fundamental tool for people suffering from severe motor disabilities [19, 20]. Nowadays, the focus of BCI research has been extended to other non-medical applications, namely, educational systems [21, 22]. It adapts the interface in accordance with the cognitive state of the user [18]. And as far as BCI concerns, we have seen an increase in research aiming to objectively estimate mental workload, through EEG [20], or functional Near-Infrared Spectroscopy (fNIRS) [18, 19, 23].

Functional Near-Infrared Spectroscopy is one promising and ongoing approach for the study of brain activity, as a non-invasive and movement tolerable brain scanner [9, 24–27]. It appears as a complement or substitute to the conventional techniques (EEG or MRI) once it reduces the complexity and cost and provides sufficient temporal and spatial resolution [28]. Functional Near-Infrared Spectroscopy is an imaging modality measuring the haemodynamic processes in the brain. This technique consists in the emission of two beams of different frequencies radiation, being the absorption profile for each beam different for different tissues [28]. Specifically, oxygenated (oxyHb) and deoxygenated (deoxyHb) haemoglobin profiles are different, enabling the detection of chromophore relative concentrations fluctuations, which gives indirect information about the activation of monitored zone due to neurovascular coupling [28]. For evaluating the effort involved in cognitive tasks, the prefrontal cortex (PFC) is often assessed as it comprises executive functions required for the cognitive processes that mental workload is grounded on [18, 29–32]. Given its advantages, fNIRS appears as an auspicious ally to the brain exploration. Actually, many relevant research trends predict that interest in fNIRS-BCI will continue to grow, giving room to communication, motor restoration and neuro-rehabilitation research [19].

Based on the above statements, in the present study, we sought to investigate fNIRS technique applicability for monitoring different states of cognitive workload with a small number of channels. In addition, it aims at (i) developing an appropriate methodology to stimulate different cognitive workload levels, (ii) performing data acquisition in a simulated environment, (iii) studying the relative cognitive concentrations estimation and (iv) identifying the features that enable differentiating between cognitive levels.

1.1 Outline of the thesis

This thesis is divided into 8 chapters. The present one serves to introduce the research question that motivates this work, its relevance, and the means to answer it.

Chapter 2 is about the human central nervous system. Here the biologically relevant theoretical concepts for this thesis are presented, namely, the anatomy and physiology of the brain and the underlying learning mechanisms.

Chapter 3 reveals the physical principles of fNIRS optical spectroscopy, explains the theory of neurovascular coupling and the haemodynamic response that underpin it, and finally, the advantages that make it of interest to the research world.

Chapter 4 presents the literature review of the modalities that allow CW monitoring, with a special focus on the fNIRS technique.

Chapter 5 includes the materials and methods used to assess the cognitive workload experienced by the participants in this study. Chapter 6 draws these results together, while chapter 7 discusses them.

Finally, chapter 8 presents the main findings and limitations of this study, as well as future work suggestions.

A chapter of supplementary material, containing some results not presented in the results chapter, can also be found at the end of the document.

2

The Human Central Nervous System

Our nervous system functions as a very powerful integrator of information, coming from the nerves and sensory organs, orchestrating the motor responses to be given by the body [33]. In fact, it comprises billions of neurons, operating as a complex network that enables an organism to communicate with its environment [33, 34]. This network encompasses sensory neurons, sensible to stimuli in the internal or external environment, and motor neurons, responsible for movement in general [34]. Neurons are cells with a very defined structure, which is represented in the Figure 2.1. Signals entering these neurons originate at synapses that occur on neuronal dendrites and in the cell body. On the other hand, the outgoing signals travel via a single axon with multiple branches, flowing from the neuron into different parts of the nervous system [33]. This forward direction (from the axon of a preceding neuron to dendrites on cell membranes of subsequent neurons), present in the majority of synapses, forces the signal to travel in the necessary direction for performing specific nervous functions [33].

The nervous system is composed of the central nervous system (CNS), which includes the brain and spinal cord, and the peripheral nervous system (PNS), which includes sensory receptors, sensory nerves, and ganglia outside the CNS. The CNS and PNS are in constant communication [34]. Sensory receptors, in turn, are indispensable in most activities, whether visual, auditory, or tactile. The sensory experience coming from these activities can originate an immediate reaction from the brain, or the memory of the experience can be stored in the brain for a long time and determine bodily responses in the future [33]. This sensory information enters the CNS via the peripheral nerves and is conducted to multiple sensory areas: (1) the spinal cord at all levels; (2) the reticular substance of the medulla, pons, and mesencephalon of the brain; (3) the cerebellum; (4) the thalamus; and (5) the areas of the cerebral cortex (see Figure 2.2) [33].

The head role of the nervous system is to regulate various bodily activities through

its motor functions. These consist of controlling the contraction of skeletal muscles throughout the body, contraction of smooth muscle in the internal organs, and secretion of active chemicals by both exocrine and endocrine agencies in various parts of the body [33].
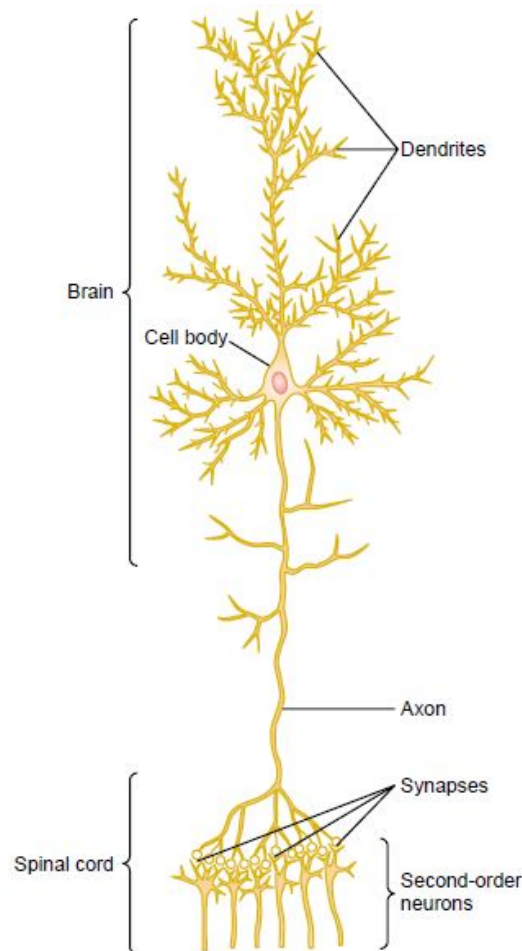
Figure 2.1: **Schematic of a neuron in the brain.** Reprinted from [33].

2.1 The Brain

Despite being the largest portion of the nervous system, we know very little about cerebral cortex and its functions. However, we do know the effects of damaging or specific stimulating in multiple regions of the cortex [33]. The functional part of the cerebral cortex is a 2 to 5 millimeters layer of neurons covering the surface of all the convolutions of the cerebrum. This layer comprehends a total area of about one-quarter of a square meter, containing around 100 billion neurons [33]. The cerebral cortex acts as a memory store. It is in constant association with the lower centers of the nervous system. Otherwise, the functions of these lower centers would be imprecise [33].

In turn, the cerebral hemispheres are composed of the cerebral cortex, subjacent white matter, and the deep nuclei, which are the basal ganglia, hippocampus, and amygdala.
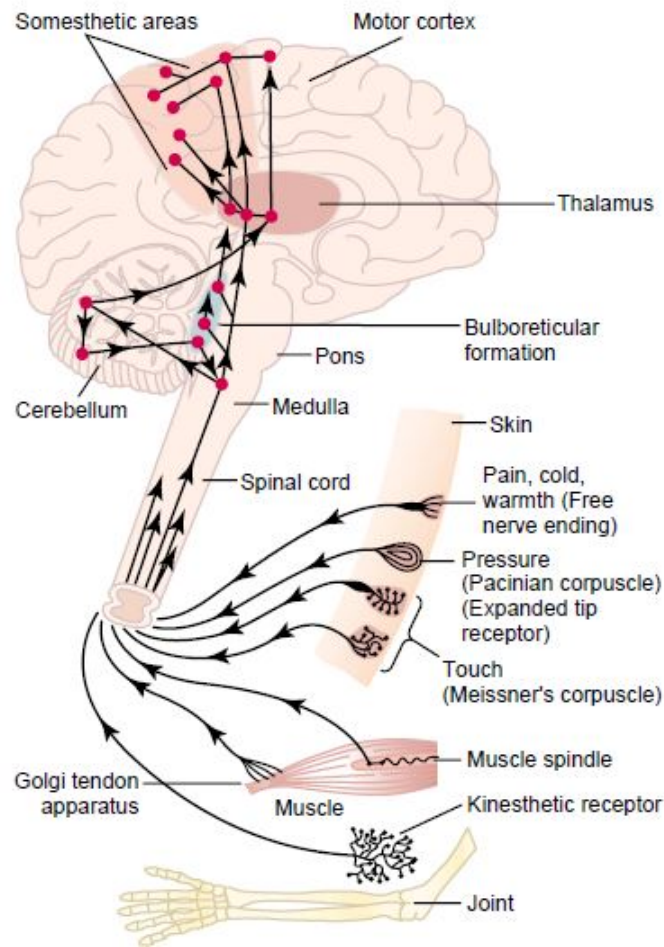
Figure 2.2: **Somatosensory axis of the nervous system.** Reprinted from [33].

Some of its functions are perception, higher motor functions, cognition, motivation, memory and emotion regulation (see Figure 2.3) [34].

Cerebral cortex. The cerebral cortex is the convoluted surface of the cerebral hemispheres and is divided into four lobes - frontal, parietal, temporal, and occipital - separated by sulci. As mentioned earlier, the cerebral cortex receives and processes sensory information and integrates motor functions. The sensory and motor areas of the cortex can also be designated as primary, secondary or tertiary, depending on their degree of sensory or motor processing. Thus, primary areas are those that require the simplest processing and therefore involve the fewest synapses, while tertiary areas require the most complex processing and the greatest number of synapses [34]. In addition to this, primary motor areas, linked to a specific muscle, are responsible for discrete muscle movements. Primary sensory areas (including primary visual cortex, primary auditory cortex, and primary somatosensory cortex), on the other hand, detect visual, auditory, or somatic stimuli and transmit that information directly to the brain from peripheral sensory organs. Finally, secondary and tertiary sensory, connected to association areas, are involved in complex processing [34]. For example, secondary sensory areas analyze

7
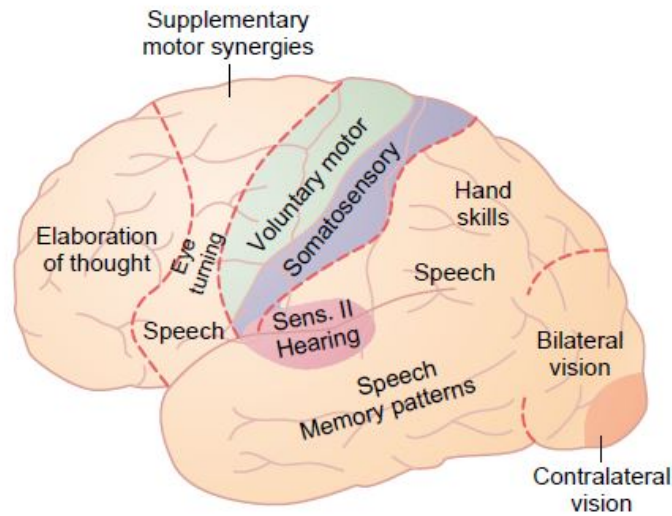
Figure 2.3: **Functional areas of the human cerebral cortex.** Reprinted from [33].

the shape and color of an object or the significance of sound tones, while tertiary areas relates to memory, abstract thinking and behaviour [33].

Association areas. The association areas integrate the information they receive simultaneously from the motor and sensory cortices for specific actions execution (see Figure 2.4). The main ones are (1) the parieto-occipitotemporal association area, (2) the prefrontal association area and (3) the limbic association area, whose explanations are given below [33, 34].
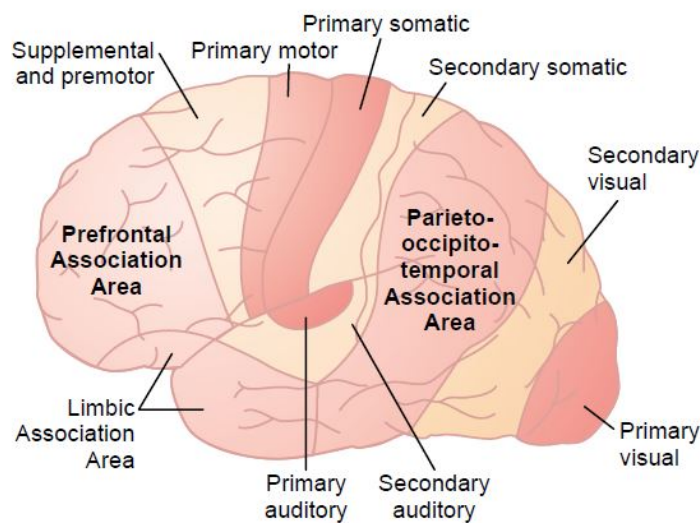
Figure 2.4: **Map of major association areas of the cerebral cortex, as well as primary and secondary motor and sensory areas.** Reprinted from [33].

- **Parieto-occipitotemporal Association Area.** It is located in the large parietal and occipital cortical space bounded by the anterior somatosensory cortex, the posterior visual cortex, and the auditory cortex laterally, being at the basis of the interpretation of signals from all surrounding sensory areas. Its functional sub-areas are shown in Figure 2.4 [33].

  1. Analysis of the Spatial Coordinates of the Body. An area that begins in the posterior parietal cortex and extends to the superior occipital cortex. It receives simultaneous visual and somatosensory information, providing continuous analysis of the spatial coordinates of all parts of the body and the surrounding environment. Figure 2.5 shows this area [33].

  2. Area for Language Comprehension. The Wernicke area is the main area for language comprehension, located behind the primary auditory cortex, in the posterior cortex part of the superior temporal lobe gyrus. A fundamental region of higher intellect since most intellectual functions are language-based. Figure 2.5 shows this area [33].

  3. Area for Initial Processing of Visual Language (Reading). This area, also known as the angular gyrus, is located posteriorly to the language comprehension area, situated mainly in the anterolateral region of the occipital lobe. It is an area of the visual association that provides the visual information transmitted by reading into Wernicke's area, of language comprehension. In its absence, a person cannot get language comprehension through reading, even if they have it through hearing. Figure 2.5 shows this area [33].

  4. Area for Naming Objects. In the most lateral portions of the anterior occipital lobe and posterior temporal lobe is located the area for naming objects. The learning of names, usually through auditory input, is essential for the understanding of auditory and visual language (functions performed in Wernicke's area). On the other hand, learning the physical nature of objects is mainly done through visual input. Figure 2.5 shows this area [33].

- **Prefrontal Association Area.** The prefrontal association area and the motor cortex work in close association in movements planning. This area is in charge of thought development and short-term working memories storage. Besides, it is capable of motor and non-motor thoughts. Figure 2.4 shows this area [33].

  Broca's Area. Broca's area, located partly in the posterior lateral prefrontal cortex and partly in the premotor area is the area of word formation, where individual words or short sentences are initiated and executed (see Figure 2.5). This region works in close association with Wernicke's area [33].

- **Limbic Association Area.** Figures 2.4 and 2.5 show, lastly, the limbic association area. This area is located in the anterior pole of the temporal lobe, in the ventral

9

portion of the frontal lobe, and in the cingulate gyrus, located at the bottom of the longitudinal cleft on the medial surface of each cerebral hemisphere and is closely related to motivation in the learning process [33].
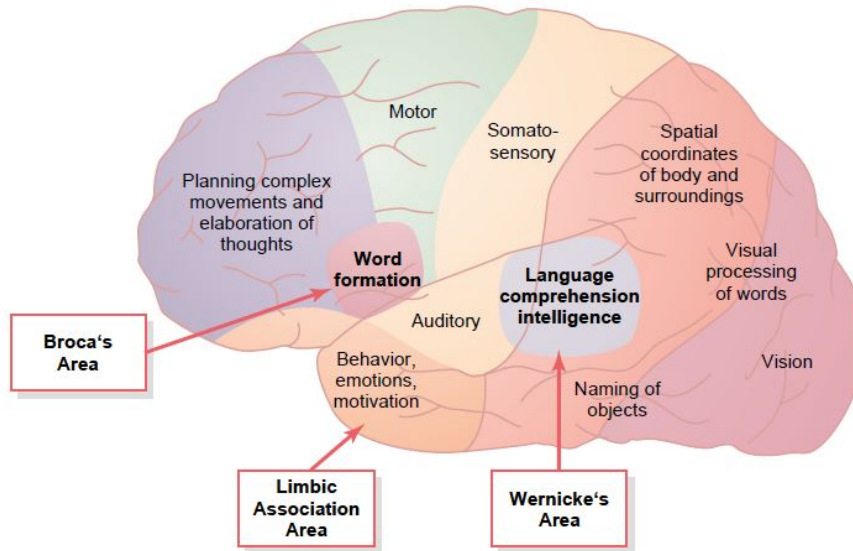
Figure 2.5: **Locations of specific cerebral cortex functional areas, such as Wernicke's and Broca's areas for language comprehension and speech production.** Reprinted from [33].

- **Basal ganglia, hippocampus, and amygdala.** The basal ganglia, constituted by the caudate nucleus, the putamen, and the globus pallidus, receive information from all areas of the cerebral cortex and its projections through the thalamus to the frontal cortex help in movement regulation. The limbic system consists of the hippocampus, involved in memory processes and the amygdala, involved in emotions [34].

The cognitive tasks assessed in this study, the n-back task and mental subtraction, activate the PFC [19, 35, 36]. And therefore, this is the brain area of interest in this work. Prefrontal cortex-related activities are also an appropriate choice for fNIRS-BCI since they involve fewer motion artefacts and signal attenuation due to hair [19].

2.1.1 The Brain Waves

The level of excitation and activation within different brain areas determine the intensity of electrical activity and its patterns [33]. In addition, EEG power reflects the number of neurons that discharge synchronously. In the same way that intelligence correlates to brain size, one may suppose that EEG power reflects cortical information processing capacity [37]. Brain waves, shown in Figure 2.6, are electrical impulses in the brain. All brainwaves are produced by synchronised electrical impulses, due to the neurons' communication with each other. Their intensities range from 0 to 200 microvolts and can

occur with specific frequencies. Brain waves are almost always irregular and do not have a specific pattern, but sometimes they appear in a form characteristic of specific brain disorders, such as epilepsy. These EEG waves are called alpha, beta, delta, gamma and theta, as we can see in Figure 2.6 [11, 33, 37].
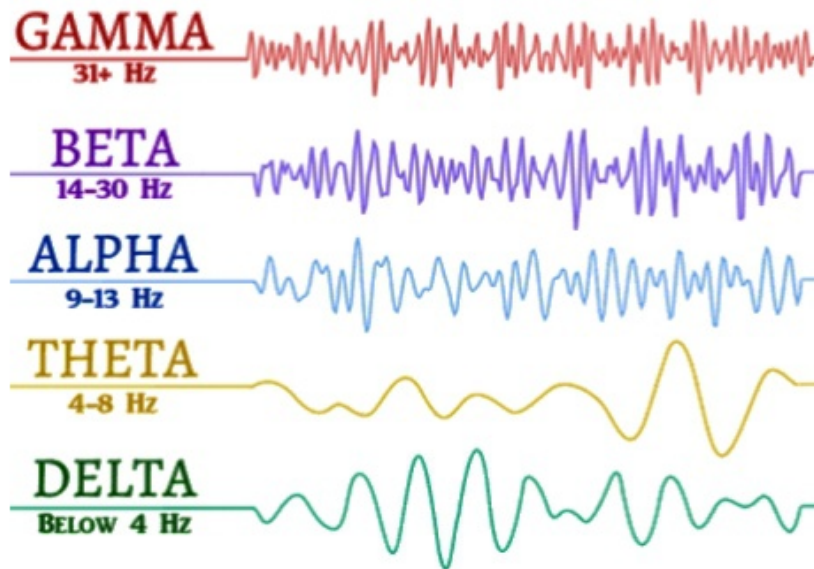
Figure 2.6: **Diagram of the five brain waves.** Reprinted from [38].

- **Delta waves (0.1-3 Hz).** More prominent at the temporal lobe, indicates deep sleep, and less focus and attention maintenance [11, 37, 39–41]. In fact, individuals diagnosed with attention disorder deficit show higher delta activity when trying to focus. Delta waves tend to be the highest in amplitude and the slowest waves [37]. They are also related to memory consolidation, concentration, motivation and focused attention [42]. They seem to increase during continuous attention tasks [11].

- **Theta waves (4-8 Hz).** Normally occur in the parietal and temporal regions during emotional stress in some adults, particularly during disappointment and frustration. These waves are related to slow activity, fantasizing and meditation [11, 37, 39, 40]. They are also related to selective attention, arousal, episodic memory and mental workload [41–44]. Theta spectral power is believed to increase accordingly to task difficulty and WM. It also seems to intensify in activities requiring sustained attention. In addition, an increased theta power may be translated into lower alertness, and it appears indicative of loss of cortical arousal [41, 43, 44].

- **Alpha waves (8-13 Hz)**. Rhythmical waves that occur most intensely in the occipital region but can also be recorded from the parietal and frontal regions of the scalp. These waves are related to relaxation, reflection and resting states and are one of the brain's most important frequency for learning once they reflect task engagement, speed of working memory and cognitive performance [11, 33, 37, 39–42, 44]. An

11

increase in alpha power is related to lower mental vigilance and alertness and, therefore, a decrease in the attention resources allocated to the task. When task difficulty increases, a suppression in alpha band is expected [44].

- **Beta waves (13-30 Hz).** Are recorded mainly from the parietal and frontal regions [33]. Beta activity means alertness, waking consciousness, frustration, anxiety, engagement and working state [11, 33, 39, 41]. These waves are especially notorious when a person is thinking or receiving sensory stimulation and they are widely used in cognitive workload assessment [11, 40, 41]. They are also related to visual memory and working memory allocation [40–42]. Higher cognitive demand and concentration, in visual WM tasks, for example, increase beta power, especially on the parieto-occipital cortex area [40, 41, 44].

- **Gamma waves (above 30 Hz).** More intense at centro-midline of the brain, are related to high mental activity, memory and learning. A good 40 Hz activity is associated with good memory, whereas its lack of activity create learning disabilities [11, 39]. Besides, it is also linked to selective and focused attention and motor task execution [40–42].

Spectral power within these frequency bandwidths study can be used for cognitive effort assessment. More specifically, previous works reported that a more challenging task originates frontal theta synchrony and frontal alpha desynchrony [37, 43]. It is well established that increased recruitment of the relevant cognitive-motor resources for task performance reflects an alpha attenuation. In addition, various studies have alleged that theta synchrony in the frontal cortical regions represents the recruitment of attentional resources for concentration and working memory (WM) [37, 40, 41, 43, 45]. Beta and Gamma variations have also been associated with cognitive processes, such as cortical activation with task increasing difficulty [40, 41, 43]. Nevertheless, these two were shown to be less representative indices of CW, compared to alpha and theta, and so, they were considered less important in this study [37, 43, 45].

2.1.2 The Learning Brain

The main barrier to understanding learning and memory lies in the fact that the neuronal mechanisms of thought are unknown, and little is known either about the mechanisms of memory. However, the destruction of large portions of the cerebral cortex does not prevent a person from having thoughts but reduces their depth and also the degree of awareness of the surroundings [33].

The PFC is the area of thought elaboration. The ability of the prefrontal areas to keep track of many pieces of information simultaneously and recalling this information instantaneously, when needed for later thoughts, is called the brain's WM. In fact, studies defend that the prefrontal intelligence areas are divided into different segments for storing different types of temporary memory, such as one area for storing the shape and

form of an object or part of the body and another for storing movement. By combining all these temporary parts of WM, we have the ability to plan the future or solve complicated tasks [33]. Besides, the PFC is related to executive functions. These functions, including inhibition, cognitive shifting, and WM are higher-level cognitive process for the attainment of specific goal [33, 46].

Learning and memory are higher-order functions of the nervous system. Learning is the mechanism that changes someone's knowledge, values, behaviour and world view based on their experiences [47]. Memory, on the other hand, stores what is learned. Learning can be non-associative or associative. In non-associative learning, i.e. learning by habituation, a repeated stimulus causes a response, but that response's intensity gradually diminishes as it is learned. In associative learning, on the other hand, there is the phenomenon of sensitisation, where a stimulus results in a greater likelihood of a response when it is learned that the stimulus is important. Synaptic plasticity is the fundamental mechanism on which learning is based. In this mechanism, synaptic function (called synaptic strength) and effectiveness depend on the previous level of synaptic traffic [34]. The memory may be of a short-term (working) type and long-term (storage) type, as it will be explained in Chapter 4 [2].

Attention factor is closely related to learning performance. According to [41], in 1983, [48] defined attention as a psychological process comprised of focus and concentration, which enhances cognition speed and accuracy. Also study [49], in 1958, alleged that learning without attention prevents identification, effective learning and memory. In other words, learning is ineffective when a learner disregards the learning content, explaining why instructors should improve learning quality by stressing learner attention and providing effective strategies [41]. Attention types affecting learning performance include sustained, focused, selective, shifting, spatial and divided attention [41]. However, the type that seems to be more correlated with learning performance is sustained attention [41]. Attention and WM share parts of the same cerebral regions. Working memory load is reflected by the activation of frontal and right parietal cerebral areas, that is, synchronization in the theta band and desynchronization in the alpha band. Similarly, attention has been related mainly to the activation of frontal and parietal areas, that is, a desynchronization of the alpha band and theta band synchronization [45, 50].

Other variable fundamental in the learning process scope is CW. Cognitive Workload was defined by Hart and Staveland, in 1988, as a relationship between mental processing capabilities and the performed task demands [12]. Various studies state that a higher CW level might break performance, resulting in cognitive task failure [2, 12]. Thus, since our brains have limited capacity to process and hold information, CW assessment is relevant to determine learning content efficiency and for those who study BCI [2]. In 1974, Dolce and Waldeier found that complex mental tasks, this is, arithmetic tasks increased delta power. Besides, an increase in theta activity may be linked to task difficulty and emotional factors. Finally, higher WM loads are thought to reflect an increase in theta and low beta bands powers in frontal midline regions [11].

At last, recent literature has been arguing the significance of emotions in learning environments due to its effect on academic performance and achievement [51, 52]. In 2008, Poggi described emotions as "multifaceted internal states, encompassing feelings and cognitive, physiological, expressive, and motivational aspects, that are triggered whenever an individual's goal is achieved" [51]. One may consider four types of emotions: positive activating emotions (e.g., enjoyment), negative activating emotions (e.g., anxiety), positive deactivating emotions (e.g., relaxation) and negative deactivating emotions (e.g., boredom) [52]. However, some studies advise research on emotions in e-learning to focus on negative activating emotions, such as anxiety, annoyance and frustration and positive activating emotions, e.g. enjoyment and curiosity [53–55].

3

FUNCTIONAL NEAR-INFRARED SPECTROSCOPY

Neurophotonics is an increasingly popular field [28]. Brain activity can be measured through changes in the optical properties of brain tissues and as such, optical technology has been paving the way in the study of the brain [56]. Functional Near-Infrared Spectroscopy is an emergive optical noninvasive neuroimaging technique that allows the quantification of concentration changes of oxyhaemoglobin (oxyHb) and deoxyhaemoglobin (deoxyHb) non-invasily in the brain, following neuronal activation [57, 58]. In 1977, Fransis Jobsis discovered that near-infrared light (NIR) could be used in our bodies when, holding a steak against visible light, he observed that red light could penetrate 4 mm thick bone [58]. This seemed to indicate that NIR light, could pass through the skull and reach the underlying tissues. Functional Near-Infrared Spectroscopy technology takes advantage of the property of skin and bone being transparent to NIR light, and is increasingly being used in a wide variety of physiological fields and applications [58].

3.1   Functional Near-Infrared Spectroscopy Principles

Functional Near-Infrared Spectroscopy is an emerging optical neuroimaging technique underlying the theory of neurovascular coupling (NVC) and optical spectroscopy [57]. Figure 3.1 resumes these two fNIRS principles. Functional Near-Infrared Spectroscopy measurements are a result of NIR light with different wavelengths transmission by a source onto the scalp. Source and detector are placed over the scalp surface to detect the change in optical density caused by the haemodynamic changes in neuronal tissue (specifically, in cortical grey matter). Therefore, light needs to travel through many extracranial and intracranial layers (the scalp skin, skull, cerebrospinal fluid), with different optical properties. Despite the very complex interaction of the NIR light with the human tissue, as the tissue is anisotropic and inhomogeneous through the different layers, this can be

simplified considering that inside the tissue the light is attenuated by absorption and scattering [28, 57, 58]. Figure 3.2 is a schematic representation of NIR photons' travel between source and detector, through the different layers of the head.
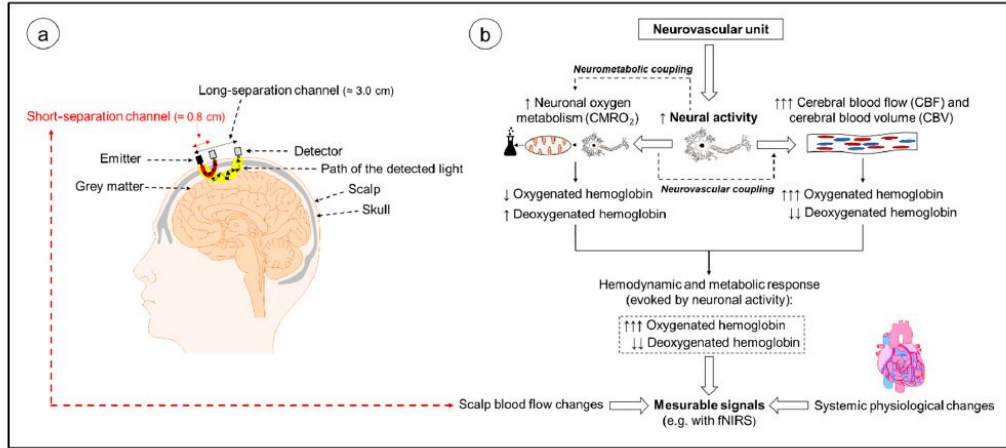
Figure 3.1: **(a) Representation of a NIRS setup on the human head and the assumed banana-shaped course of detected light. (b) Schematic illustration of the neurovascular unit and the changes in cerebral hemodynamics and oxygenation induced by neural activity.** $CMRO_2$ stands for cerebral metabolic rate of oxygen; $\uparrow$ stands for increase; $\downarrow$ stands for decrease. Reprinted from [57].
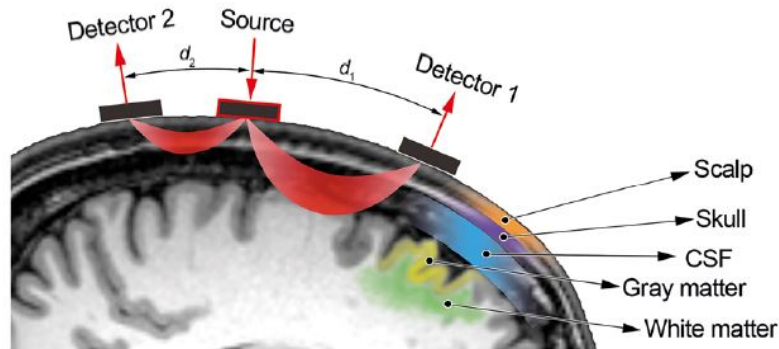
Figure 3.2: **Schematic representation of NIR photons' travel between source and detector, through the different layers of the head.** The penetration depth of the light is proportional to the source-detector distance (*d1* represents the deeper channel and *d2* the superficial channel). A channel is "composed by the pair source-detector and is located at the midpoint between the source and the detector and a depth of around the half of the source-detector separation". Reprinted from [58].

Haemoglobin, the oxygen-carrying part of blood, is the key absorbing chromophore for the NIR spectrum and very physiological dependent. Light in the NIR zone of the electromagnetic spectrum (620 nm - 1000 nm) disperses through most biological tissues like bones and skin and haemoglobin absorbs NIR light. It may appear in its oxygenated

form (oxyHb) and its deoxygenated form (deoxyHb), depending on its saturation. For oxyHb its NIR light absorption is higher for $\lambda > 800$ nm and for deoxyHb is higher for $\lambda < 800$ nm. This difference may be found in blood's colour: oxygenated blood is more red and venous blood more purple [58]. While crossing biological tissue, the NIR light is also scattered besides being absorbed. Scattering is 100 times more frequent than absorption and leads to light attenuation. The more a photon is scattered, the longer is the travelled path and the greater is the probability of being absorbed. The light shined into the head is scattered, diffused, and able to penetrate several centimeters through the tissue [57, 58]. Figure 3.3 is a schematic illustration of light propagation through the multiple neuronal tissues. It presents possible paths of the photons, some of them with the scattering phenomenon. Thus, if we position a detector at the head's surface, we can estimate changes in light attenuation by measuring the non-absorbed components of the scattered light (see Figure 3.3). In turn, these changes in attenuation appear, at a given wavelength, as a linear combination of the concentration changes of oxyHb and deoxyHb since the absorption in the NIR zone is mainly due to these two chromophores. These principles of fNIRS are summarized in Figure 3.3. Functional Near-Infrared Spectroscopy devices can be of three types: time-domain, frequency-domain and continuous wave. Time-domain fNIRS emits a very short pulse of light into the tissue and measures the arrival times of photons emerging from the tissue. This technique is the most complex of the three, producing the most information [59]; frequency-domain fNIRS modulates the intensity of the emitted light, then measures the intensity of the detected light as well as the phase shift, corresponding to the time of flight [59]; continuous-wave fNIRS uses an emitting light at a constant intensity, measuring only changes in the intensity of the light that has passed through the tissue [59]. Most commercially available fNIRS systems are continuous wave and so this is what we will focus on. These devices measure light attenuation (A) by estimating the ratio of injected light (I_{IN}) to output light (I_{OUT}) (see Figure 3.3). The changes in attenuation (ΔA) are estimated and used to derive the changes in oxyHb and deoxyHb concentrations. Thus, ΔA depends only on changes in absorption by the haemoglobin chromophores and not on other factors such as melanin or water concentration. This method is known as modified Beer-Lambert's law (see Figure 3.3) [57, 58].

3.1.1 Modified Beer-Lambert Law

The Modified Beer-Lambert Law (MBLL), represented in Equation 3.2, is an extension of the Beer–Lambert law introducing a scattering dependent light intensity loss parameter, here denoted by G. The law describes the loss of light intensity (I) in tissue (referred as optical density (OD) or attenuation (A), unitless) as a function of the chromophore concentrations (c,[M]), molar extinction coefficients (ε, [$M^{-1} cm^{-1}$]) differential path length factor (DPF, unitless; accounts for the increased distance the light travels due to the reduced scattering coefficient μ_s'), source–detector separation (d, [cm]) and G
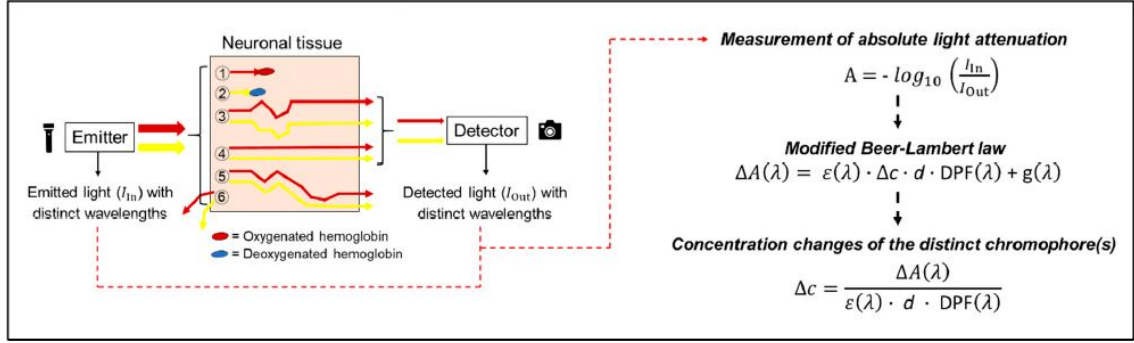
Figure 3.3: **Schematic illustration of light propagation through the neuronal tissue.**
The left side of the figure shows possible photon paths for different wavelengths (the red
color symbolize wavelengths of $\lambda > 800$ nm, mainly absorbed by oxyHb — see Photon 1),
while yellow color symbolize wavelengths of $\lambda < 800$ nm (mainly absorbed by deoxyHb
— see Photon 2). Path 3 is a representation of a photon suffering some scattering events
before reach the detector. Path 4 stands for a ballistic photon. Path 5 shoes a photon not
detected due to scattering events (lost due to forward scattering). And, finally, path 6
represents a photon lost due to backward scattering. The formulas to calculate concentra-
tion changes in chromophores, which relates to continuous-wave NIRS, are presented on
the right side of the figure. The symbols means: *A:* light attenuation, or *ΔA(λ):* changes
in light attenuation at a certain wavelength (λ); I_{IN}: intensity of emitted light; I_{OUT}:
intensity of recorded light; $\varepsilon(\lambda)$: the extinction coefficient of the chromophore at a certain
wavelength *(λ)*; *λc:* changes in chromophore concentration; *d:* distance source-detector;
DPF(λ): differential path length factor DPF for a particular wavelength (λ); g(λ): scat-
tering at a certain wavelength (λ), where g is ignored because only light attenuation is
considered. Reprinted from [57].

(unitless) [59]:

$$OD(t,\lambda) = -\log_{10} \frac{I(t,\lambda)}{I_0(t,\lambda)} = \sum_i \varepsilon_i(\lambda)c_i(t)DPF(\lambda)d + G(\lambda) \qquad (3.1)$$

The index i represents all the chromophores used in this study, oxyHb and deoxyHb,
and I_0 represents the intensity of the emitted light. Please note that Equation 3.1 use the
base 10 logarithm (i.e. $I = I_0 10^{-\varepsilon cd}$) and so use molar extinction coefficients, instead of the
absorption coefficients which are related to the natural logarithm. Both express the level
of absorption per concentration (μM/mM) and per length (cm) but diverge by a scaling
factor of ln(10). Taking into account that the change in scattering is small relative to the
change in absorption, we can presume that G is time-invariant. Thus, it can be ignored
when determining the change in optical density ($\Delta OD(\Delta t,\lambda) = OD(t_1,\lambda) - OD(t_0,\lambda)$) for a
time point t_1 relative to an initial time point t_0. Furthermore, the emitted intensity I_0 is
assumed constant and therefore this term cancels out [59].

$$\Delta OD(\Delta t,\lambda) = -\log_{10} \frac{I(t_1,\lambda)}{I(t_o,\lambda)} = \sum_i \epsilon_i(\lambda)\Delta c_i(t)DPF(\lambda), \qquad (3.2)$$

where $\Delta c_i = c_i(t_1) - c_i(t_0)$ denotes the temporal change in chromophore concentration. The MBLL applies to homogeneous change in oxyHb and deoxyHb in homogeneous tissue. The fact that the head is inhomogeneous is not a concern, once the inhomogeneity remains constant and is mostly covered by the constant G, which cancels out for measuring concentration changes. The fact that the concentration change in oxyHb and deoxyHb is not homogeneous, i.e. it occurs only in the brain and not in other tissues such as skin and skull, leads to an error in quantification, i.e. the MBLL highly underestimates the size of the variations in oxyHb and deoxyHb. Normally this error is rectified by considering partial differential pathlengths, but usually, there is no need for this correction to be performed, once the trend of the signals is correct and quantification is not essential in brain research. It is sufficient to detect the presence of activation and where it is taking place. In this way, we can compare signals in multiple locations and under different stimulation paradigms [59].

The DPF is a dimensionless correction scaling factor that indicates how many time farther than distance d the detected light has traveled [59]. This factor accounts for the increase in optical path length caused by scattering (i.e. the change in direction of motion of a particle due to a collision with another particle) of light in biological tissue, and is multiplied by the source-detector separation to estimate the actual length corresponding to the path that the light has travelled [57]. This factor is considered age, sex and wavelength dependent and variant between subjects and different tissues. Finally, the haemoglobin concentration changes are estimated by solving Equation 3.3 and determining Δc, i.e. Δ[oxyHb] and Δ[deoxyHb]:

$$\begin{bmatrix} \Delta[deoxyHb] \\ \Delta[oxyHb] \end{bmatrix} = (d)^{-1} \begin{bmatrix} \varepsilon_{deoxyHb\lambda_1} \varepsilon_{oxyHb\lambda_1} \\ \varepsilon_{deoxyHb\lambda_2} \varepsilon_{oxyHb\lambda_2} \end{bmatrix}^{-1} \begin{bmatrix} \Delta OD(\Delta t, \lambda_1/DPF(\lambda_1)) \\ \Delta OD(\Delta t, \lambda_2/DPF(\lambda_2)) \end{bmatrix} \qquad (3.3)$$

The totalHb concentration is calculated from the sum of the oxyHb and deoxyHb concentrations.

3.2 Haemodynamic response

In the neurometabolic coupling phenomenon, an increasing neural activity leads to increasing aerobic metabolism, necessary to meet the energetic demands of neural tissue [57, 58]. The neural activity also triggers changes in cerebral haemodynamics that induce cerebral blood flow (CBF) intensification in the active brain areas. This other phenomenon is called neurovascular coupling or functional hyperaemia and is responsible, for example, for the mechanism of change in the diameter of capillaries. As the local oxygen supply is superior to its consumption, there is an increase in oxyHb concentration and a decrease in deoxyHb, being these concentrations estimated based on light attenuation changes, measured by fNIRS [57, 58]. Figure 3.4 B shows the haemodynamic response being measured with fNIRS and Figure 3.4 C shows the different locations of the cerebral cortex that are being measured. Data in Figure 3.4 refer to a memory task,

19

recorded over the PFC. This task resulted in an increase in oxyHb concentration and a decrease in deoxyHb concentration (see Figures 3.4 A and B) in the measured brain region (see Figure 3.4 C), as would be expected [58].
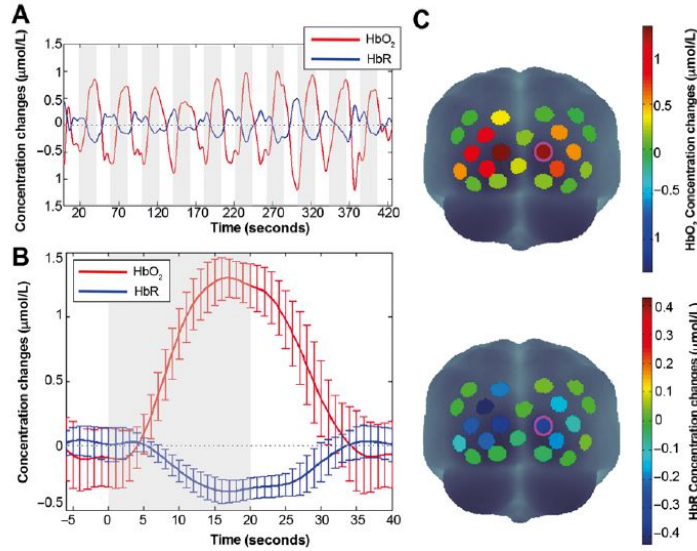
Figure 3.4: **Representation of haemodynamic response uring a block-designed flashing checkerboard experiment.** The gray areas represent the stimulation period. Panel B shows the block-averaged haemodynamic response (mean ± SD) from the channel surrounded with pink line in Panel C, computed by averaging the oxyHb (HbO$_2$, in red) and deoxyHb (HbR, in blue) signals presented in A, across the 10 task blocks. It is characterized by simultaneous oxyHb increase and deoxyHb decrease. Panel C presents the distribution of the maximum block-averaged concentration changes within the gray block shown in Panel B across all the channels, both for oxyHb (top) and deoxyHb (bottom). Reprinted from [58].

A typical haemodynamic response peaks at $\approx$ 6 s after stimulus onset and returns to its baseline $\approx$ 10 s after the stimulus presentation [58]. The characteristics of the response, such as intensity and duration, may vary between different regions of the cerebral cortex and according to the age of the participants [58]. Several studies indicate that there is a metabolic correlation between neural activity measured by fNIRS and the fMRI technique, i.e., the blood oxygenation level-dependent (BOLD) response. Specifically, there is a positive correlation between the BOLD signal and oxyHb and a negative correlation between the BOLD signal and deoxyHb [58].

The study and quantification of CW using fNIRS is an increasingly common practice. Increased workload causes an increase in prefrontal oxyHb, correlated with increased task engagement. However, oxyHb decreases when the cognitive task is too difficult; so does task engagement and performance [60].

3.3 Advantages and disadvantages of fNIRS

The advantages of fNIRS, compared to EEG, MEG, fMRI and PET techniques, make it a great alternative for the study of cognitive performance and cerebral haemodynamics. Table 3.1 indicates some of those. Compared to more conventional neuroimaging techniques, it is a non-invasive tool, with a sufficiently good spatial (1.0 - 3.0 cm) and temporal (usually up to 10 Hz) resolution, portable, with low noise level during operation, low cost, robust against motion artefacts (avoiding strict immobilisation or sedation of the participants), suitable for periods of extended monitoring relative to other neuroimaging techniques and with the possibility of including individuals with metal implants or claustrophobia [57, 60]. Table 3.2 presents the comparison between fNIRS and other conventional techniques. Furthermore, fNIRS allows the quantification of oxyHb and deoxyHb concentrations, unlike fMRI which focuses on the paramagnetic properties of deoxyHb. Thus, it enables the inclusion of other markers such as tissue oxygenation (TOI) and total haemoglobin concentration (totalHb) [57]. However, fNIRS also has disadvantages, namely being limited to cortical layers, being vulnerable to changes in systemic physiology, and not having a standard protocol for data processing [57].

Table 3.1: **Advantages and disadvantages of fNIRS.** Adapted from [58].

Advantages	Disadvantages
Safe	
Better spatial resolution than EEG	
Better temporal resolution than fMRI	Lower temporal resolution than EEG/MEG
Tolerance to motion artifacts	Lower spatial resolution than fMRI
Possibility to monitor oxyHb and deoxyHb	Low penetration depth (1.5-2 cm)
Portability	Impossibility to gather structural images and anatomical information
Low cost	Systemic interferences
Silent	Variable SNR
Availability of miniaturized and wearable systems	Optodes placement can be time consuming in case of hairy regions and a high number of sources and detectors
Suitability for long-time continuous monitoring	
Feasibility for multimodal imaging	Higher susceptability to motion errors and less confort in case of
Compatibility with other electrical and magnetic devices	high-coverage measurements with fiber optics coupled to the head
More participant friendly than fMRI	Lack of standardization in data analysis
All participants are eligible (all ages, no exclusion criteria)	

Table 3.2: **Comparison between fNIRS and other neuroimaging techniques.** Adapted from [58].

	fNIRS	fMRI	EEG/MEG	PET
Signal	oxyHb deoxyHb	BOLD (deoxyHb)	Electromagnetic	Cerebral blood flow Glucose metabolism
Spatial resolution	2-3 cm	0.3 mm voxels	2-3 cm for EEG/ 1-2 cm for MEG	4 mm
Penetration depth	Brain cortex	Whole head	Brain cortex for EEG/ deep structures for MEG	Whole head
Temporal sampling rates	Up to 10 Hz	1-3 Hz	> 1000 Hz	< 0.1 Hz
Range of possible tasks	Good	Limited	Good/Limited	Limited
Robustness to motion	Very good	Limited	Limited	Limited
Range of possible participants	Everyone	Limited	Everyone	Limited
Sounds	Silent	High noise	Silent	Silent
Portability	Yes, for portable systems	None	Yes, for portable EEG systems	None
Cost	Low	High	Low for EEG/high for MEG	High

Functional Near-Infrared Spectrocopy and Cognitive Workload (CW): State-Of-the-Art

This chapter presents the literature review of the modalities that allow CW monitoring, with a special focus on the fNIRS technique, taking into account the subject of this thesis.

4.1 Cognitive Workload

Cognitive Workload refers to the degree of mental processing capacity imposed by a task on an individual or the amount of mental resources a person needs to employ to perform a particular task in a given environment and within a limited time period [9, 61]. Cognitive Workload increases accordingly to task difficulty due to the reduction in available cognitive resources. Cognitive Workload so high that approaches the individual's cognitive capacity, will lead to human errors and sub-optimal decisions. However, prolonged low-level mental activities, in which task demand does not increase at all, can also lead to depletion of cognitive resources. And low CW may cause boredom, distraction and human errors [44]. Given that human mental capacity is limited, it is extremely important trying to optimize this amount of resources toward specific tasks and to increase performance to obtain better results on it [7, 12, 35, 44]. Situations of high CW must be especially avoided in critical decision-making applications, such as in air traffic control or military operations, in which there is no room for error [2]. Excessive CW may lead to fatigue, memory deficits, lack of reasoning, irritability and learning difficulty [7]. Also, fatigued subjects are more likely to commit errors. In fact, the study [7] states that a large number of anesthesiologists in the United States and in Australia have commited

fatigue-related medical errors. Given the idea that the individual's attentional capacity
is finite, two components of the cognitive workload appears: the mental load and the
mental effort. The mental load is relates with the characteristics of the task itself, such as
its complexity or the way how it is presented. Mental effort, on the other hand, refers to
the quantity of cognitive resources required for the task [62].

Cognitive load theory (CLT) belongs to the human cognitive architecture that consists
of three processors - perceptual, cognitive and sensitometric - and memory. These three
processors, in turn, interplay with working memory (WM) and long-term memory (LTM),
according to cognitive process demands [11, 63]. On the education scope, cognitive
load assessment is a way of evaluating the productivity of learning materials in order to
enhance student apprenticeship [2]. This measurement is based on CLT, which searches
for the understanding of how and why a particular content is more difficult to learn than
others and it defends that our brain uses two types of memory, the short-term (working)
and long-term (storage) memory [2, 11].

- Long-term memory (LTM) exists according to all theoretical views, being an exten-
  sive store of knowledge and past events [64];

- Short-term memory (STM) is thought to be related to the faculties of the human
  mind that can hold a limited amount of information in a very accessible state tem-
  porarily [64]. The term "short-term memory" differs from the term "primary mem-
  ory", once the last one might be considered to be more restricted. Possibly not every
  temporarily accessible idea is, or even was, in conscious awareness [64]. This con-
  ception may explain why if you are speaking to a person with a foreign accent and
  inadvertently change your speech to match speaker's accent, you are influenced by
  an unconscious and uncontrollable aspect of your STM [64].

- Working memory is not completely distinct from STM. It is a term that was used
  in 1960 to refer to memory as it is used to plan and carry out behavior. One relies
  on WM to retain the partial results while solving an arithmetic problem without
  paper, for example [64]. Working memory may be seen as a temporary storage
  and manipulation system of the information necessary for complex cognitive tasks.
  These tasks include language comprehension, speaking, learning, and reasoning [2,
  37]. One reason to pursue the term WM is that measures of WM have been found to
  correlate with intellectual aptitudes (and especially fluid intelligence) better than
  measures of STM and, in fact, possibly better than measures of any other particular
  psychological process.

Stores of STM and LTM mainly differ in two ways: in duration, and in capacity. A
duration difference means that items in short-term storage decay from this sort of storage
as a function of time. A capacity difference means that there is a limit in how many
items short-term storage can hold. If there is only a limit in capacity, a number of items

smaller than the capacity limit could remain in short-term storage until they are replaced by other items. Both types of limit are controversial. Therefore, in order to assess the usefulness of the short-term storage concept, duration and capacity limits will be assessed in turn. As mentioned in [37], in the daily cognitive process of recognizing a familiar object, a sensory code is established, semantic information in LTM is accessed and used to identify the perceived object. If the matching process has a positive result, the object is recognized and it is created a STM code. This idea meets the originally stated by Shiffrin and Geisler: "The process of encoding is essentially one of recognition: the appropriate image or feature is contacted in LTM and then placed, i.e., copied in STM". More complex cognitive processes such as speaking and thinking are also a result of WM system and LTM system interaction [37]. Short-term and long-term memory systems are fundamental to cognitive processes and most IQ tests include memory span assessment, a basic STM function and analogies decoding, which relates to the LTM system. The increase in cognitive performance in childhood as well as the decrease in performance at an older age is due to performance changes in the WM and LTM systems [37].

According to the existing Multiple Resources theory, running different tasks requires a subject to exploit a separate set of resources, which are limited in capacity and distributable across tasks. In general, these resources can be of four types: processing stage (perception or cognition vs. response), perceptual modality (visual vs. auditory), visual channel (focal vs. ambient), and processing code (verbal vs. spatial) [23].

4.2 Monitoring Cognitive Workload

The methods of mental workload measurement include (1) self-reporting and subjective ratings; (2) behavioural measures; and (3) physiological measures. Subjective methods "are based on the subjective experience of task demands. Data are in the form of beliefs, values, preferences and attitudes and is collected via completion of self report questionnaires" [61]. Performance assessed by surveys may be misleading as the same level of performance may reflect multiple levels of mental workload [23]. Behavioural methods also known as task performance methods "provide assessments in the form of either speed/accuracy trade off and/or error rate data for primary tasks and/or measures of cognitive resource availability assessed via completion of secondary tasks [61]. "Self-reporting and behavioral based information tends to be delayed, sporadic, and intrusive to obtain" [23]. On the other hand, physiological measures can have little or no interference with task execution and do not require any specific behaviour by the participants [23]. Physiological methods "are based on the recording of the physiological states that are thought to correlate highly with the subjective experience, and possibly the performance decrements, associated with high and low CW situations" [61]. The physiological approach assumes that CW changes result in autonomic nervous system responses, possibly measured through physiological parameters, such as pupillometry or brain activity [23, 65]. Therefore, this work will focus on this type of CW measurement.

In the discipline of human factors, the study of CW is essential to quantify the transaction between operators and a range of task demands or technological systems or operational protocols, and for predicting the probability of performance impairment during operational scenarios, which may be safety-critical [66]. Moreover, in BCI and e-learning contexts, quantification of CW is essential for the detection and assessment of learner status. It also allows, through intelligent teaching systems, for the content of lessons to be modified according to these learner states, adapting the lesson's content to be the most fruitful as possible for the learner [14].

4.2.1 Electroencephalography

Electroencephalography is a domain concerning recording and interpretation of the electroencephalogram [67]. Electroencephalogram (EEG) is a record of the electric signal generated by the cooperative action of brain cells, or more precisely, the time course of extracellular field potentials generated by their synchronous action. Electroencephalography recorded in the absence of an external stimulus is called spontaneous EEG and EEG generated as a response to external or internal stimulus is called an event-related potential (ERP). Electroencephalography is a measure of potential difference; in the referential (or unipolar) setup, it is measured relative to the same electrode for all derivations. This reference electrode is usually placed on the earlobe, nose, mastoid, chin, neck, or scalp center [67].

Studying the brain's electrical activity is considered the gold standard to CW assessment. Electroencephalography allows the quantification of various levels of CW through brain waves power spectral density study at different cortical locations. Despite being a popular framework in the literature, where EEG indicators have been identified, it is still unclear which of the power bands are the most appropriate indicators of mental workload. The contribution of different rhythms to the EEG depends on the age and behavioral state of the subject, mainly the level of alertness. Considerable intersubject differences in EEG characteristics also exist. Electroencephalography pattern is influenced by neuro-pathological conditions, metabolic disorders, and drug action [67].

Nevertheless, alpha, beta and theta bands, addressed in Chapter 2, and ratios beta/(alpha + theta), theta/alpha and theta/beta, explained below, seems to appear in the literature as the most promising CW indicators [44].

- The ratio **beta/(alpha + theta)** interests to the study of alertness, task engagement, attention and mental effort. When alpha reduction was observed to correlate with increases in activity in frontal-parietal cortical areas, beta power increased while theta decreased, indicating a state of high vigilance. And a higher activity in the occipital and parietal regions, which seems to correlate with an alpha reduction, results in beta decrease and theta increase, reflecting drowsiness, boredom or low vigilance [44].

- **Theta/alpha** ratio (or Task Load Index, TLI) is another workload index. This one assumes that an increase in mental load causes an increase in theta power (at the frontal cortex regions) and a decrease in alpha power (at parietal regions). On the other hand, a higher theta/alpha ratio means a higher level of fatigue [44].

- A higher **theta/beta** ratio means a shorter, faster reaction time. This index may be used for attention levels assessment once it assumes that task engagement and alertness cause an increased beta power and a decreased theta power [44]. This ratio has also been linked to attention deficit, hyperactivity disorder and WM issues at young ages [44].

The resume of EEG indicators for CW assessment founded in the literature is present in Table 4.1.

Table 4.1: **Summary of EEG indicators of CW founded in literature.** Adapted from [44].

Indicator	Type of cognitive behavior	Description
Theta	Workload, vigilance and concentration.	Theta spectral power is thought to increase with increase cognitive resources demand. Theta increases in tasks requiring a sustained focus of concentration and vigilance.
Alpha	Workload, cognitive fatigue and attention.	Alpha band increases in relaxed states with eyes closed and decreases when the eyes are open. An increase in alpha power is related to lower mental vigilance and alertness.
Beta	Workload, visual attention and concentration.	An increase in beta power is associated with elevated mental workload levels during mental tasks and concentration. Beta band activity reflects an arousal of the visual system during increased visual attention.
Beta / Alpha + Theta	Mental Effort, vigilance and attention.	It has been used to study alertness and task engagement, mental attentional investment, and mental effort.
Theta/Alpha	Workload, mental effort.	This index is based in the assumption that an increase of mental load is associated with a decrease in alpha power and an increase in theta power.
Theta/Beta	Working memory, attention and sleepiness.	This index is based in the assumption that an increase in alertness and task engagement result in an increase in beta power and a decrease in theta power.

Study [68] used a EEG-based Workload index consisting of EEG measurement, eye-tracking and subjective measures to study the mental workload of drivers in traffic. This work concluded that both traffic and road types have a meaningful impact on the drivers' CW, with the advantage of being under real-life settings, instead of conducted at laboratories as the majority of studies [68]. In [45] they used the same approach for CW assessment in a multitasking environment. They concluded that both alpha and theta power increased while performing one to three sub-tasks simultaneously and better performance was linked to lower band power. The work [43] aimed at quantifying CW of participants performing multiple level cognitive tasks, while seated or walking. The results revealed higher theta and theta/alpha ratio power to an elevated level of CW, as well as a low alpha power [43]. Study [40] investigated alpha, beta, delta and theta power for building a human attention indicator, reaching an accuracy of up to 76.82% and using a support vector machine (SVM) classifier. According to [2], works [69] and [70] used a combination of frequency and time EEG features to classify different cognitive tasks, obtaining accuracies of 84.72%-98.95% and 97.78%. Study [71] used a method combining Independent Component Analysis (ICA) and SVM for CW classification and achieved an accuracy of 79.8%. Work [72] aimed at estimate CW level by classifying arithmetical tasks divided into multiple intervals of complexity, achieving an accuracy as high as 91% with K-Nearest Neighbors (KNN) classifier. Study [73] applied as feature wavelet-based entropy and SVM and reported an accuracy of 87.5%-93%. In [74] they achieved 78.31% accuracy using neural network as classification technique and applying EEG-power features for cognitive tasks [2]. In [75] wavelet-based features and SVM use resulted in 67.96%-80.71% accuracy [2]. Other study [76] employed an artificial neural network (ANN) in the cognitive task-based classification, yielding an accuracy of up to 80% [2]. The authors in [77] used discrete-wavelet transform-based feature extraction technique for the classification of learning EEG states, achieving an accuracy of 98% using SVM [2]. Study [2] collected EEG data while participants assisted multimedia learning material. They analyzed alpha, beta, delta and theta and used as features spectral entropy, approximate entropy and sample entropy. They then classified the cognitive states using the classifiers Naïve Bayes, radial basis function (RBF), and Linear kernel among the three learning conditions. Their results revealed alpha waves as being the best waves for different learning states discrimination as they achieved the highest classification accuracy. Finally, study [41] aimed at assessing student's attention level based on their EEG signals. They used gamma-approximate entropy, gamma-total variation, beta-approximate entropy, beta-total variation, beta-skewness, alpha-total variation and theta-energy as features and accomplished an accuracy as high as 89.52%, meaning their proposed attention recognizer could successfully determine whether students were attentive or not. Other works focused on using using ML to predict learner's performance. Study [78] tried to predict the success or failure of the learn on the following task, based on certain mental states extracted from EEG. These to mental states, engagement and workload, were also considered fundamental for this purpose in work [79]. In [80] the possibility of using user

performance levels to build future learning models with adaptive tutorials was studied. Many previous studies seeking to predict EEG chose a classifier instead of a regressor approach [81].

4.2.2 Functional Magnetic Resonance Imaging

Functional Magnetic Resonance Imaging (fMRI) is another way of quantifying CW once this technique is able to assess neurovascular coupling (NVC) [82, 83]. This technique provides three-dimensional images of the activated brain regions through time. More specifically, when a brain region is activated, there is an intensification in the blood flow and fMRI measures the relationship between oxyHb and deoxyHb at those locations [82, 83]. The blood oxygen level-dependent (BOLD) response is often used as an indicator of neural activity once it is generally accepted that local neural activity influences the blood oxygen level [82, 83].

Functional Magnetic Resonance Imaging is considered a gold technique for the assessment of brain activity due to its capacity of measuring across the whole brain with a high spatial resolution (<4.0 mm). Nevertheless, fMRI has some limitations that hinder its use in real-life scenarios: (i) acquisition costs relatively high; (ii) susceptibility to movement artefacts, requiring rigorous head stabilization; (iii) noisy measurements; (iv) a relatively poor temporal resolution (e.g., $\approx$0.5 Hz); (v) impossibility of being conducted outside the hospital or the laboratory and (vi) the impossibility of its use in special cases, such as individuals with metallic implants or claustrophobia [57].

In studies [83] and [84] fMRI was used to evaluate CW, through cortical activation, during n-back tasks. In this type of task, participants have to decide whether a stimulus is the same as the n-th letter before the stimulus letter [18]. The two of them observed activations in the PFC, concordant to the increasingly memory load required by the tasks. In the study [84] they also noted that the most difficult condition did not cause a reduction in cortical activity, which may indicate that in this case, a high CW did not drop the performance. However, in [85] they concluded the opposite: a considerable cognitive load harms the task accomplishment. In fact, WM load investigations using fMRI have been showing consistent increases in BOLD responses in the PFC areas of the brain [86]. Finally, study [82] used machine learning to classify mental workload states via fMRI. They were able to identify whether individuals were describing food, people, or buildings.

4.2.3 NASA Task Load Index

Another way to assess the level of CW is to ask people directly how they feel while performing a certain task through a questionnaire. The NASA Task Load Index (NASA-TLX) uses six dimensions to assess mental workload: mental demand, physical demand, temporal demand, performance, effort, and frustration [87]. Table 4.2 presents the definitions for NASA-TLX dimensions. The physical component may include activities such as

Table 4.2: **NASA-TLX rating scale definitions and endpoints.** Adapted from [87].

Title	Endpoints	Descriptions
Mental demand	Low/High	How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical demand	Low/High	How much physical activity was required (e.g. pushing. pulling, turning, controlling, activating, etc.)? Was the task easy or demading, slow or brisk, slack or strenuous, restful or laborious?
Temporal demand	Low/High	How much time pressure did you feel due to the rate or pace at which the task or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Performance	Good/Poor	How successful do you think you were in accomplishing the goals of the task set by the experimenter? How satisfied were you with your performance in accomplishing these goals?
Effort	Low/High	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration level	Low/High	How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

lifting, pulling or pushing; the temporal component relates to the time pressure for task conclusion; the effort component quantifies the level of mental and physical work needed during the task; the frustration component assesses how stress and joy connect to task accomplishment [7, 62]. Compared to other neurophysiological tools, a questionnaire is cost-effective and easy to administer, which makes it suitable for a large number of applications. The authors in [88] considered the NASA-TLX questionnaire, particularly, the subjective score on the mental demand, very sensitive to CW changes during walking in young adults [62].

4.2.4 Pupillometry

The relative size of the pupil has been assessed to quantify CW in the past 50 years [89]. Larger dilations of the pupil are associated with an increased cognitive load while performing a task. However, the following considerations should be borne in mind: (i) the pupil also dilates and contracts, depending on the amount of light detected, meaning that the luminosity must be carefully controlled to avoid false positives; (ii) the dilation and contraction of the pupil is too slow to precisely measure the cognitive response when stimuli succeed each other rapidly or overlap. Thus, to get over these constraints, the Index of Cognitive Activity (ICA) was developed [89, 90]. The pupil responds rapidly with a reflex reaction when in cognitive effort. The ICA estimation is through the count of the number of fast pupil's increases in area in a particular time interval. But at the

same time, the pupil also reacts to light variations, and ICA accurately differentiates the dilation reflex from the light change reflex [89, 90]. Thus, ICA is especially suitable for experiments needing high time resolution due to its low latency and auto-correlation. A low auto-correlation means that the frequency of rapid dilations in a certain time frame shows very little correlation with the frequency of rapid dilations in a previous time frame [89]. It may be obtained using conventional eye-tracking and because of its portability, it is more easily used in naturalistic activities such as driving, compared to the other conventional technique, such as EEG [89]. Contrary to expectation, [91] discovered that the frequency of ICA events decreased under dual-task conditions involving language comprehension and driving comparatively to single-task driving. Nevertheless, overall pupil size did increase with cognitive load under dual-task condition [89]. The study [89] used ICA to investigate two dual-task settings, combining both language comprehension and simulated driving with a memory task. Their findings confirmed that toughest linguistic processing result in larger ICA but, on the other hand, ICA did not increase in the dual-task condition compared to the single task, and, consistent with earlier findings, decreased with a more difficult secondary task.

4.3 Functional Near-Infrared Spectroscopy to Monitor Cognitive Workload

As mentioned in the previous chapters, the fNIRS technique is suitable for CW assessment because of its portability and temporal and spatial resolutions. In fact, the interest in this approach for cognitive studies is growing and believed to keep [19, 57, 58].

Relative to the fNIRS signal acquisition protocol, in particular, to the optodes' placement, 21 studies reviewed in [57] used the international EEG 10-20 system and 17 applied source-detector separations of 3.0 cm. The majority placed the optodes over the PFC. In this regard, a study using the PFC to assess cognitive effort in everyday tasks was able to detect CW changes in a reading task but not in a writing task, suggesting that the PFC may not be involved in all cognitively demanding tasks and therefore, we must design the acquisition protocol taking into account the type of task we want to access [65].

Regarding baseline brain activation recording position, 29 studies used the sitting position and it lasted from 2 s to 10 min. However, it should be $\approx$ 10-30 seconds to ensure appropriate signal-to-noise ratio. Moreover, studies using block-designs should use baselines and inter-stimulus with quite the same duration as the stimulus because of (i) the refraction time, i.e. the period with reduced responsiveness, duration is similar to the stimulation phase, and (ii) the stimulus-evoked cortical haemodynamic responses need a few seconds to return to the baseline level. Furthermore, baseline periods between the tasks should not be a multiplier of the Mayer-wave (e.g., n × 0.1 Hz). It is also recommended to vary the duration of the baseline period, i.e. to choose 12–18 s instead of a consistent 12 s, to avoid resonance effects [57].

Although it stills unclear which is the best indicator of neuronal activation, it is recommend to use at least oxyHb and deoxyHb because (i) usually neuronal activity is reflected by an increase of oxyHb and a decrease in deoxyHb; (ii) less physiological noise is present in deoxyHb signals but oxyHb signals have a higher signal-to-noise ratio as compared to deoxyHb signals; (iii) both decrease in deoxyHb and increase in oxyHb relates to an increase in the BOLD contrast obtained in fMRI; (iv) oxyHb exhibits an acceptable high reproducibility, while deoxyHb is spatially more focused; (v) deoxyHb may shows an arbitrary and paradoxical signal changes, whereas oxyHb is assumed to be the more sensitive marker of regional blood flow changes; (vi) pathologies may influence neurovascular coupling so that an decrease in deoxyHb does not necessarily reflect an increase in neural activity; and (vii) single measures (oxyHb or deoxyHb) may not be sufficient to characterize the neurovascular response of neuronal tissue [57].

The fNIRS signal has three main sources of noise: (i) instrumental noise, such as low-frequency drifts and short noise produced by light instabilities of light sources, (ii) motion-related artefacts, such as baseline shifts evoked by movements, and (iii) physiological oscillations, such as heartbeat (0.5 to 2.0 Hz), Mayer waves (0.07 to 0.13 Hz) and respiration (0.2 to 0.4 Hz). Thus, to remove physiological noise, band-pass filters with a cut-off frequency of 0.7 Hz or 0.3 Hz for the low-pass filter and 0.04 Hz for the high-pass filter should be applied. Recent reviews also recommend cut-off frequencies in the range of 0.5 Hz for low-pass filters and 0.01 Hz for high-pass filters. However, the selection of appropriate filter frequencies in functional neuroimaging needs to take into account the stimulus protocol. Thereby, cut-off frequencies for filtering must be carefully chose to avoid the unintended removal of task-evoked cortical haemodynamic responses [57].

Concerning motion-related artefacts removal in fNIRS data, there is a wide range of choices, including wavelet-based filters, Wiener filter, task-related component analysis, principal component analysis, correlation-based signal improvement, autoregressive algorithm-based filters, accelerometer-based filter methods, Temporal Derivative Distribution Repair method and artificial neural network methods. Work [57] also suggests the use of a heart rate monitor for the removal of the physiological artefacts once heart rate is associated with changes in blood flow and CW and provides information about the autonomic nervous system. Furthermore, future studies should incorporate multiple physiological parameters in fNIRS measurement, to identify the real source of the observed oxygenation changes over the head and to avoid false-positive results. It is also suggested to perform a baseline correction/normalization and averaging across channels, after filtering, to reduce individual variability of fNIRS data [57]. It was also alleged that usage of mean values is preferable compared to the use of peak values because peak values are more dependant on the accurate removal of motion and other artefacts [57]. The temporal window for the analysis should be selected considering that cortical haemodynamic response does not normally go back to the baseline level before ≈10 s (≈16 s) after stimulus presentation and that there is a delay (≈6 s) after stimulus presentation and cortical haemodynamic peak [57].

After data preprocessing, the different brain activities can be classified according to determined features. In fNIRS-BCI approach, although some studies use features extracted directly from detected light-intensity signals, most are extracted from haemodynamic measures, once they provide more options for a suitable features selection [19]. The study [36] performed a classification study of cognitive task that have been shown to create haemodynamic responses in the PFC. The classification of mental arithmetic from relax state, result on an accuracy of 71%, which is in line with that obtained in [92]. Moreover, the word generation task yielded an accuracy of 70%, while the mental rotation task had a result of 62%. These findings confirm that mental arithmetic and word generation task are efficient models for fNIRS based BCI [36]. The study [18], on the other hand, opted to assess CW using only a 3 level n-back task, achieved an accuracy of 71.5% in 1-back, 80.3% in 2-back and 80.5% in 3-back task. Their results indicate that it is possible to successfully discriminate workload induced by simpler tasks from resting states.

However, Aghajani, Peck and Solovey advocate the combination of fNIRS + EEG as a promising and forward-looking approach to brain study that generates better results than individual modalities alone. Functional Near-Infrared Spectroscopy provides complementary information to EEG, by measuring the changes in CBF and related haemoglobin concentrations. It is comparable to EEG in portability and resembles the BOLD response from fMRI, considered the gold standard for measuring cerebral haemodynamics. It also has the advantage of not having electromyographic and blink artefacts. Thus, Electroencephalography and Functional Near-Infrared Spectroscopy combination offers innovative features more suitable to the assessment of NVC and not only sensitive to neural activity, such as EEG or haemodynamics, such as BOLD response [23].

Concerning this, study [18] aimed to show how both EEG and fNIRS are a good approach to classify five different degrees for memory load. They accomplished accuracies of up to 93% for the binary classification between very high and very low workload. Two levels of workload yielded an accuracy of 74%, and discriminating full five classes was possible with an accuracy of 44% on average. They concluded that the combination of both modalities gave solidity to classification results. Work [23] also studied the capability of a hybrid fNIRS+EEG technique to quantify human mental workload during a n-back task. Their fNIRS-based features included oxyHb and deoxyHb, slope, standard deviation, skewness and kurtosis and the EEG-based set of features were determined from the PSD of alpha, beta, delta, and theta bands, phase-locking value (PLV), phase-amplitude coupling (PAC), and the asymmetry of frequency band power between right and left hemispheres. They used a SVM classifier and 10-fold cross-validation. Their results showed an accuracy of 79.6%-97.3% in classifying between task and rest and between the different n-tasks, advocating the use of EEG+fNIRS, over a single EEG or fNIRS approach, in BCI development and other applications requiring CW monitoring. Finally, [93] used fNIRS+EEG, more specifically, EEG spectral power and spatiotemporal characteristics of haemodynamic responses, and achieved an accuracy of 76.9% during a

word generation task.

Materials and Methods

This chapter presents the materials and methods used in this work. Thus, it includes the description of the (i) study sample (ii) experimental procedure used, including the cognitive tasks (iii) fNIRS and EEG signal processing, (iv) data analysis, (v) feature extraction and selection and (vi) ML models used.

5.1 Study Design and Population

Eight individuals (4 females), aged between 20 and 27 years old (M = 22.9, SD = 2.1) were recruited at Nova School of Science and Technology, at Costa da Caparica. All participants were right-handed and none reported to suffer from psychological or neurological disorders or taking medication other than contraceptive pills. The exclusion criteria were neurological pathologies and drug use since they influence the EEG pattern, namely, altering the spectral content and topographic features [67]. Prior to the biosignals acquisition, participants provided written informed consent and none objected wearing the sensors. The experimental protocols were approved by the Ethics Committee of the University. Our sampling was of convenience since we took advantage of the cases we had at our disposal due to the current pandemics situation.

5.2 Experimental Procedure

In the first place, all the room and equipment were disinfected at each acquisition, due to the current pandemic situation, following government guidelines set by the government health entity of Portugal, *Direção-Geral de Sáude*. Next, when participants entered the room, they were explained the whole procedure and also about the cognitive tasks they were going to experience. That said, they were asked to read and sign an informed

consent. After this, the sensors were connected and required software was launched for the participant to interact with. We also used a user-computer interaction monitoring software, which included image capture of the participants every 10 s [94]. The experimental procedure is represented in Figure 5.1. The data were thus collected using biosignalsplux acquisition devices, from PLUX Wireless Biosignals, at 1000 Hz and 16-bit resolution. Concerning fNIRS sensors, the two optodes emitted at 660 nm and 860 nm and emitters and detector were separated by 2 cm. Both EEG and fNIRS sensors were of only two channels each, minimizing the impact of the setup on the learning environment. These were placed at positions F7 and F8 of the EEG 10-20 system, at the prefrontal region [95]. The reference electrode of EEG was placed on the earlobe. The brightness of the room was controlled so as not to impair the results with the measurements using fNIRS. And EEG noise sources, such as electronic devices with WiFi or Bluetooth were kept away from the sensors to avoid contamination of biosignals. A breathing sensor was also placed above the participant's abdomen to aid in the process of labelling cognitive states - boredom, frustration and interest. Finally, a push-button was delivered and the participant was asked to push it every time that an interruption in task occurred (i.e., a doubt, if they needed to go to the toilet, etc.). Regarding the experimental procedure, the participants were asked to fill a sample characterization survey with some of their personal information. The cognitive tasks that followed included standard psychological tests, namely the n-back task and the mental subtraction task, which are going to be explained next. Both tasks gradually increase in difficulty and were developed using *PsychoPy* software [96]. Having completed these cognitive tasks, participants were asked to solve a Python tutorial. This tutorial aimed to simulate a lesson and the learning process. The teaching process was through asking simple questions. Sometimes, before the questions, some theoretical content was presented to help in the realization of the proposed exercises. Since this lesson was intended for students who did not know to program, the Python contents taught were introductory. Finally, a questionnaire was presented on the participants' opinion of the whole experience (whether the duration of the experience was appropriate, whether the language was understandable, whether the difficulty of the questions was adjusted, whether they had learned, etc.). Figure 5.1 resumes the experimental procedure.

5.2.1 Cognitive Task Procedure

The cognitive task procedure included the n-back task and mental subtraction. Figure 5.2 exemplifies the cognitive task procedure.

N-back task

N-back task, first introduced by Kirchner in 1958, mainly evaluates WM and sustained attention and is one of the most common WM paradigm for CW assessment [23, 97]. It is a continuous-performance task for measurement of WM capacity, which has been used frequently in the field of cognitive neuroscience. Studies [98] and [99] revealed that high task-load of n-back tasks increased EEG theta activity in the frontal midline and
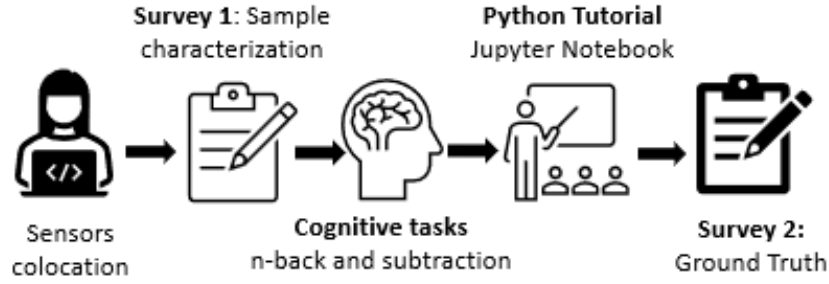
Figure 5.1: **Experimental procedure.**
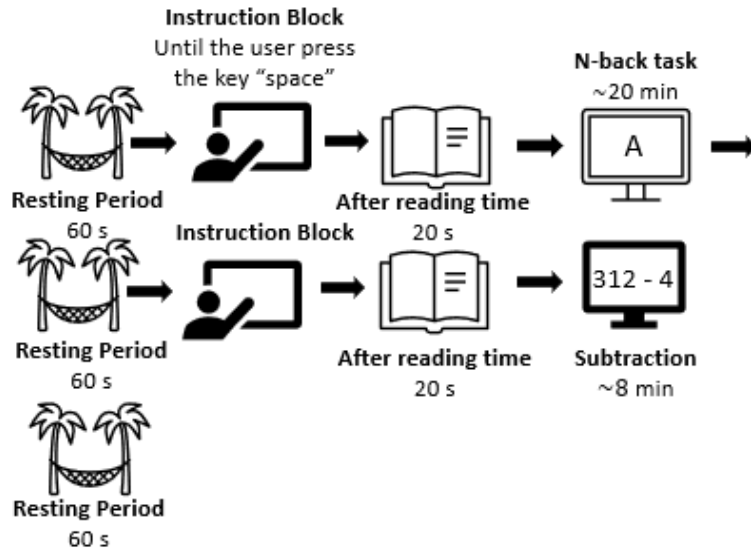
Figure 5.2: **Cognitive task procedure.**

attenuated alpha activity. In addition, fNIRS revealed WM load while performing n-back task activates PFC [18, 100–103]. The n-back task engages WM and becomes more demanding as the value of n increases. We have therefore used the n-back task as our experimental paradigm with n ranging from 0 to 3, allowing us to tune the task difficulty. N-back task consists of, first, an attention test and then a memory task. For the first level of difficulty, 0-back, the subject is asked to find the letter "A". Thus when the letter "A"appears the subject should press the key "Y"(for yes). As n increases the difficulty of the task increases. In the literature usually 0-back task has been used as a control state. In the third, 2-back, the subject must remember the letter that appear before the last, i.e. the two letters before, and press "Y"(for yes) if they correspond, or "n"(for no) if they do not. The activity began with a 60 s resting period, followed by the instruction block with the task explanation, in which the participant needed to press the key "space"or "enter"to move on. Figure 5.3 shows the instruction block presented in the activity. In the rest

period, the word "Rest"would appear on the screen and the participant should focus on the white cross on the screen until it disappeared. After this, there was a period of resting of 20 s to make sure the levels of brain activation returned to their baseline values after the reading. Then, the n-back task would begin. Each level had 60 experiments, i.e., 60 letters. Letters were A, B, C, D, and E. Each level started with an instruction block that is displayed for 5 s on the screen and informed the subject about which type of the n-back tasks was about to start (instruction block). Each letter remained on the screen for 3 s, but if the participant responded before then, the letter changed. At the end of each level, there was a 10 s resting block, defined by the word "Rest". During this block, the subject remained relaxed and fixated at a cross on the screen to let the brain activation return to its baseline and get ready for the next n-back session [18]. Before each level begins, there was an instruction block informing which levels was about to start. Total recording time was around 20 min but it depended on the subject's response time. This experience was built using the *PsychoPy* software [96]. The information about the subject's response time, and also whether the presented letter was a target or not, i.e. right and wrong answers, was recorded by this software and stored for later checking. The performance of the subjects within each level was then calculated from this information.
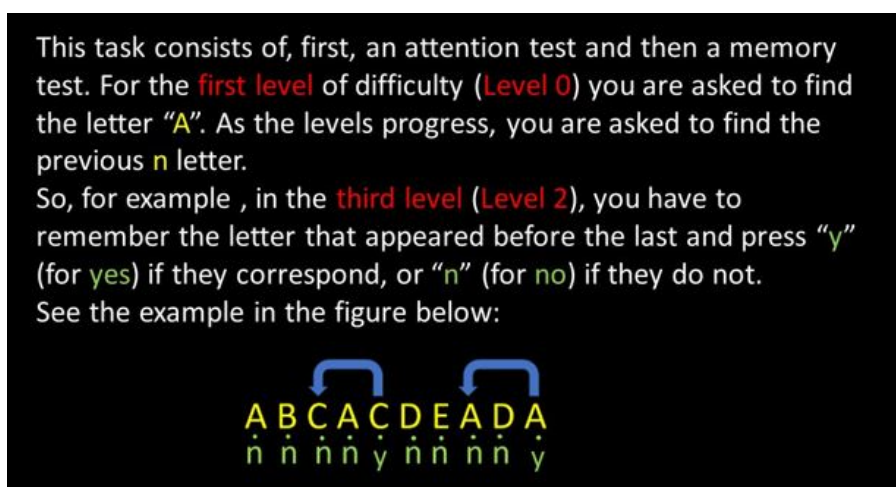
Figure 5.3: **N-back task instruction block.** Instructions that appeared on participants' screens before the task began. The duration of the task depended on the speed of response of the participant but was around 20 min.

Mental subtraction

Mental subtraction refers to performing calculations using the brain without any help in the form of paper, pen, calculator or computer. It activates the PFC and since it does not involve any body movement, it it widely used for fNIRS-BCI [19]. Backwards subtraction is the most used mental arithmetics, which involves subtraction of a small number (two-digit number) from a large number (three-digit number) with successive subtraction of a randomly appearing small number from the result of the previous subtraction [19]. The activity began with a 60 s resting period, followed by the instruction block and then a

resting time of 20 s to ensure the activation levels could return to its baseline values. Figure 5.4 shows the instruction block presented in the activity. In mental subtraction task, there were two types of crosses that could appear on the screen. The blue one was relative to resting time. The subject should focus his eyes in the cross and try to relax. Then, an subtraction should appear (p.e., 312-4) and the subject should memorize it until the orange cross appears. When it appears, the subject should focus his eyes in the cross and perform continuously the calculation (i.e., 312 - 4 = 308, 308 - 4 = 304, 304 - 4 = 300...) until the rest message and blue cross. The activity consists of 20 calculations of 12 s each followed by a resting period of 12 s each either. In the end of the task, there was a resting period of 60 s.

Figure 5.4: **Mental subtraction task instruction block.** Instructions that appeared on participants' screens before the mental subtraction task began. The task lasted from 12 min.

5.3 Signal Processing

Signal processing techniques are essential for proper analysis of results and the conclusions that arise from them. They can improve signal quality by greatly reducing noise and artefacts and, as such, they are also fundamental in Machine Learning (ML) applications [104].

In order to avoid contamination of the biosignals with external factors, the parts of the signal between push-buttons pressings were discarded. Participants were asked that if for some reason they had to pause the task, they should press the push-button to mark the beginning of the pause and then press it again when the pause was over to facilitate the removal of uninteresting parts of the signal.

5.3.1 Functional Near-Infrared Spectroscopy

The acquired fNIRS signals can contain three main sources of noise: (i) instrumental noise, such as low frequency drifts (ii) motion-related artefacts or experimental errors and (iii) physiological oscillations, such as heart beat and Mayer Waves. Since the instrumental noise and experimental errors are not related to the brain activity, it is better to remove

them prior to converting the raw optical density signals to the concentration changes of oxyHb and deoxyHb through the modified Beer-Lambert law [19].

The first step in signal processing was the ADC sample values to physical units (μV) conversion. Then, a band-pass filter was employed. This filter consists of a low-pass filter with a cut-off frequency of 0.2 Hz to remove physiological noise, such as heart rate combined with a high-pass filter with a cut-off frequency of 0.01 Hz to remove instrumental noise as low-frequency drifts. Then, a third-order Savistzky-Golay filter of window size of 11 s, to reduce spikes noise of oxyHb, experimental errors due to optodes' movement and motion artefacts was applied [19].

Thereafter, the conversion of the optical density data (red, 660 nm and infrared, 860 nm) acquired from fNIRS to oxyHb and deoxyHb concentrations was made by the use of the modified Beer-Lambert law[1], represented by Equation 3.3. The totalHb concentration, in turn, arises from the sum of the oxyHb and deoxyHb concentrations.

The signals from n-back task and mental subtraction were then divided into baseline, 0-back task, 0-back baseline, 1-back task, 1-back baseline, 2-back task, 2-back baseline, 3-back task, 3-back baseline, subtraction task and subtraction baseline signals. The baseline segments concern the most general periods of rest, namely, the longest periods (60 s) before or at the end of cognitive tasks (see Figure 5.2). This division allows a comparison to be made, firstly, between task and rest, and secondly, between different levels of task difficulty where the cognitive workload is expected to increase. Next, these signals were segmented into windows of 10 s, concerning that the cortical haemodynamic response does not normally go back to the baseline level before $\approx$10 s after stimulus presentation and that there is a delay ($\approx$6 s) after stimulus presentation and cortical haemodynamic peak [57].

A normalization, showed in Equation 5.1, was also applied to the signals. Normalization consists of adjusting the values of the signal to fit a certain scale, without modifying signal characteristics. It is a common step used in ML for data standardisation, regarding inter-subject variability.

$$x' = \frac{x - \overline{x}}{max(x) - min(x)} \tag{5.1}$$

Finally, the signals were segmented into 2 s windows (overlap = 0 s) for feature extraction, which will be presented later.

5.3.2 Electroencephalography

The acquired EEG signals were, in the first place, converted from ADC sample values to physical units (μV). They were also, similarly to fNIRS signals, divided into baseline, 0-back task, 0-back baseline, 1-back task, 1-back baseline, 2-back task, 2-back baseline, 3-back task, 3-back baseline, subtraction task and subtraction baseline signals. Next, the

---

[1]The conversion between optical density and the concentrations of oxyHb and deoxyHb was performed using the package *mes2hb* [105]

signals were segmented into 10 s windows and normalized from the Equation 5.1. The filtering technique that followed included a band-pass filter in which the upper cut-off frequency and the lower cut-off frequency varied according to the wanted band extraction. Thus, for example, if the band to extract was delta (0.1 - 3 Hz), the lower cutoff frequency would be 0 Hz but if instead, it was the beta band (13 - 30 Hz), the cut-off frequency would change for 3 Hz. In a similar way, the upper frequency was 30 Hz for all the bands except for gamma, in which it was 50 Hz. This filter with cutoff frequencies of 3 and 30 Hz eliminates noise and artefacts, keeping the frequencies in the range of brain activity of awake adults. Table 5.1 resumes the cutoff frequencies used in EEG signals processing.

Table 5.1: **Lower and upper cut-off frequencies for EEG bands extraction.**

Band	Lower Cutoff Frequency	Upper Cutoff Frequency
Delta (0.1 - 3 Hz)	0	30
Theta (4 - 8 Hz)	3	30
Alpha (8 - 13 Hz)	3	30
Beta (13 - 30 H)	3	30
Gamma (> 30 Hz)	3	50

5.4 Data Analysis

This work focuses on two issues: (i) identifying whether a subject is in the cognitive task or at rest and (ii) quantifying the CW experienced by the subject while performing a cognitive task. The statistical test used was Wilcoxon signed-rank test, a non-parametric version of the T-test for paired and dependent samples from the same distribution. It tests the null hypothesis that the underlying distribution of sample x is the same as sample y [106]. The significance threshold was set at 0.05.

5.4.1 Task versus Rest

The difference between the levels of oxyHb and deoxyHb concentrations of task and resting-state can be very notorious, as shown in Figure 5.5. This figure is relative to the mean wave of fNIRS signal from a single subject. The mean wave was set as a function that returns a segment with the means of every segment at the same point, as represented in Equation 5.2, in which x_i represents a segment of fNIRS data, t represents the instant of time (that should be the same to every segment) and n represents the number of segments of fNIRS data.

$$y(t) = \sum_{i=1}^{n} \frac{x_i(t) + ... + x_n(t)}{n} \tag{5.2}$$

This approach was used especially for data visualization once it enables a more general notion of the signal' behaviour. When a subject is interested and engaged in a cognitive

exercise, there is an intensification of blood flow in activated brain areas used to perform the activity. As so, and because the local oxygen supply is higher than its consumption, an higher local oxyHb concentration and a lower local deoxyHb concentration are expected. During resting periods, these concentrations should return to their baseline values, once the individual is not supposedly engaged in a specific task. The Wilcoxon signed-rank test statistical test was performed for the relevance evaluation of the differences between the values of EEG power bands found for the task and the baseline. The results are presented in Chapter 6.1.
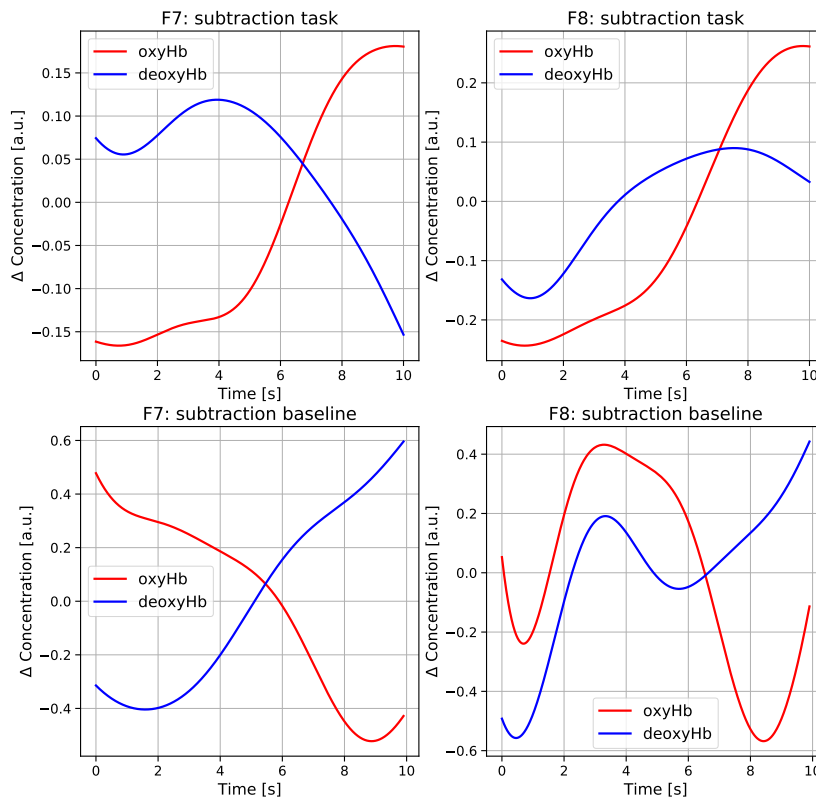
Figure 5.5: **Mean oxy and deoxyHb concentrations during subtraction task and rest.** Signal from a single participant recorded from the positions F7 and F8 of the 10-20 system, during the subtraction task (on top) and at rest (on bottom).

At the spectral level, differences can also be seen in the power spectra of the average oxyHb and deoxyHb waves as shown in Figure 5.6.

Understanding when a subject is or is not in cognitive task is relevant for e-learning platforms, in which the absence of the teacher makes it difficult to detect whether or not the student is paying attention. For this reason, a classification between task and rest was carried out, which will be discussed and deepened in the following sub-chapters.
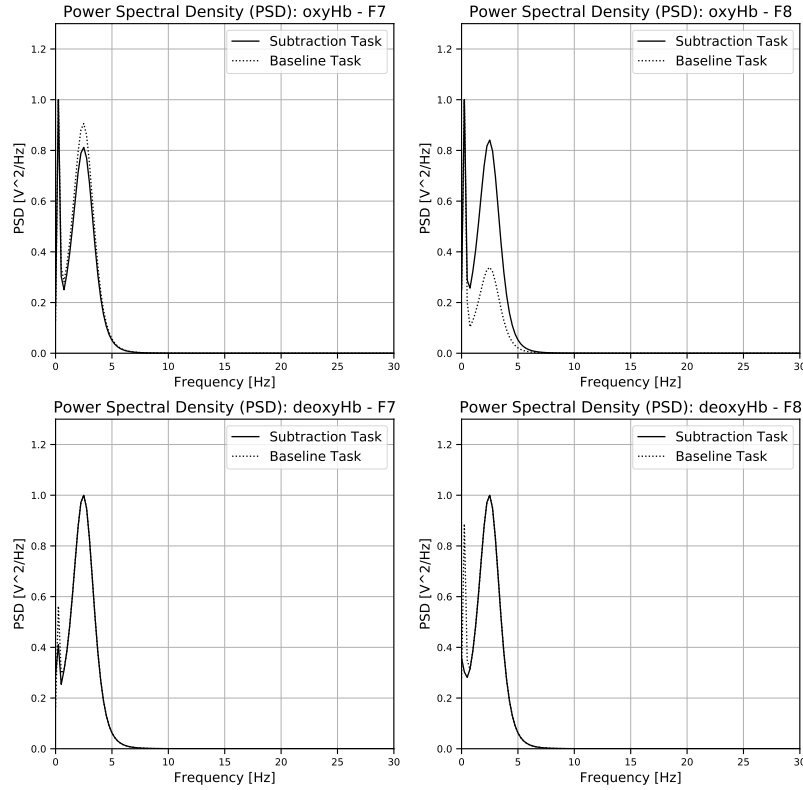
Figure 5.6: **Power spectra of the mean oxyHb and deoxyHb signals.** Power spectral density (PSD) of the oxyHb (on top) and deoxyHb (on bottom) mean waves from the positions F7 (on the left) and F8 (on the right) of the 10-20 system, during the subtraction task and at rest. The deoxyHb medium wave spectra show that the maximum frequency is different during the subtraction task and the resting period.

5.4.2 Cognitive Workload Estimate

Cognitive performance can be improved by adapting to the individual's mental needs. Generally, cognitive load is correlated with task difficulty; more complex tasks increase the cognitive load experienced [86]. The Task Load Index (TLI) is a CW quantification tool, which studies theta and alpha powers, as stated in Chapter 4. Actually, the TLI is defined as the ratio of the mean frontal midline theta energy to the mean parietal alpha energy [107]. Since our study assessed alpha power in the frontal area and not in the parietal one, we calculated the ratio of the mean frontal midline theta energy to the mean frontal alpha energy - Modified Task Load Index (MTLI). For this, the Wilcoxon signed-rank test statistical test allowed us to understand if the differences between alpha and theta were significant for the calculation of the MTLI. It was first analysed in an attempt to distinguish between cognitive task and rest period, being calculated for each of these

43

modes. The results are presented in Chapter 6.1. Secondly, to assess the relationship between task difficulty and cw across n-back tasks. The results of this analysis can be found in Chapter 6.2. And thirdly, this index was employed to understand the association between cognitive states and workload. The results of this analysis can be found in the Chapter 6.3.

5.4.3 Cognitive States Quantification

Cognitive states analysis is also essential in the context of learning. In intelligent tutoring systems and e-learning environments, the detection and evaluation of some cognitive states, such as boredom, frustration and interest, are important for their design. For example, automatically detecting whether or not the student is engaged in the task, allowing the lesson content to be adjusted according to this variation, could improve student performance. Similarly, a system that uses ML algorithms to classify whether the student is bored at a given moment in a given task and its quantification, can be used to adjust the learning strategy and provide the student with more appropriate content and/or feedback, avoiding disengagement situations [14].

For this reason, a regression was carried out trying to quantify the cognitive states, which will be discussed and deepened in the following sub-chapters.

To this end, a labelling process was conducted, in which we used image captures of the participants at various instants of the experiment, response time and signals from the respiration sensor to label the subjects' state as bored, frustrated, interested and at rest. We also had access to the electrodermal activity (EDA), electrocardiogram (ECG) and accelerometer sensors but these were not considered in the labelling process.

In Russell's Circumplex model shown in Figure 5.7, the x-axis is valence and the y-axis is arousal. Valence reflects the pleasantness of stimuli. For example, hapiness has a positive valence and disgust has a negative valence. The other dimension, arousal, reflects the activation level [108]. Thus, interest has a high arousal, while boredom has a low arousal. Regarding this model, boredom is placed in the third quadrant with low arousal and low valence and frustration is placed in the second quadrant with medium arousal and low valence.

Boredom is a temporary feeling of low-arousal and unpleasant emotions induced by environmental factors. Boredom impairs attention, making it difficult to perform cognitive tasks [13]. The expression of boredom is characterised by closed lips without flexion of the zygomatic muscle, i.e. the facial muscle that raises the angle of the mouth. The eyes should be open, but with the eyelids slightly drooping. The gaze may be directed at nothing in particular, which means a lack of interest and attention. The eyes may also be rolled back as a manifestation of boredom. The posture is slouching and shrinking. Overall, the expression and characterization of boredom have received little research attention [13].

Frustration is a fundamental negative emotion possibly defined as irritable distress,
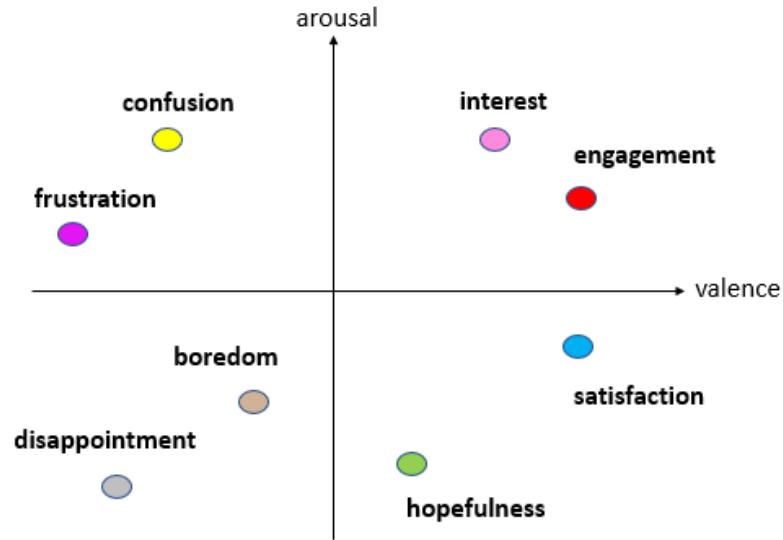
Figure 5.7: **Russell's Circumplex Model.** The x-axis is referent to valence and the y-axis to the arousal. Adapted from [109].

which arises in response to limitation, failure and exclusion [16]. The intensity of frustration depends on the reward value of the goal (i.e. the more satisfaction the completion of a task gives us, the more frustration its failure will cause us) and one's capacity for self-control [16]. Frustration can contradictorily improve and worsen cognitive performance [15].

Interest is central to learning. It explains why we select and persist in processing certain types of information over others. Interest is also very close to mental engagement, which may be defined as the level of mental vigilance and alertness, while performing a task [14].

For the labelling exercise, the following considerations were taken into account: (i) bored mode links to deeper and slower breathing, sighs, narrower eyes, lean back in the chair, resting head on the hand, looking away, and slow response time; (ii) frustrated mode links to faster but not deeper breathing, eyes open, more body movements than usual (moving the head, leaning back, etc) and (iii) interested mode links to regular breathing, eyes open, forward-leaning posture and normal or fast response time. The resting mode consists of the nonexistent time answer. In case of doubt, it was assumed that the previous label was maintained. We used this labelling process to construct the regression index. In which, a value of 0 would mean, for example, not at all interested and 1 fully interested.

We used the MTLI to quantify the CW associated with each cognitive state and study their relationship. The Wilcoxon signed-rank test statistical test was also used to assess the statistical relevance of the difference between the alpha and theta powers required to calculate the MTLI. Finally, we calculated the engagement index defined by beta power/(theta power + alpha power) for the interest state, comparing these values with

the regression results for each [107].

5.5   Features Extraction and Selection

In the classification process, the segments of 10 s from fNIRS and EEG were then segmented into windows of 2 s and overlap of 0 s, originating 2583 samples. Initially, 176 features from the temporal, statistical and spectral domain were extracted from oxyHb, deoxyHb and totalHb concentrations segments, but only the 15 most relevant features were included, by the feature selection method SelectKBest provided by scikit-learn python library [110]. This method scores the features based on the distance between the means of different classes and variance of each single class. In the feature selection technique, the features in our dataset that contribute the most to the target variable are chosen. In other words, we choose the best predictors for the target variable.

SelectKBest method combines the univariate statistical test with selecting the K-number of features based on the statistical result between all features. As an example, consider that we have two classes, and we want to find a score for each feature, looking for how well this feature discriminates between two classes. Figure 5.8 presents 2 classes, red and blue and 2 features, on x and y axes. The x feature is a more effective separator than y, since by projecting data onto the x-axis, we get two completely separate classes, unlike y, where the two classes overlap at the centre of the axis [111]. In line to x, the two classes are far from each other. In other words, the distance between means of class distributions on x is more than y. According to x once again, the scatter of classes do not fall on each other but according to y they do. It means that according to x, classes are more compact so more probable to not have an overlap with another class. Meaning that the variance of each single class according to x is less than those of y. Thus, the higher the score $\frac{distance\_between\_classes}{compactness\_of\_classes}$ is, better the features discriminates between classes [111]. This definition explains what good and bad features mean. The following math formulations are a way of quantify the quality of features.

Selecting the best features based on univariate statistical tests can be seen as a preprocessing step to an estimator once applying these tests before the classifier increases the classifier weight attributed to the significant features, thus possibly improving classification [112].

These tests are called univariate tests because they do not look at all features collectively. In particular, they do not take into account the interaction between features. Instead, they examine each features separately and assess whether there is a significant relationship between that feature and the target [112].

In this work, the univariate test chose for the classification was the the ANOVA F-value for the provided sample [110]. The analysis of variance (ANOVA) is a statistical analysis tool used to determine the influence that independent variables have on the dependent variable in a regression study [112]. It compares the variation between each species with the variation within each species. Through some statistical assumptions, it proves that
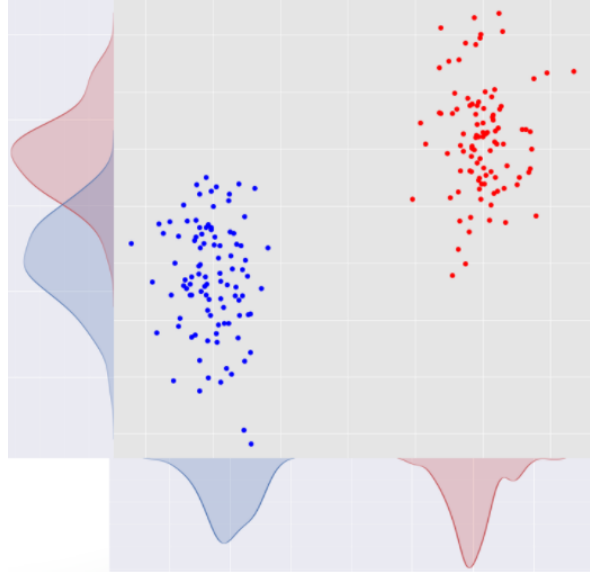
Figure 5.8: **Diagram exemplifying how *SelectKBest* selects features with** $f_{classif}$. x and y represent features, while red and blue represent classes. The x feature is a more effective separator than y, since by projecting data onto the x-axis, we get two completely separate classes, unlike y, where the two classes overlap at the centre of the axis. Reprinted from [111].

the ratio between the two variations follows the F-distribution, resulting in a high F-score and p-value to be evaluated [112].

Table 5.2 shows features names of the temporal, statistical and spectral-domain. These features, which are briefly explained below, were calculated from the *Time Series Feature Extraction Library* [113]. We construct several datasets, allowing the evaluation of different classification perspectives. Thus, the fNIRS dataset is composed only of fNIRS features, the EEG dataset is also composed only by EEG features and the fNIRS + EEG dataset combines the features from both sensors. In addition to these, we also created a dataset exclusive to the n-back task and another dataset exclusive to the subtraction task. Table 5.3 presents the selected features of fNIRS, EEG and fNIRS + EEG datasets. Since the methodology used was Stratified 10-Folds cross-validation, features chosen were not always the same in all 10 iterations. Table 5.4 shows selected features of n-back and subtraction. Table 5.5 presents the selected features of fNIRS, EEG and fNIRS + EEG datasets using Leave One Participant Out, instead of the Stratified K-Fold method. Finally, Table 5.6 presents the selected features of n-back and subtraction datasets using Leave One Participant Out method, instead of the Stratified k-Fold method.

- **Mean of Differences** - computes mean of differences of the segment (derivatives) [113].

- **Total Energy** - computes the total energy of the segment [113].

- **Area Under Curve** - computes the area under the curve of the segment computed with trapezoid rule [113].

- **Absolute Energy** - computes the absolute energy of the segment [113].

- **Peak to Peak Time** - computes the peak to peak distance of the segment [113].

- **Entropy** - computes the entropy of the segment using the Shannon entropy [113].

- **Polarity** - computes the absolute value of the division between the maximum and minimum values of a segment.

- **Slope** - computes the slope of the segment by doing a linear regression, using the library SciPy [114].

- **Zero Crossing Number** - function that counts the total number of times that the segment crosses the zero value, i.e. counts changes from positive to negative or vice versa [113].

- **Maximum Value** - computes the maximum value of the segment [113].

- **Minimum Value** - computes the minimum value of the segment [113].

- **Mean Value** - computes mean value of the segment [113].

- **Variance** - computes variance of the segment [113].

- **Standard Deviation** - computes standard deviation of the segment [113].

- **Kurtosis** - computes kurtosis of the segment [113]. Kurtosis is a measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution. Thus, datasets with high kurtosis tend to have heavy tails, or outliers. Datasets with low kurtosis tend to have light tails, or no outliers. A uniform distribution would be the extreme case [115].

- **Skewness** - computes skewness of the segment [113]. Skewness is a measure of symmetry. A distribution is symmetric if it looks the same to the left and right of the center point [115].

- **Fundamental Frequency** - computes fundamental frequency of the segment, i.e. the frequency corresponding to the first harmonic, the lowest, the first peak of the wave [113].

- **Maximum Frequency** - corresponds to the frequency where the cumulative sum of the Fourier transform magnitude reaches 95% [113].

- **Power Band Width** - computes power spectrum density bandwidth of the segment [113].

- **Spectral Distance** - computes the segment spectral distance, distance of the segment's cumulative sum of the spectral features elements to the respective linear regression [113].

- **Median Frequency** - computes median frequency of the segment [113].

- **Spectral Entropy** - computes the spectral entropy of the segment based on Fourier transform [113].

Table 5.2: **Features used in this work from the temporal, statistical and spectral domain.** This table sums all features that were calculated in this work. The best features for classification were then selected by *SelectKBest*.

Temporal	Statistical	Spectral
MD - mean of differences		
TE - total energy	MAX - max value	f0 - fundamental frequency
AUC - area under curve	MIN - min value	fmax - maximum frequency
ABS_E - absolute energy	MEAN - mean value	PBW - power band width
dP_P - peak to peak distance	VAR - variance	SD - spectral distance
S - entropy	STD - standard deviation	MF - median frequency
Pl - polarity	KURT - kurtosis	SS - spectral entropy
m - slope	SKEW - skewness	
ZCR - zero crossing rate		

Table 5.3: **Selected features with the Stratified 10-Fold method of fNIRS, EEG and hybrid fNIRS + EEG datasets.** These features were selected by *SelectKBest*. The selected features of the hybrid dataset are the same as in the EEG-only dataset, which indicates that the EEG features allow to better differentiate the task state from the resting one, compared to the fNIRS features. Features names followed by oxyHb/deoxyHb/totalHb refer to fNIRS features, while features names followed only by Left/Right refer to EEG features.

fNIRS	EEG	fNIRS + EEG
dP_P oxyHb Left	AUC Left	AUC Left
dP_P deoxyHb Left	AUC Right	AUC Right
dP_P deoxyHb Right	dP_P Left	dP_P Left
dP_P totalHb Left	dP_P Right	dP_P Right
dP_P totalHb Right	TE Right	TE Right
STD oxyHb Left	ABS_E Right	ABS_E Right
STD deoxyHb Left	VAR Left	VAR Left
STD deoxyHb Right	VAR Right	VAR Right
STD totalHb Left	MAX Left	MAX Left
SD oxyHb Left	MAX Right	MAX Right
SD oxyHb Right	MIN Left	MIN Left
SD deoxyHb Left	MIN Right	MIN Right
SD deoxyHb Right	STD Left	STD Left
SD totalHb Left	STD Right	STD Right
SD totalHb Right	SD Left	SD Left
fmax deoxyHb Left [2]	SD Right	SD Right
fmax deoxyHb Right		

49

Table 5.4: **Selected features with the Stratified 10-Fold method of n-back and subtraction datasets.** These features were selected by *SelectKBest* and are from fNIRS and EEG. The selected features of the n-back dataset are mostly from fNIRS, while in the subtraction dataset the features are mostly from EEG. Features names followed by oxyHb/deoxyHb/-totalHb refer to fNIRS features, while features names followed only by Left/Right refer to EEG features.

n-Back	Subtraction
dP_P oxyHb Left	
dP_P deoxyHb Left	AUC Left
dP_P deoxyHb Right	AUC Right
dP_P totalHb Left	dP_P Left
VAR deoxyHb Left	dP_P Right
STD oxyHb Left	TE Right
STD deoxyHb Left	ABS_E Right
STD deoxyHb Right	VAR Left
STD totalHb Left	VAR Right
SD oxyHb Left	MAX Left
SD oxyHb Right	MAX Right
SD deoxyHb Left	MIN Left
SD deoxyHb Right	MIN Right
SD totalHb Left	STD Left
SD totalHb Right	STD Right
fmax deoxyHb Left	SD Left
fmax deoxyHb Right	SD Right
fmax Right	
f0 Right	

The regression dataset included the fNIRS+EEG features used for the classification approach, i.e for distinguish between cognitive task and rest and features of delta, theta, alpha, beta and gamma bands powers.

The band power characteristics were extracted by transforming each instant from the time domain to the frequency domain using the Welch method. The Welch method averages the Fast Fourier Transform (FFT) results from several overlapping Hamming window segments. Thus, a window size of 2 s (2000 points) and overlap of 0 s.

5.6 Machine Learning Models

In this work, ML approaches were used for classifying between cognitive task and resting and for quantifying cognitive states, namely, interest, boredom and frustration, by the use of regression. Machine Learning is a category of artificial intelligence that enables computers to think and learn on their own, acting without being explicitly programmed [116]. It is one of today's most promising areas, where computer science and statistics intersect, and at the heart of artificial intelligence and data science [117].

Machine Learning algorithms can detect patterns in data and then use the discovered

Table 5.5: **Selected features with the Leave One Participant Out method of fNIRS, EEG and hybrid fNIRS + EEG datasets.** These features were selected by *SelectKBest*. The selected features of the hybrid dataset are not the same as in the EEG-only dataset. Features names followed by oxyHb/deoxyHb/totalHb refer to fNIRS features, while features names followed only by Left/Right refer to EEG features.

fNIRS	EEG	fNIRS + EEG
	TE Left	dP_P deoxyHb Left
	TE Right	dP_P deoxyHb Right
	ABS_E Left	dP_P Left
	ABS_E Right	dP_P Right
dP_P oxyHb Left	dP_P Left	STD oxyHb Left
dP_P deoxyHb Left	dP_P Right	STD deoxyHb Right
dP_P deoxyHb Right	S Left	STD totalHb Left
dP_P totalHb Left	S Right	STD Left
STD oxyHb Left	Pl Left	STD Right
STD deoxyHb Left	ZCR Left	SD oxyHb Left
STD deoxyHb Right	MIN Left	SD oxyHb Right
STD totalHb Left	MIN Right	SD deoxyHb Left
STD totalHb Right	MAX Left	SD deoxyHb Right
VAR deoxyHb Left	MAX Right	SD totalHb Left
VAR deoxyHb Right	VAR Left	SD totalHb Right
SD oxyHb Left	VAR Right	SD Left
SD oxyHb Right	SKEW Left	SD Right
SD deoxyHb Left	KURT Right	fmax deoxyHb Left
SD deoxyHb Right	fmax Left	fmax deoxyHb Right
SD totalHb Left	fmax Right	AUC Left
SD totalHb Right	SS Left	AUC Right
fmax deoxyHb Left	AUC Left	ABS_E Right
fmax deoxyHb Right	AUC Right	MAX Left
	STD Left	MAX Right
	STD Right	MIN Left
	SD Left	MIN Right
	SD Right	VAR Right

patterns in decision-making scenarios for future's data prediction [118]. For the detection of these patterns they may use two types of learning: the supervised learning and unsupervised learning. In the supervised learning, one learns from the behaviour of the dataset by studying a train set where each data point is labelled and categorised. Here, when we are given a set of labelled data we already have an idea of what our correct output will be, considering that since there is a relationship between the input and the output, they should be similar. Classification and regression techniques use this type of learning. On the other hand, in unsupervised learning, the behaviour of a data set is learnt without prior knowledge about it [118].

The classification approach tries to solve a problem in which the output can be only

Table 5.6: **Selected features with the Leave One Participant Out method of n-back and subtraction datasets.** These features were selected by *SelectKBest* and are from fNIRS and EEG. The features chosen for the n-back task were the same with Stratified k-Fold and Leave One Group Out methods. Features names followed by oxyHb/deoxyHb/totalHb refer to fNIRS features, while features names followed only by Left/Right refer to EEG features.

n-Back	Subtraction
	dP_P deoxyHb Left
	dP_P deoxyHb Right
	dP_P totalHb Right
	dP_P Left
	dP_P Right
dP_P oxyHb Left	VAR deoxyHb Left
dP_P deoxyHb Left	VAR deoxyHb Right
dP_P deoxyHb Right	VAR Right
dP_P totalHb Left	STD deoxyHb Left
STD oxyHb Left	STD deoxyHb Right
STD deoxyHb Left	STD totalHb Right
STD deoxyHb Right	STD Left
SD oxyHb Left	STD Right
SD oxyHb Left	SD oxyHb Left
SD oxyHb Right	SD oxyHb Right
SD deoxyHb Left	SD deoxyHb Left
SD deoxyHb Right	SD deoxyHb Right
SD totalHb Left	SD Left
SD totalHb Right	SD Right
fmax deoxyHb Left	MAX Left
fmax deoxyHb Right	MAX Right
fmax Right	MIN Left
SS Left	MIN Right
f0 Right	fmax deoxyHb Left
	AUC Left
	AUC Right
	TE Right
	ABS_E Right

one of a fixed number of output classes known a priori (p.e. yes/no or true/false). Depending on the number of output classes, the problem can be a binary or multi-class classification problem. On the other side, we use regression algorithms to deal with problems using continuous and numeric outputs, answering questions such as *how much* or *how many* [116].

5.6.1 Classification

For the classification between task and rest, a Random Forest (RF) algorithm was employed, together with the Stratified k-Fold and Leave One Participant Out methods. Random Forest is a supervised learning algorithm. The "forest"it builds, is an ensemble of decision trees that it trains in parallel, with bootstrapping followed by aggregation, usually referred as "bagging" [119]. Figure 5.9 shows a diagram of RF implementation technique, exemplifying bootstrapping and aggregation phenomenons. When training, each tree in a RF learns from a random sample of the data points. The samples are drawn with replacement, known as bootstrapping. Every decision tree consists of decision nodes, leaf nodes, and a root node. The leaf node of each tree is the final output produced by that specific decision tree. For the final decision, RF classifier aggregates the decisions of individual trees. The selection of the final output follows the majority-voting system, thus, RF classifier exhibits good generalization [119]. Random Forest classifier tends to outperform most other classification methods in terms of accuracy without issues of overfitting [119]. The general idea of the bagging method is that a combination of learning models increases the overall result allowing for a more stable and more accurate forecast.
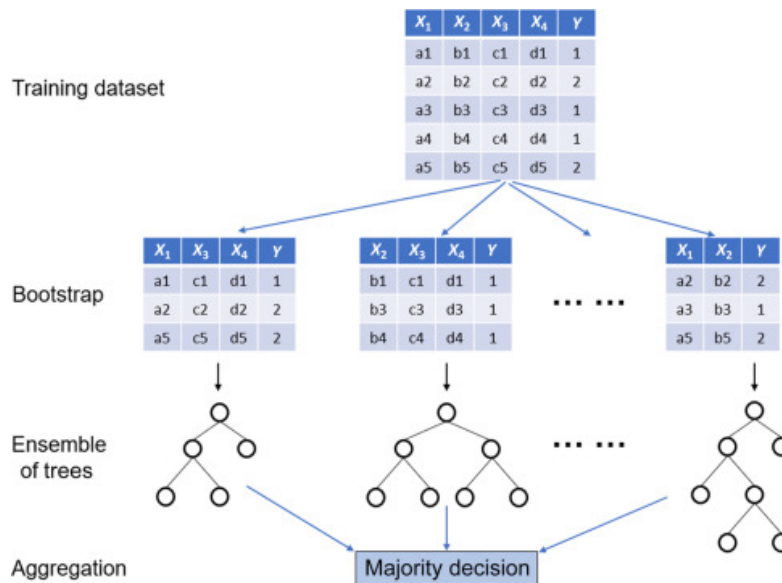
Figure 5.9: **Diagram of Random Forest classification structure.** The represented dataset is composed of 4 features (X_1, X_2, X_3 and X_4) and 2 classes (Y=1 and 2). Reprinted from [119].

53

The k-Fold technique divides all samples into groups of samples (folds) of equal sizes. The prediction function uses k-1 folds for training and leaves one fold out for testing. The Stratified k-Fold method is a variation of k-fold in which each fold contains approximately the same percentage of classes in the training set and the test set [110]. The shuffle method was also used to mix the samples in the dataset.

For training the models across subjects (i.e. user-independent models) the Leave One Participant Out method was used, which is similar to k-fold cross-validation except that the training and test sets are completely independent. Using this technique, we reserved one participant for the test set and the remaining k-1 participants for the training set. This was repeated k times, where k = number of subjects, so that each subject is used as the test set once [120].

The metrics chosen for the evaluation of the classifier performance were accuracy, f-score, precision, recall and Area Under the ROC Curve (AUC-ROC). A brief explanation of each of the metrics appears below.

The evaluation of the best solution for binary classification problems is based on confusion matrix, shown in Table 5.7. In this table, the row represents the predicted class, while the column represents the actual class. From this confusion matrix, *TP* and *TN* denote the number of positive and negative instances correctly classified, while *FP* and *FN* denote the number of misclassified negative and positive instances, respectively. From Table 5.7, several metrics can be generated, which are explained below.

Table 5.7: **Confusion matrix for binary classification.** Reprinted from [121].

	Actual Positive Class	**Actual Negative Class**
Predicted Positive Class	True Positive (TP)	False Negative (FN)
Predicted Negative Class	False Positive (FP)	True Negative (TN)

- **Accuracy** - reflects the ratio of correct predictions over the total number of instances evaluated [121]. Accuracy is calculated by the equation

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{5.3}$$

- **AUC-ROC** - The Receiver Operator Characteristic (ROC) curve is an evaluation metric for binary classification problems. It is a probability curve that plots the *TP* against *FP* at various threshold values, separating the signal from the "noise". AUC-ROC reflects the overall ranking performance of a classifier [121]. It is calculated by the formula

$$AUC - ROC = \frac{S_p - n_p(n_n + 1)/2}{n_p n_n}, \tag{5.4}$$

where S_p is the sum of all positive examples ranked, while n_p and n_n denote the number of positive and negative examples respectively[121].

- **Precision** - measures the positive samples that are correctly predicted from the total predicted samples in a positive class [121]. Precision is calculated by the equation

$$Precision = \frac{TP}{TP + FP} \tag{5.5}$$

- **Recall** - measures the fraction of positive samples that are correctly classified [121]. Recall is calculated by the equation

$$Recall = \frac{TP}{TP + TN} \tag{5.6}$$

- **F-Score** - reflects the harmonic mean between recall and precision values [121]. F-score is calculated by the equation

$$F - score = \frac{2 * P * R}{P + R}, \tag{5.7}$$

where P stands for precision and R for recall.

5.6.2 Regression

For the regression technique, trying to quantify the cognitive states - boredom, frustration and interest - the RF algorithm, introduced previously, and the Extra Trees algorithm, which will be discussed next, were used. For splitting the training and test set the Stratified k-Fold method was employed with k = 10 and with shuffle, to allow the samples to be mixed. The Random Forest, as explained earlier, works by constructing several of decision trees, choosing the class that is the mode of the classes, in classification, or the average prediction of the individual trees, in regression [122]. Figure 5.10 shows the RF regression structure.

Extra Trees (ET), the short name for Extremely Randomized Trees, works like an RF in that it builds multiple trees and splits nodes using random subsets of features. However, the two algorithms have two key differences: ET does not bootstrap observations, i.e., it samples without replacement and instead of choosing the best possible threshold for each tree at each node, splitting nodes in the best combination based on the subset of features selected, it simply opts for a random edge, splitting nodes randomly [123].

The metrics chosen for the evaluation of the regressors performance were R^2, mean square error (MSE) and mean absolute error (MAE). A brief explanation of each of the metrics appears below.

- **R^2** - reflects the variability in dependent variable that can be explained by the model. It is the square of the Correlation Coefficient (R) [124]. R^2 is calculated by the equation

$$R^2 = 1 - \frac{SS_{regression}}{SS_{total}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2}, \tag{5.8}$$

where y_i represents the real output, $\hat{y}_i$ represents the predicted output and $\bar{y}_i$ represents the baseline model, the mean. The R-Squared formula compares our fitted

55
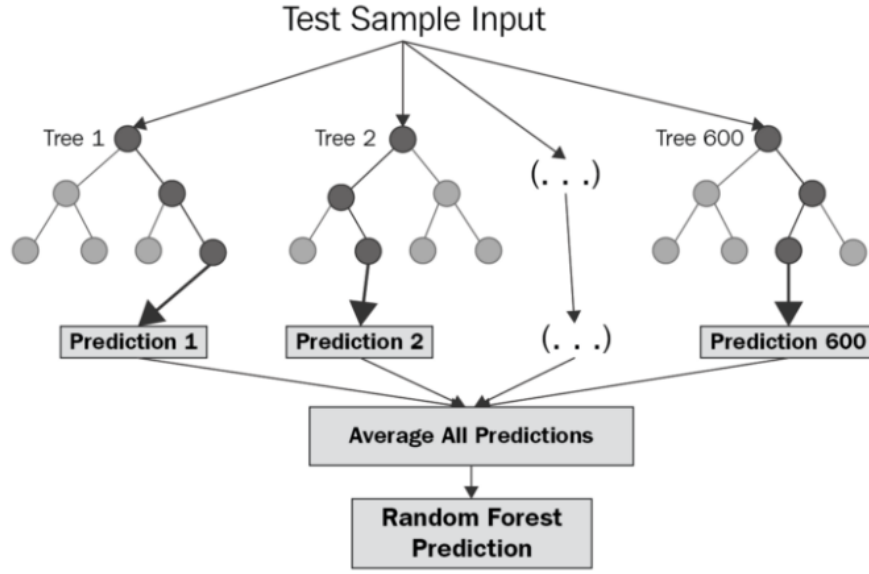
Figure 5.10: **Diagram of Random Forest regression structure.** Reprinted from [122].

regression line to a baseline model. This baseline model is considered the "worst" model. The R-squared value is normally between 0 and 1 and the higher this value, the closer the predicted value is to the actual value [124]. When this value is negative, it means that the model selected does not follow the trend of the data, therefore leading to a worse fit than the horizontal line.

- **Mean Square Error** - while R Square is a relative measure of the fit between the model and dependent variables, Mean Square Error (MSE) is an absolute measure of the quality of the fit. MSE is calculated by the equation

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2, \tag{5.9}$$

where y_i represents the real output and $\hat{y}_i$ represents the predicted output. This metric reflects the absolute number on how much the predicted results deviate from the actual value. It is sensitive to outliers [124].

- **Mean Absolute Error** - Mean Absolute Error (MAE) is close to MSE but instead of the sum of square of error in MSE, MAE is taking the sum of the absolute value of error. MAE is calculated by the equation

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i|, \tag{5.10}$$

where y_i represents the real output and $\hat{y}_i$ represents the predicted output. MAE is a more direct representation of sum of error terms than MSE once MSE penalises big prediction error by square it, while MAE does not distinguish between errors. It is not sensitive to outliers [124].

56

R E S U L T S

This chapter is divided into three parts: results concerning (i) the distinction between mental task and rest cognitive states; (ii) the study of the influence of cognitive task difficulty on experienced CW and (iii) the study of the relationship between cognitive states - boredom, frustration and interest - and CW.

6.1   Task Workload and Baseline Workload

This section shows the results of the distinction between the workload experienced during cognitive tasks and during the rest period.

Figure 6.1 shows the mean oxyHb and deoxyHb concentrations of the 8 subjects in the subtraction task and resting period for the F7 position of the 10-20 system. From this figure, we conclude that the graphic information provided by fNIRS is not sufficient to distinguish between task and rest based on the variations of the concentrations of the haemoglobin chromophores. The signals provided do not allow us to evaluate the trend of the chromophores, nor to distinguish between task and rest, from the level of brain activation. Figure A.1 presents the mean wave of the 8 subjects in the subtraction task and resting period but this time for the F8 position of the 10-20 system. The figure can be found in the Appendix A.

The EEG bands powers, for the MTLI assessment, were calculated for the distinction between the subtraction task and resting period (see Tables A.1 and A.2). However, the results showed that the difference between alpha and theta powers were not statistically significant in the resting state. Therefore, we did not compare the MTLI between cognitive task and rest period (see Table A.3).
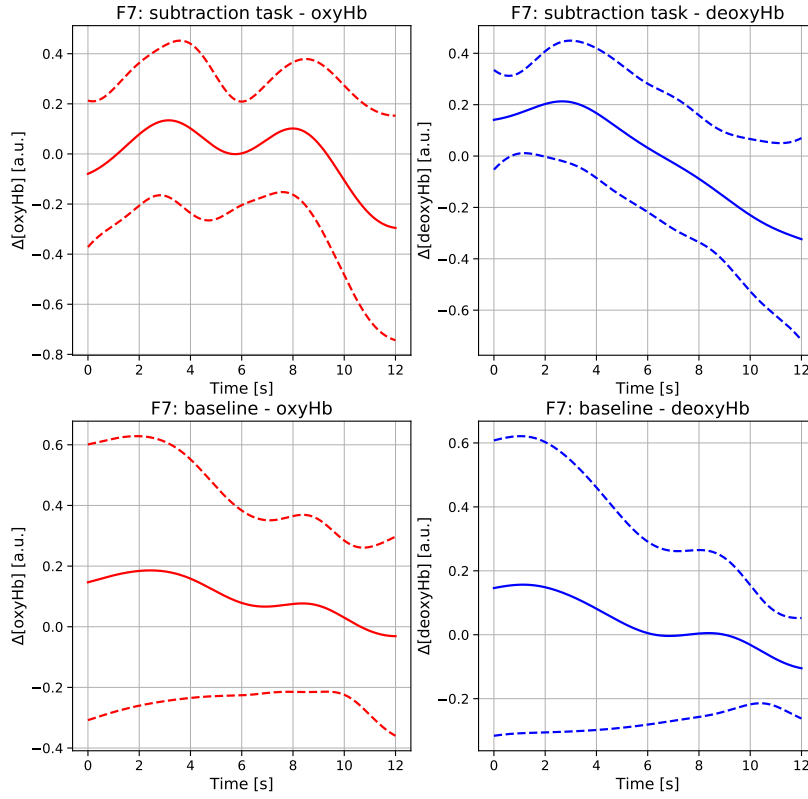
Figure 6.1: **Mean oxyHb and deoxyHb concentrations of the 8 subjects in the subtraction task and in the rest period (F7).** On the left side of the figure, the oxyHb signals and on the right side the deoxyHb signals are represented. In the same figure, the signals corresponding to the subtraction task, at the top, and, at the bottom, the signals corresponding to the subjects' rest period are represented. These signals were measured at position F7 of the 10-20 system. The dashed lines represent the standard deviation.

6.1.1 Classification

The classification between task and resting state was performed using datasets composed only by fNIRS features, only by EEG features and composed by fNIRS +EEG features. Classification was also conducted only for the n-back task and only for mental subtraction. For this purpose, two different approaches were followed, the Stratified k-Fold method and the Leave One Participant Out method. The final 15 features were selected by *SelectKBest* from *scikit-learn* library, and the chosen classifier was RF, as discussed in Chapter 5. Table 6.1 shows the classification results for the fNIRS, EEG and fNIRS+EEG datasets using the Stratified k-Fold method. According to this table, the hybrid fNIRS+EEG approach resulted in a predictive model as good as the one using only EEG features, achieving an accuracy score of 83.55 ± 1.33%, whereas using only

fNIRS scored an accuracy score of 80.05 ± 1.59%. This may be explained considering that the Select K Best algorithm selected the same features in both datasets EEG and fNIRS + EEG. Considering precision, the values were very similar to those of accuracy. The hybrid fNIRS+EEG approach and the one using only EEG features obtained 83.84 ± 1.18%, while fNIRS approach obtained 80.96 ± 0.98%. Recall values were higher, with the hybrid fNIRS+EEG and EEG approach reaching 97.66 ± 1.68% and the fNIRS approach reaching 97.17 ± 0.83%. These precision and recall values, along with the accuracy result, confirm an efficient classification model and a balanced dataset, i.e., with an adequate proportion of classes. Table 6.2 shows the classification results for the same datasets, but using the Leave One Participant Out method instead. Observing this table, we can conclude that with this, the dataset with exclusively fNIRS features is the one that achieves a better classification result, reaching an accuracy of 78.71 ± 3.72%, compared to the EEG dataset that obtains an accuracy of 73.98 ± 4.74% and the hybrid dataset that obtains an accuracy of 75.98 ± 3.68%. Once again, the precision and recall values, along with the accuracy result, validate this model, indicating a balanced dataset. This time, the features selected in the EEG-only and fNIRS + EEG approaches are not the same, as can be seen by the results in Table 6.2. Table 6.3 shows the classification results of fNIRS+EEG dataset for each participant. As can be seen from the table, the individual classifier, i.e., for each subject, is the one that obtains the best results, reaching accuracy values as high as 99.72 ± 0.83%, accompanied by precision values of 99.67 ± 1.00% and recall values of 100.00 ± 0.00%. This value of recall indicates that the model is fully able of finding all the relevant cases within the dataset. Both the n-back and subtraction tasks proved adequate for distinguishing between CW states, achieving accuracies of 76.16 ± 1.99% and 77.24 ± 3.43%, as shown in Table 6.4. Table 6.5 shows the classification results for the same datasets, but using the Leave One Group Out method instead. With this method the grading results are slightly down, to values of 70.77 ± 4.41% accuracy for n-back and 58.37 ± 4.59% for subtraction. Finally, Tables 6.6 and 6.7 presents the classification results for n-back and subtraction tasks, respectively. For this datasets, excellent results were obtained for participant #1 (100% for all metrics) as shown in the tables.

Table 6.1: **Classification results of fNIRS, EEG, and fNIRS+EEG datasets.** This approach used the Stratified 10-Fold Cross Validation method

Signal	Accuracy	F-Score	Precision	Recall	AUC-ROC
fNIRS	80.05 ± 1.59	88.33 ± 0.91	80.96 ± 0.98	97.17 ± 0.83	65.70 ± 5.04
EEG	83.55 ± 1.33	90.21 ± 0.81	83.84 ± 1.18	97.66 ± 1.68	82.50 ± 2.16
fNIRS + EEG	83.55 ± 1.33	90.21 ± 0.81	83.84 ± 1.18	97.66 ± 1.68	82.50 ± 2.16

6.2 Task Difficulty and Workload

This section shows the results of the differentiation between the workload experienced during the various levels of the n-back task. Figures 6.2 and 6.3 show the mean oxyHb and

Table 6.2: **Classification results of fNIRS, EEG, and fNIRS+EEG datasets.** This approach used the Leave One Participant Out Cross Validation method.

Signal	Accuracy	F-Score	Precision	Recall	AUC-ROC
fNIRS	78.71 ± 3.72	87.49 ± 2.34	80.10 ± 4.94	96.75 ± 2.60	64.87 ± 8.07
EEG	73.98 ± 4.74	84.51 ± 3.39	77.84 ± 4.51	93.09 ± 7.29	56.87 ± 14.45
fNIRS + EEG	75.98 ± 3.67	85.74 ± 2.75	78.87 ± 4.86	94.56±5.87	62.62± 7.89

Table 6.3: **Classification results for fNIRS+EEG of each participant.** The method of train and test split was Stratified k-Fold with k=10. There is no data of subtraction task from participant #5, and this may account for his worse results.

Participant	Accuracy	F-Score	Precision	Recall	AUC-ROC
1	99.72 ± 0.83	99.83 ± 0.51	99.67 ± 1.00	100.00 ± 0.00	99.30 ± 2.10
2	81.85 ± 4.17	89.41 ± 2.33	83.88 ± 3.11	95.82 ± 2.64	87.00 ± 7.89
3	88.18 ± 4.63	92.77 ± 2.77	89.01 ± 3.48	96.99 ± 3.67	87.30 ± 8.01
4	82.94 ± 3.04	89.58 ± 2.15	83.87 ± 1.91	96.27 ± 4.37	70.40 ± 10.04
5	68.08 ± 7.87	77.12 ± 7.55	72.85 ± 5.41	83.00 ± 12.59	67.00 ± 12.23
6	79.81 ± 5.81	87.73 ± 4.03	82.70 ± 1.94	93.63 ± 7.19	65.90 ± 11.83
7	82.53 ± 2.90	90.07 ± 1.80	83.79 ± 1.98	97.51 ± 3.64	70.30 ± 11.69
8	82.99 ± 4.29	89.78 ± 2.52	83.75 ± 3.58	96.83 ± 2.39	72.10 ± 10.52

Table 6.4: **Classification results of n-bask and subtraction datasets, using Stratified k-Fold Cross Validation.**

Signal	Accuracy	F-Score	Precision	Recall	AUC-ROC
n-Back	76.16 ± 1.99	84.10 ± 1.50	76.92 ± 0.98	92.80 ± 2.76	76.30 ± 3.98
Subtraction	77.24 ± 3.43	82.21 ± 2.48	77.28 ± 3.60	88.00 ± 3.48	83.60 ± 4.82

deoxyHb concentrations (from fNIRS)[1] at position F8 of the 10-20 system (right side) from the 8 subjects for the four n-back conditions. Looking at the signals in Figures 6.2 and 6.3, no clear trend is observed in the oxyHb and deoxyHb signals. In some cases, it is not even possible to assess the trend of the concentration of the chromophores, i.e., to understand whether their concentrations increase or decrease over time. Still, in some instances, as for task 0-back, oxyHb concentration seems to increase over time (see Figure 6.2), and deoxyHb concentration seems to decrease over time (see Figure 6.3), as would be expected in a situation of task engagement. However, in task 1-back, for example, this trend does not hold (see Figures 6.2 and 6.3). Thus, the graphical information provided by fNIRS is not sufficient to assess the cognitive state of the participant, nor to distinguish the workload at different levels of effort and therefore more data need to be included.

The medium waves of oxyHb and deoxyHb signals for the n-back conditions from the position F7 of the 10-20 system may be consulted in Appendix A. Figure 6.4 presents the percentage of missed answers, the percentage of wrong answers and the classification results achieved for each of the four n-back conditions assessed in this work. Through

---

[1]OxyHb and deoxyHb signals refer to the chromophores' concentration's variation signals.

Table 6.5: **Classification results of n-bask and subtraction datasets, using the Leave One Participant Out Cross Validation method.** The dataset included fNIRS+EEG features.

Signal	Accuracy	F-Score	Precision	Recall	AUC-ROC
n-Back	70.77 ± 4.41	80.71 ± 3.06	73.05 ± 5.05	90.62 ±4.57	68.13 ± 5.67
Subtraction	58.37 ± 4.59	68.27 ± 3.15	62.99 ± 4.27	74.98 ± 5.16	55.71 ± 8.01

Table 6.6: **Classification results for for the n-back task of each participant.** The dataset included fNIRS+EEG features. This approach used the Stratified-10 Fold Cross Validation method. If we notice, participant #5 shows exactly the same results as in Table 6.3, since he only has the n-back part of the task.

Participant	Accuracy	F-Score	Precision	Recall	AUC-ROC
1	100.0 ± 0.00	100.0 ± 0.00	100.0 ± 0.00	100.0 ± 0.00	100.0 ± 0.00
2	63.33 ± 7.98	73.60 ± 6.19	68.94 ± 6.00	79.40 ± 8.66	66.80 ± 8.16
3	85.94 ± 6.63	90.25 ± 4.12	87.78 ± 8.51	93.75 ± 5.59	93.10 ± 5.30
4	81.67 ± 6.54	86.76 ± 4.76	81.30 ± 7.17	93.90 ± 7.53	76.30 ± 9.27
5	68.08 ± 7.87	77.12 ± 7.55	72.85 ± 5.41	83.00 ± 12.59	67.00 ± 12.23
6	75.26 ± 9.29	82.62 ± 6.00	77.91 ± 7.50	88.24 ± 5.20	73.70 ± 12.01
7	76.11 ± 7.03	85.47 ± 4.20	78.79 ± 4.98	93.53 ± 4.22	69.50 ± 11.50
8	80.91 ± 6.36	85.89 ± 5.04	83.14 ± 5.16	89.62 ± 9.15	85.90 ± 8.88

Table 6.7: **Classification results for for the subtraction task of each participant.** The dataset included fNIRS+EEG features. This approach used the Stratified-10 Fold Cross Validation method. Participant #5 has no data regarding this task.

Participant	Accuracy	F-Score	Precision	Recall	AUC-ROC
1	100.0 + 0.00	100.0 + 0.00	100.0 + 0.00	100.0 + 0.00	100.0 + 0.00
2	83.24 + 6.85	87.33 + 5.05	84.32 + 8.36	91.09 + 5.38	91.90 + 6.01
3	83.01 + 6.62	85.86 + 5.72	83.89 + 7.84	88.64 + 8.26	81.64 + 6.70
4	75.06 + 5.31	80.05 + 3.94	78.52 + 8.98	83.79 + 10.59	77.10 + 8.77
6	69.04 + 8.75	75.65 + 7.27	71.23 + 7.71	82.27 + 13.03	69.60 + 15.29
7	78.63 + 5.90	83.59 + 4.26	78.84 + 8.01	90.00 + 7.55	75.50 + 9.82
8	89.62 + 7.98	90.77 + 7.41	93.94 + 9.08	90.00 + 13.76	97.80 + 2.71

the graphs in the Figure 6.4, we can observe that the percentage of answers not given was higher in the first level of difficulty of the n-back task, probably because it is the level of habituation of the subjects to the task. In fact, in the following levels these values decreased, except for the last level, where the values increased again. This time, this increase may be due to the degree of difficulty of the task. Concerning the number of wrong answers in the n-back task, these increase with increasing task difficulty. This seems to suggest that CW associated with the complexity of cognitive tasks worsens the performance. Considering the classification results in the four phases of the n-back task, the performance of the classifier does not seem to be influenced by the complexity and difficulty of the cognitive task.

The EEG bands powers were calculated for comparison between the multiple levels of n-back task. Table 6.8 presents the MTLI results, a version of TLI which is widely used
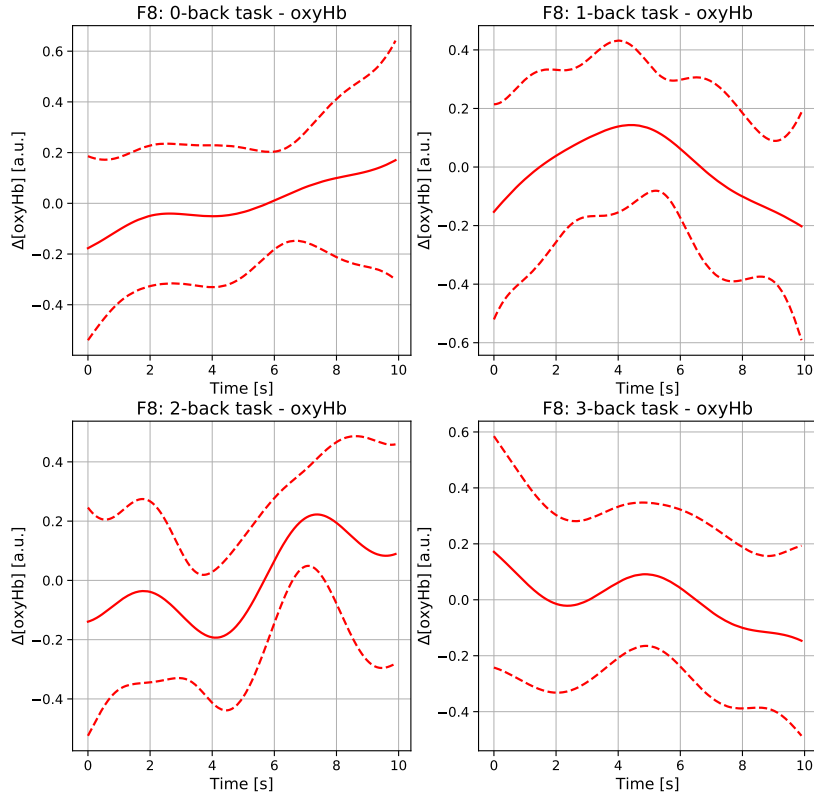
Figure 6.2: **OxyHb signals for 0/1/2 and 3-back tasks (F8).** The signals are the mean waves of oxyHb signals (fNIRS) for 0/1/2 and 3-back tasks at position F8 of the 10-20 system. The dashed lines represent the standard deviation.

to quantify CW [23, 37, 44, 45]. According to this table, the CW level seems to slightly increase with increasing task difficulty and demand on both analyzed sensor positions. However, on the left side, the CW value decreases slightly at level 3-back compared to level 2-back, while on the right side, this value diminishes in the transition from 1-back to 2-back. The complete table with all calculated powers is available for consultation at Appendix A (see Tables A.4 and A.5).

Table 6.8: **MTLI for 0/1/2 and 3-back tasks.** Alpha and theta powers for the MTLI calculation were measured in positions F7 and F8 of the 10-20 system.

MLTI	F7	F8
0-back	2.05 ($p = 0.007$)	3.54 ($p = 0.04$)
1-back	2.69 ($p = 0.04$)	3.92 ($p = 0.04$)
2-back	3.54 ($p = 0.008$)	3.76 ($p = 0.01$)
3-back	3.44 ($p = 0.008$)	4.86 ($p = 0.007$)
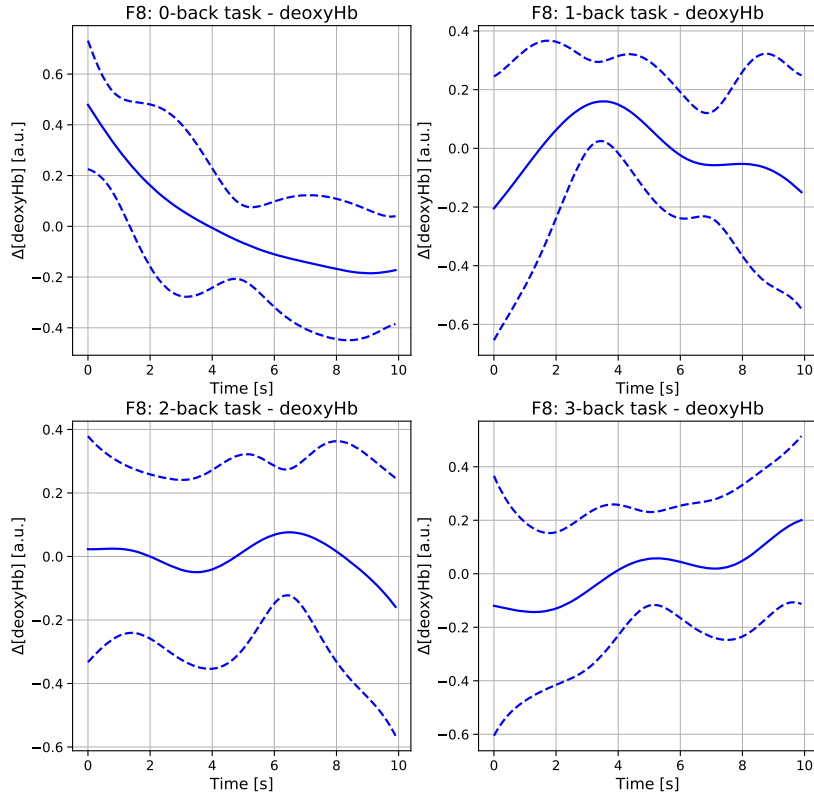
Figure 6.3: **DeoxyHb signals for 0/1/2 and 3-back tasks (F8).** The signals are the mean waves of deoxyHb signals (fNIRS) for 0/1/2 and 3-back tasks at position F8 of the 10-20 system. The dashed lines represent the standard deviation.

Tables 6.9, 6.10, 6.11 and 6.12 present the results of the distinction between rest/0/1, 0/1, 0/1/2 and rest/0/1/2/3 back levels, respectively. The results suggest the applicability of fNIRS and EEG techniques for discrimination between distinct levels of difficulty.

According to the Table 6.10, we can see that some participants have very disparate ranking results (see participants #2 and #5). The standard deviation values are justified by the reduced number of samples in each dataset and the variability of the physiological sensors. The overall precision and recall values indicate that the datasets are balanced, containing a similar number of samples from each class. However, overall results highlight the applicability of fNIRS and EEG techniques for discrimination between distinct levels of difficulty, and CW. Regarding Table 6.11, distinguishing between all 3 levels of difficulty results in worse results than distinguishing between only 2. According to the Table 6.12, the distinction between the 4 levels of difficulty results in the lowest scores. However, participant #3 demonstrates that this discrimination is possible.
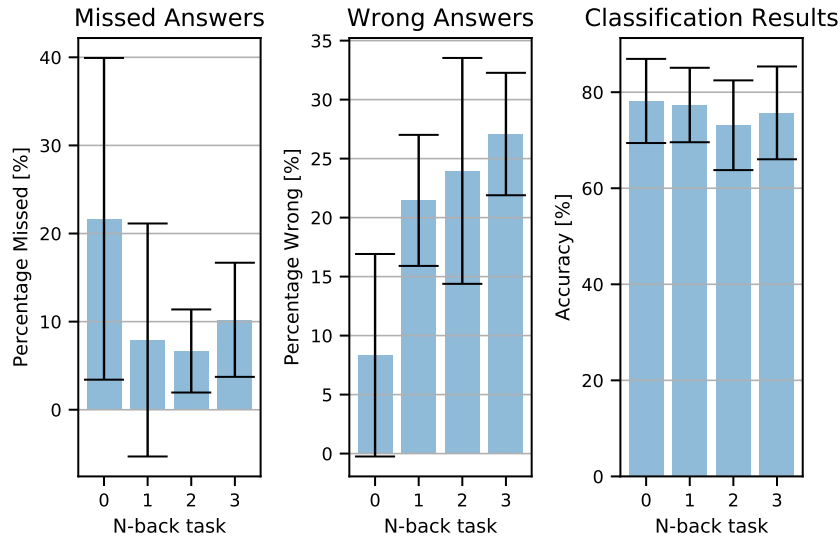
Figure 6.4: **Bar charts with the results for the n-back task.** The first bar chart presents the percentage of missed answers, the second one presents the percentage of wrong answers and the third one presents the classification accuracies obtained for the four n-back conditions.

Table 6.9: **Classification results for the distinction between rest, 0 and 1-back levels for each participant.** The dataset included fNIRS+EEG features. This approach used the Stratified-10 Fold Cross Validation method.

Participant	Accuracy	F-Score	Precision	Recall
1	86.67 + 5.83	85.33 + 13.53	87.69 + 10.80	86.61 + 20.90
2	83.90 + 5.47	79.74 + 24.88	80.46 + 23.47	83.77 + 30.65
3	77.39 + 4.26	74.21 + 25.17	75.24 + 19.94	77.45 + 31.44
4	84.87 + 6.46	81.16 + 27.69	80.02 + 26.80	84.73 + 30.70
5	82.21 + 3.71	78.21 + 22.71	80.47 + 19.98	82.37 + 30.21
6	80.00 + 3.87	74.75 + 30.92	71.83 + 28.82	80.33 + 35.54
7	71.58 + 8.78	66.09 + 27.77	69.11 + 25.97	71.90 + 36.80
8	75.05 + 5.63	71.72 + 27.02	72.34 + 27.77	74.79 + 32.91

6.3   Cognitive States and Workload

This section presents the results of the relationship between the cognitive states - boredom, frustration and interest - and cognitive workload. Does boredom or frustration worsen cognitive performance? Does interest in the content help in its learning? These are some of the questions that motivate the search for understanding this relationship.

Figures 6.5, 6.6 and 6.7 show the mean oxyHb and deoxyHb concentrations of the 8 subjects during the states of boredom, frustration and interest, respectively, for positions F7 and F8 of the 10-20 system. From the figures, we conclude that the graphical information provided by fNIRS is not sufficient to assess any of the cognitive states since we can not immediately ascertain the variation in chromophores concentrations.

Table 6.10: **Classification results for the distinction between 0 and 1-back levels for each participant.** The dataset included fNIRS+EEG features. This approach used the Stratified-10 Fold Cross Validation method.

Participant	Accuracy	F-Score	Precision	Recall
1	81.56 ± 13.34	81.20 ± 15.25	82.73 ± 17.07	81.51 ± 17.63
2	51.50 ± 21.45	49.22 ± 26.78	50.35 ± 29.39	52.92 ± 30.71
3	78.75 ± 12.56	77.40 ± 15.85	82.39 ± 17.50	78.95 ± 23.28
4	75.00 ± 18.97	71.27 ± 28.29	73.58 ± 31.14	75.91 ± 33.30
5	91.67 ± 08.33	90.99 ± 09.91	93.71 ± 11.19	91.22 ± 15.60
6	57.67 ± 19.21	53.92 ± 26.44	57.05 ± 31.09	57.31 ± 32.56
7	79.00 ± 10.44	78.79 ± 10.72	81.64 ± 15.18	79.83 ± 17.11
8	69.67 ± 13.48	68.89 ± 14.23	72.71 ± 18.50	69.74 ± 20.90

Table 6.11: **Classification results for the distinction between 0, 1 and 2-back levels for each participant.** The dataset included fNIRS+EEG features. This approach used the Stratified-10 Fold Cross Validation method.

Participant	Accuracy	F-Score	Precision	Recall
1	76.54 ± 12.77	76.11 ± 18.81	80.07 ± 20.66	76.98 ± 22.53
2	60.22 ± 9.37	56.16 ± 25.55	58.40 ± 29.14	61.31 ± 33.41
3	74.17 ± 10.83	72.55 ± 19.84	75.83 ± 20.61	74.12 ± 25.39
4	74.03 ± 19.15	73.53 ± 29.78	75.99 ± 32.12	74.61 ± 31.32
5	66.45 ± 7.08	62.93 ± 22.48	67.56 ± 27.25	65.64 ± 28.48
6	63.00 ± 14.87	60.07 ± 28.13	59.82 ± 29.72	63.79 ± 31.83
7	39.37 ± 11.20	39.01 ± 18.89	40.52 ± 21.09	39.97 ± 21.65
8	54.49 ± 12.15	52.58 ± 21.92	54.40 ± 24.91	54.22 ± 24.51

Table 6.12: **Classification results for the distinction between 0, 1,2 and 3-back levels for each participant.** The dataset included fNIRS+EEG features. This approach used the Stratified-10 Fold Cross Validation method.

Participant	Accuracy	F-Score	Precision	Recall
1	51.78 ± 8.97	50.72 ± 26.12	53.93 ± 30.39	52.32 ± 27.64
2	52.31 ± 8.86	50.12 ± 21.57	54.79 ± 28.27	52.31 ± 25.60
3	71.25 ± 12.25	69.63 ± 24.29	71.86 ± 23.83	71.64 ± 28.76
4	52.31 ± 12.66	49.90 ± 23.06	54.57 ± 28.47	52.26 ± 28.56
5	55.33 ± 12.07	54.05 ± 22.02	60.15 ± 29.06	55.06 ± 24.40
6	62.62 ± 10.32	59.87 ± 25.88	62.72 ± 28.99	63.19 ± 31.48
7	52.73 ± 6.49	50.18 ± 21.84	51.94 ± 22.70	52.51 +± 26.10
8	30.06 ± 6.70	28.73 ± 15.67	31.22 ± 19.09	30.41 ± 19.56

The MTLI was estimated for the states of boredom, frustration and interest. This index is calculated from the theta and alpha power of the EEG. Table 6.13 presents the results of this calculation. According to the table, there is no significant relationship between the alpha and theta powers within each cognitive state. Thus, no conclusions can be drawn from this MTLI analysis.

For the regression technique, we used the labelling process to construct the regression

Figure 6.5: **Mean oxyHb and deoxyHb concentrations of the 8 subjects during boredom state (F7 and F8).** On the left side of the figure, the oxyHb signals and on the right side the deoxyHb signals are represented. In the same figure, the signals corresponding to the position F7 of the 10-20 system, at the top, and, at the bottom, the signals corresponding to the position F8. The dashed lines represent the standard deviation.

index. In which, a value of 0 would mean not at all bored/frustrated/interested and 1 fully bored/frustrated/interested. We thus created 4 regressors, one for each of the cognitive states and another that joined the 3 states together.

Table 6.14 exhibits the results of regression for the label boredom. As can be seen from the table, using all subjects data, results in a R^2 of 0.21, a MSE of 0.16 and a MAE of 0.32. The MSE and MAE values are a little higher than desired perhaps because there is a large variability in fNIRS and EEG measurements between subjects. Still, participant #5 achieves an R^2 of 0.21, participant #6 achieves an R^2 of 0.17 and participant #8 achieves an R^2 of 0.19.

Table 6.15 exhibits the results of regression for the label frustration. As can be seen from the table, using all subjects data, results in a R^2 of 0.19, a MSE of 0.11 and a MAE of 0.24. In this case, the participants were only 6, since the remaining 2 participants
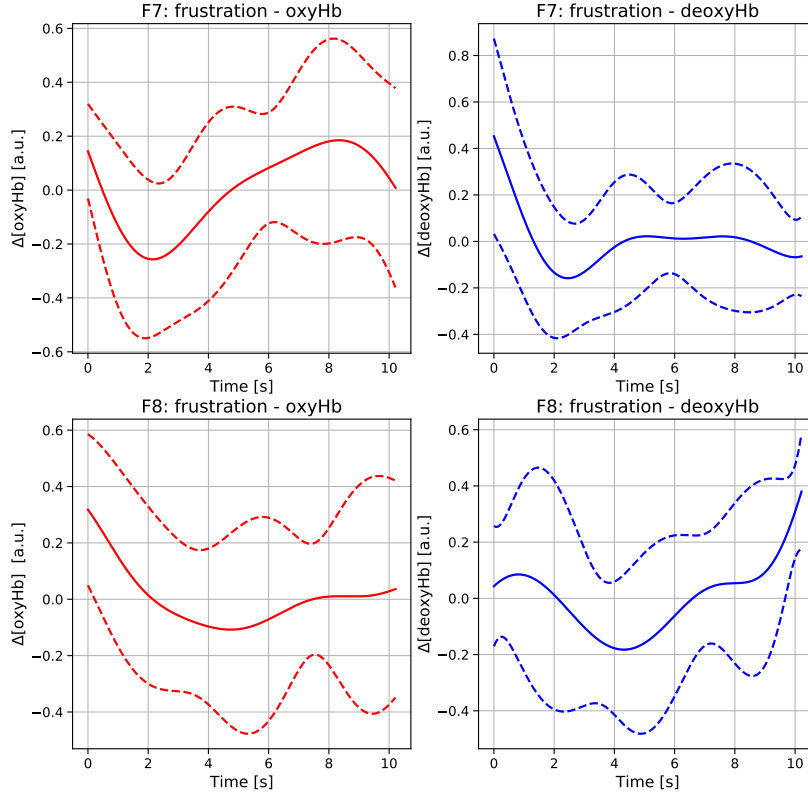
Figure 6.6: **Mean oxyHb and deoxyHb concentrations of the 8 subjects during frustration state (F7 and F8).** On the left side of the figure, the oxyHb signals and on the right side the deoxyHb signals are represented. In the same figure, the signals corresponding to the position F7 of the 10-20 system, at the top, and, at the bottom, the signals corresponding to the position F8. The dashed lines represent the standard deviation.

showed no facial and bodily evidence of this cognitive state. The MSE and MAE values are a little higher than desired perhaps because there is a large variability in fNIRS and EEG measurements between subjects. Still, participant #1 achieves an R^2 of 0.32 and a low MSE of 0.11 (considering the low number of samples), participant #3 achieves an R^2 of 0.22 and participant #6 achieves an R^2 0.17 and a low MSE of 0.11.

Table 6.16 exhibits the results of regression for the label interest. As can be seen from the table, using all subjects data, results in a R^2 of 0.12, a MSE of 0.21 and a MAE of 0.43. The MSE and MAE values are a little higher than desired perhaps because there is a large variability in fNIRS and EEG measurements between subjects. Still, participant #1 achieves an R^2 of 0.40, and participant #4 achieves an R^2 of 0.18.

Considering Tables 6.14, 6.15 and 6.16, the R^2 value can be justified considering that this is a study of human behaviour and that the variables used in the regression are highly
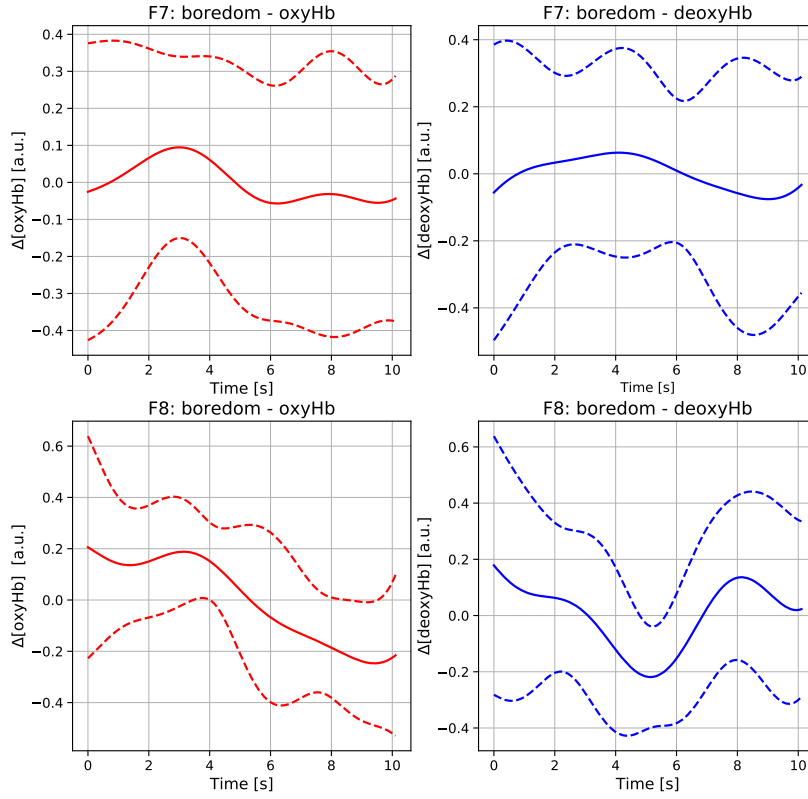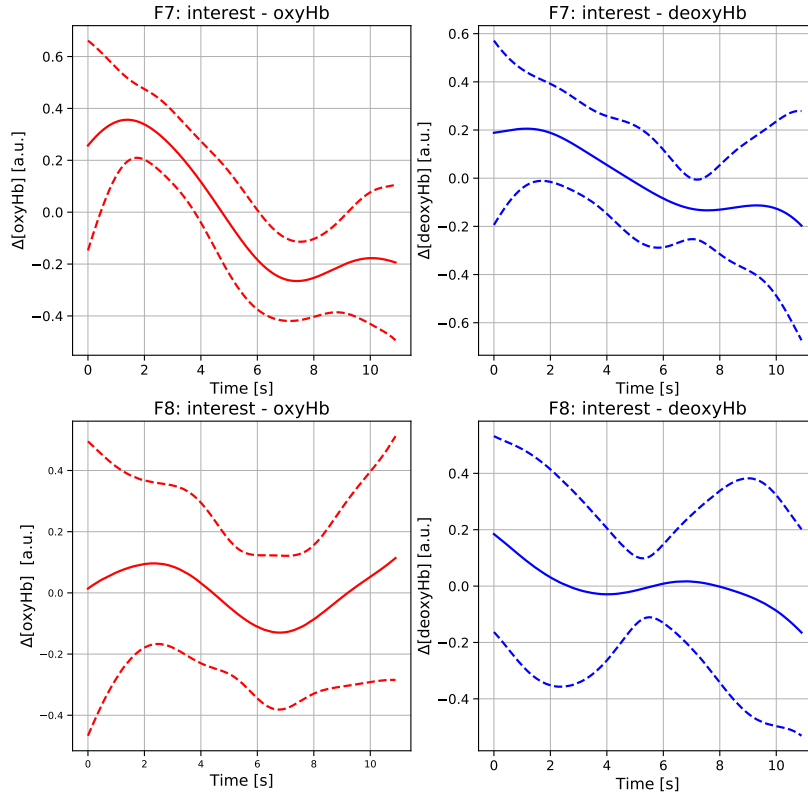
Figure 6.7: **Mean oxyHb and deoxyHb concentrations of the 8 subjects during interest state (F7 and F8).** On the left side of the figure, the oxyHb signals and on the right side the deoxyHb signals are represented. In the same figure, the signals corresponding to the position F7 of the 10-20 system, at the top, and, at the bottom, the signals corresponding to the position F8. The dashed lines represent the standard deviation.

subjective. Regarding all metrics, the fact that the subjects had to wear a mask due to the current pandemic also hampered the facial expression assessment in the labelling process when performing this regression and may help to justify these low results.

Table 6.17 presents the engagement index results and compares them to the R^2 of the regression for label interest. Through the table, we can see that the engagement index obtained by the EEG rhythms powers does not directly relate to the R^2 for the regression of the label interested.

Table 6.13: **MTLI for different the cognitive states.** The Modified Task Load Index (MLTI) was calculated using theta and alpha powers at F7 and F8 positions of the 10-20 system. In the table the values of the powers presented are relative to F7/F8 positions. P-values are also showed for statistical relevance assessment. P(Theta) stands for theta power and P(Alpha) stands for alpha power. The unit of power is μV^2.

Cognitive State	P(Theta)	P(Alpha)	MLTI
Boredom	3.65±4.04/5.81±6.71	2.11±0.94/1.89±1.21	1.73 (p=0.19)/3.07 (p=0.25)
Frustration	6.34±3.61/7.75±7.71	3.39±2.34/3.03±2.58	1.87 (p=0.06)/2.56 (p=0.09)
Interest	4.46±4.83/4.23±5.10	3.10±1.63/3.02±2.00	1.44 (p=0.64)/1.40 (p=0.84)

Table 6.14: **Regression results of the label boredom.** The table shows the regression results of the label boredom, using the dataset composed by all participants and for each of the participants. The method Stratified 10-Fold Cross Validation was used. ET stands for Extra Trees, described in Chapter 5.

Dataset	Method	R^2	MSE	MAE
All subjects	ET	0.21	0.16	0.32
Subject 1	ET	0.10	0.02	0.05
Subject 2	ET	0.12	0.19	0.40
Subject 3	ET	0.14	0.09	0.19
Subject 4	ET	0.08	0.10	0.20
Subject 5	ET	0.21	0.19	0.39
Subject 6	ET	0.17	0.15	0.30
Subject 7	RF	0.03	0.22	0.44
Subject 8	ET	0.19	0.19	0.40

Table 6.15: **Regression results of the label frustration.** The table shows the regression results of the label frustration, using the dataset composed by 6 participants and for each of the participants. The method Stratified 10-Fold Cross Validation was used. ET stands for Extra Trees, described in Chapter 5.

Dataset	Method	R^2	MSE	MAE
All subjects	ET	0.19	0.11	0.24
Subject 1	ET	0.32	0.11	0.24
Subject 3	ET	0.22	0.17	0.36
Subject 4	ET	0.04	0.07	0.15
Subject 5	ET	0.04	0.03	0.06
Subject 6	ET	0.17	0.11	0.25
Subject 7	ET	0.01	0.18	0.36

Table 6.16: **Regression results of the label interest.** The table shows the regression results of the label interested, using the dataset composed by all participants and for each of the participants. The method Stratified 10-Fold Cross Validation was used. ET stands for Extra Trees, described in Chapter 5.

Dataset	Method	R^2	MSE	MAE
All subjects	ET	0.12	0.21	0.43
Subject 1	ET	0.40	0.15	0.33
Subject 2	ET	0.11	0.20	0.41
Subject 3	ET	0.03	0.23	0.46
Subject 4	ET	0.18	0.18	0.37
Subject 5	RF	0.04	0.22	0.45
Subject 6	RF	0.11	0.22	0.45
Subject 7	ET	0.05	0.18	0.37
Subject 8	RF	0.03	0.16	0.32

Table 6.17: **Engagement index results.** The table shows the results of the engagement index, obtained from the expression Beta Power/(Theta Power + Alpha Power), for each participant. It also presents the R^2 of the regression for comparison. P(Beta) stands for Beta Power; P(Theta) stands for Theta Power; P(Alpha) stands for Alpha Power; Reg. stands for Regression. The unit of power is μV^2.

Subject	P(Beta)	P(Theta)	P(Alpha)	Engagement Index	Interest Reg. R^2
1	11.47/8.99	0.77/0.69	5.67/5.77	1.78/1.39	0.40
2	0.78/0.74	0.05/0.06	0.96/0.67	0.77/1.0.1	0.11
3	8.59/7.97	0.57/0.77	4.29/4.64	1.76/1.47	0.03
4	13.14/6.34	2.47/0.54	1.82/0.48	3.06/6.21	0.18
5	1.32/1.36	3.74/6.20	1.15/1.29	0.27/0.18	0.04
6	3.06/1.64	13.31/9.11	4.41/4.13	0.19/0.12	0.11
7	8.29/8.37	11.80/15.05	4.12/5.20	0.52/0.41	0.05
8	4.60/6.75	3.00/1.42	2.42/2.03	0.85/1.96	0.03

70

DISCUSSION

In addition to causing stress and mental exhaustion, increased CW can lead to human error in critical tasks, jeopardising safety, mental health and work performance. Its study is essential for BCI applications and in e-learning platforms conception. In our study, we, therefore, sought to study and quantify the CW event.

Briefly, to differentiate between cognitive states during mental task and rest, i.e. baseline, we acquired fNIRS and EEG signals, while participants completed the n-back and mental subtraction tasks. For the analysis of the relationship between cognitive task complexity and CW, we distinguished between the various levels of the n-back task. Related to this analysis, we also used the EEG signals for the calculation of MTLI, as an estimate of CW through spectral analysis of brain rhythms. Finally, to quantify the cognitive states of boredom, frustration and interest, we built a regression that combined temporal, statistical and spectral information from fNIRS and EEG signals and the powers of brain rhythms. The major findings are: (i) despite the small sample size considered in this study, our results seem to suggest that the combination of information obtained by fNIRS and EEG and ML algorithms allows us to effectively distinguish the cognitive state of mental tasks from the rest and (ii) it seems possible to build a cognitive states estimator based on fNIRS and EEG data and an ML approach. These results support the hypothesis that the fNIRS system is a suitable technique for CW monitoring. It can also be allied to the EEG system and ML algorithms for a more solid study of the phenomenon. We discuss these findings below, interpreting them in more detail.

7.1 Distinguishing between cognitive task and rest

To differentiate the cognitive state of mental task from rest (baseline) we resorted to visual inspection of fNIRS signals of the two modes, we used the powers of the EEG bands

to estimate the MTLI in the two cases and, finally, we classified the two states, using ML techniques.

The fNIRS and EEG measurements were acquired at the PFC, at positions F7 and F8 of the 10-20 system. Since it is widely established that the frontal lobes are commonly involved in complex tasks, involving novel demands or requiring considerable attention level [35].

Due to the shape and the standard deviations of the mean variation of concentration of the chromophores of haemoglobin (see Figure 6.1), it is not possible to predict whether subjects are tackling the mental subtraction task or are in a period of rest. Thus, the visual information offered by fNIRS is not sufficient to distinguish between cognitive states of mental task and rest. In effect, several studies that have attempted to establish a consistent relationship between CW and haemodynamic changes have failed [35].

Regarding spectral analysis, both beta and delta power stand out in an attempt to differentiate the cognitive states of mental subtraction and the rest period (see Table A.2). The high standard deviation value may be explained by the fact that we are analysing physiological measures of great inter-subject variability (fNIRS and EEG) and considering a very small sample size (of only 7 subjects in this specific case). If we do not consider these standard deviations, we can observe that beta power in the resting state is almost half (54%) of the beta power of the cognitive task state. In truth, beta rate is an indicator of alertness, visual attention and WM, increasing in more complex visual WM tasks and concentration [40, 41, 44]. Also, the delta power of the cognitive task state is almost half (54%) of the delta power in resting state. In fact, delta power is associated with deep sleep states and indicates less focus and attention maintenance [11, 37, 39, 40]. Delta power is also higher for the mental subtraction compared to the n-back task (see Tables A.2, A.4 and A.5). This observation is in line with study [125], which claimed an increase in delta power during complex tasks, such as arithmetics. Work [11] also agrees that sustained attention tasks intensify this power. In sum, these results corroborate what is expected in the literature. It was also not possible to calculate the MTLI for the cognitive task and rest states since the difference between the two did not prove to be statistically significant (see Table A.3).

However, the classification results demonstrate that the data acquired with fNIRS and EEG combined with ML algorithms allow to successfully distinguish the cognitive states of mental task and rest (see Tables 6.1, 6.2, 6.3, 6.4, 6.5, 6.6 and 6.7). The average of the individual results is 83.26 ± 8.22 %. In this case, the standard deviation is slightly higher since we are studying physiological measures with a lot of inter-subject variability and considering a reduced number of samples. In the classification using all participants, the hybrid approach, i.e., the one using fNIRS and EEG features, and the one using only EEG features obtained equal and better results than the alternative of using only fNIRS features, achieving an accuracy of 83.55 ± 1.33% against 80.05 % ± 1.59% (see Table 6.1). Once the Select K Best algorithm chose the same 15 features for both EEG and fNIRS + EEG datasets. These results lead us to believe that this classification system

could be of both types (i) user-tuned, i.e., a user-dependent system prepared for each user, which would require a calibration process before use and (ii) generically-tuned, a more comprehensive system that does not require a calibration process and is prepared for various types of user. Also in study [23], their results suggested that the hybrid approach should be preferred over only fNIRS or EEG in developing BCI and other application needing to monitor users' CW.

The EEG+fNIRS modality is a promising alternative as a more accurate practical technique than individual modalities alone, presenting a strong potential for future neurorehabilitation and neurofeedback applications [19]. Functional Near-Infrared Spectroscopy complements EEG by quantifying CBF changes via a CBF light source/detector on the scalp. Functional Near-Infrared Spectroscopy has no electromyography (EMG) or blinks artefacts, and its signal approaches the blood oxygen level-dependent (BOLD) signal of functional Magnetic Resonance Imaging (fMRI), the gold standard of cerebral haemodynamics measurement. Moreover, EEG+fNIRS introduces new types of features based on NVC, which are not exclusively originated by neural activity (such as EEG) or by haemodynamics (such as BOLD) [23].

On the other hand, considering Leave One Participant Out Cross Validation, the best result is obtained for the modality using only fNIRS features, reaching an accuracy of 78.71 ± 3.72% (see Table 6.2). This results suggests that fNIRS may be the most generalizable solution.

The highest score, with 100% accuracy (participant #1 in Tables 6.6 and 6.7), was obtained in the individual approach for the separated n-back and subtraction tasks, using fNIRS + EEG features and the method Stratified 10-Fold (see Tables 6.6 and 6.7).

7.1.1 Is it possible to distinguish between cognitive task and rest by combining fNIRS and EEG information with ML techniques?

As observed, ML algorithms can successfully distinguish the cognitive state associated with mental tasks from that associated with rest, contrary to what was verified using only visual inspection of fNIRS signals and using only spectral analysis of EEG bands. Our best result was for the individual classification of the n-back and subtraction tasks, where 100% was achieved in all metrics (see participant #1 in Tables 6.6 and 6.7). The complexity of ML algorithms allows them to identify patterns that we can not with the naked eye. We are limited to our two-dimensional perception of the physiological signal, opting for the most basic features, such as slope, to assess the trend of the chromophores. Machine Learning algorithms, on the other hand, approach the problem in its 15 dimensions (i.e. the 15 selected features), using temporal, statistical and spectral-domain tools and perceiving the issue in a way that we would never be able to. In addition, the reduced number of samples considered does not allow assumptions regarding the spectral study of the EEG powers.

7.2   Relationship between task complexity and CW

To study how the difficulty of a mental exercise affects CW, we used visual inspection of the variation of mean oxyHb and deoxyHb concentrations for the 4 levels of the n-back task (with increasing difficulty), calculated the MTLI for the 4 cases by spectral analysis of the powers of the EEG bands, compared the percentage of missed and wrong answers with the classification results at the different levels of the task and, finally, we discriminate the various levels of the n-back task, with the help of ML techniques.

The signals of the oxyHb and deoxyHb chromophores obtained through fNIRS do not allow their behaviour assessment (see Figures 6.2 and 6.3). Recognising patterns or trends of chromophores behaviour through visual inspection is not possible. Since it is fundamental for the mental task and baseline categorisation, the visual information obtained by fNIRS is not sufficient to study and understand the relationship between the difficulty along the various levels of the mental task and the CW expressed by brain activation.

The results of Table 6.8 indicate that there is a significant relationship between the alpha and theta powers required to compute MTLI ($p < 0.05$). However, the standard deviations of the alpha and theta powers necessary to this index estimation are too high (see Tables A.4 and A.5). We may justify this standard deviation by the high variability associated with physiological fNIRS and EEG measurements and the small number of samples considered. There seems to be an increasing tendency of CW along with the levels of the task and of its difficulty (see Table 6.8), nonetheless, the standard deviation values do not allow any conclusion.

Considering the graph in the Figure 6.4 and the classification results in the four phases of the n-back task, the classifier performance does not seem to be influenced by the complexity of the cognitive task. Although the number of wrong answers increases throughout the levels, the classification results remain around 80%. This indicates that wrong answer does not mean a lack of attention on the part of the participant. Therefore, even if the participant gets it wrong or don't answer, he can be focused on the cognitive task, and it may be enough to distinguish it from the resting state, where we have no concrete focus.

The results of the attempted discrimination between various levels of difficulty of the n-back task suggest that fNIRS and EEG techniques are capable of monitoring various CW stages (see Tables 6.10, 6.11, and 6.12). The differences found in some participants (p.e., see participants #1 and #6 in Table 6.10) may be justified by the neural efficiency hypothesis of intelligence. According to this theory, some individuals will have to allocate a substantial amount of mental resources for a given task completion and a certain level of performance, while others will achieve the same results with much less mental effort [126].

7.2.1 Does an increase in cognitive task difficulty rise CW?

Despite the small number of subjects in our sample, which does not allow us to know whether an increase in complexity in a task increases the CW level experienced, our results seem to indicate that it does. In fact, both studies [100] and [102] found consistent increases in brain activation as a function of memory load in PFC, while performing n-back task.

7.3 Influence of boredom, frustration and interest on the level of CW experienced

To understand cognitive states' influence on CW, we visually inspected the fNIRS signals and analysed the spectral content of the EEG bands, estimating MTLI for the different states. Finally, we used a regression technique for the quantification of boredom, frustration and interest states. The regression index was obtained through a labelling process that consisted of, through the breath sensor, response time and facial expression, assess the cognitive state of the subjects as bored, frustrated, interest or at rest. Looking at the oxyHb and deoxyHb signals during boredom, frustration and interst modes, we conclude that it is not possible to directly access the subject's cognitive state from the fNIRS data (see Figures 6.5, 6.6 and 6.7). Thus, we conclude that the graphical information offered by fNIRS is insufficient for cognitive modes assessment and the understanding of its relation with CW. Alpha and theta powers for the different cognitive states showed no significant relationship ($p > 0.05$) (see Table 6.13). Thus, to continue this spectral analysis, we would need to increase the number of participants in the sample. In light of the results obtained for the regression technique, there are indications that it is possible to quantify the states of boredom, frustration and interest by combining brain information provided by fNIRS and EEG, breathing and facial expression with ML algorithms (see Tables 6.14, 6.15 and 6.16). Increasing the number of physiological sensors, such as including Electrocardiography (ECG) and electrodermal activity (EDA) sensors, as well as enlarging the sample size would allow improving the regression results. The engagement index, obtained by analyzing the alpha, beta and theta powers, indicates alertness, attention and task engagement [44]. We calculated this index individually on the parts of the signal that we labelled interest, expecting a relationship between it and the R^2 scores of the regressor of interest. However, we found no relationship between them (see Table 6.17).

7.3.1 Is it possible to build a cognitive state calculator combining information from fNIRS and EEG with ML algorithms?

Our results indicate that it is possible to construct a cognitive state calculator, based on a regression approach combining the brain information from fNIRS and EEG, respiration

and facial expression. However, a larger sample of participants would be needed to prove this theory.

CONCLUSIONS

Cognitive Workload is the amount of mental effort and cognitive resources imposed to complete a task within a limited time period [9]. It increases accordingly to task difficulty, and an increased amount may lead to sub-optimal results. However, low CW also relates to boredom, distraction and human errors [44]. Thus, optimizing this amount is crucial for performance enhancing [7, 12, 35, 44].

The study and quantification of the CW phenomenon is of interest for the design of e-learning platforms [2, 12]. The world is moving in a digital direction. Furthermore, the pandemic we have experienced has accelerated this transition. Home-working and home-learning are increasingly common realities. In the field of digital education, an optimal amount of CW is essential for effective learning. In addition to CW, the cognitive states of the learner also influence the assimilation of the taught contents [13–16]. While boredom may shorten attention spans, interest might increase them and motivate the understanding of a given subject [13, 14].

Furthermore, fNIRS has proven to be a promising tool for CW analysis and in the context of BCI research, due to its simple setup and robust results [18, 19, 36]. Brain-computer interfaces are the hope of people with severe motor disabilities to communicate again or use an injured limb, so their applications are of immeasurable and incalculable value [19, 20].

Thus, the research conducted in this thesis attempted to investigate the applicability of the fNIRS technique in monitoring different CW states through a reduced number of channels. In addition, it aimed at developing a methodology to stimulate different CW levels, comprising the relative concentrations estimate and identification of features that enable these levels' discrimination.

To this end, our experimental procedure included a n-back and a mental subtraction task and a Python tutorial, attempting to simulate the learning process. Both the n-back

and subtraction tasks were shown to create haemodynamic responses in PFC, our area of interest, which is also the most common in BCI research [18]. In addition, the 4-level n-back task aimed to analyse the relationship between the difficulty of the cognitive task and the level of CW experienced. The brain information we collected from the participants throughout these exercises, through fNIRS and EEG, allowed us to analyse and answer three questions:

1. Is it possible to detect mental states changes using both fNIRS and EEG data?

To answer, we (i) visually analysed the variation of average chromophorus concentrations relative to the mental subtraction task and relative to the rest period (baseline) and (ii) implemented a classification system to automatically distinguish between cognitive task and rest. In this respect, we built several dataset versions in the search for the optimal classifier: using all participants, using one by one, using only the n-back task, using only the subtraction task and using two different cross-validation methods, Stratified 10-Fold and Leave One Participant Out.

2. Can we discriminate between the different levels of CW experienced during the n-back task by combining fNIRS and EEG information?

To answer, we (i) visually analysed the fNIRS signals for each difficulty level of the n-back task, (ii) calculated the MTLI, through the alpha and theta powers measured with EEG, for each level, allowing us to assess the level of CW experienced at each difficulty stage and (iii) we distinguish the various levels of the n-back task, using fNIRS + EEG features and ML algorithms.

3. Are we able to quantify the cognitive states, i.e., how much a subject is bored, frustrated or interested in front of a particular task?

To answer, we (i) visually analysed the fNIRS signals corresponding to the participants' states of boredom, frustration and interest and (ii) built a regressor that included the features used in the classification between task and baseline and new features with the powers of the EEG bands, to quantify the level of boredom, frustration and interest felt by the participants during the experimental procedure used. For the regression index, we performed a labelling process that consisted of, through the breath sensor, response time and facial expression, assessing the cognitive state of the subjects.

In response to these questions, we conclude that (1) an ML approach using fNIRS+EEG features seem to successfully distinguish between users' cognitive states associated with mental tasks and rest; (2) although the small sample size considered in this work does not allow us to properly answer the question using spectral analysis, some of our results from classification indicate that yes. In the classification process, we used ML algorithms and fNIRS+EEG features to discriminate between different levels of task (e.g. 0-back vs 1-back), as was done in [23]; and (3) evidence shows that an ML approach using the regression technique and combining features of fNIRS+EEG can quantify the boredom, frustration and interest of users when performing cognitive tasks.

In this way, the classifiers and regressors architected in this thesis can be used and optimised for e-learning applications, whose monitoring of the user's cognitive states will

allow the adaptation in real-time of the lesson contents presented, inciting the student's engagement and contributing to more effective learning. On the other hand, our results prove the potential of fNIRS and fNIRS+EEG as modalities for BCI and user state monitoring, although further investigation is necessary to discriminate between various levels of CW.

However, we should underline that the sample considered in this study is too small and that these conclusions are only preliminary.

8.1 Limitations

- **Reduced sample size**

  The current pandemic situation we are experiencing has made the process of recruiting volunteers for this study difficult. Not only because safety and hygiene measures to prevent the SARS-CoV-2 virus discourage personal face-to-face contact, but also because some of these measures made it impossible to use the faculty's laboratories. Furthermore, some information from the acquisitions was lost since the experimental procedure was computationally demanding (collecting data from several sensors at a sampling frequency of 1000 Hz and taking photographs of the participants every 10 s). In total, we lost all data from one subject and the subtraction task part from another. As a result, the current findings should be considered preliminary only.

- **Sampling choice method**

  Our sampling was of convenience since we took advantage of the cases we had at our disposal due to the current pandemics situation. This method presents strong limitations since the results and conclusions only apply to the built sample and can not be generalized trusty to the population. On the other hand, it may be useful at the beginning of an investigation.

- **Same CW does not mean same brain activity**

  The assumption that the 3-back task, for example, generates greater CW in all participants than the 2-back task, or that, on the other hand, one of the cognitive tasks generates higher CW than the resting state has its limitations. Since, according to the neural efficiency hypothesis of intelligence, for the execution of a given task with a certain level of performance, some individuals will have to allocate a substantial amount of mental resources, while others will achieve the same results with much less mental effort. Thus, we may expect to see higher brain activity in a task considered more complicated compared to another simpler task, or even expect a specific fNIRS signal for the task state and another one for the baseline state, but this may not happen because a given individual may be able to perform the task without a significant amount of mental resources [126].

8.2 Future Work

Now that we have seen that it is possible to use the fNIRS technique for monitoring different levels of CW, and have created means to do so, future studies are needed to strengthen and optimise this monitoring.

- **Expanding the population**

  First and foremost, it would be of the utmost interest to add more participants to this study, employing the same experimental design. By increasing the sample of participants, the results would reach statistical significance, probably leading to more general conclusions. In particular, the influence of task difficulty on CW and the relationship between CW and the cognitive states, based on EEG bands powers, could be studied in more detail.

- **Adding sensors**

  Adding sensors to the regression approach for quantifying cognitive states would allow the number of features to be increased, providing a panoply of new options to the regressor that most likely improves its results. Sensors such as electrodermal activity (EDA) and electrocardiogram (ECG) could be very promising for cognitive states detection. Namely, for boredom, one would expect a lower level of arousal, i.e. a lower amplitude in the EDA signal and a lower frequency in the ECG signal (slower heartbeat) compared to the state of frustration. Sensors like this would also allow emotional monitoring, which we have seen as essential in the learning process [17, 51, 52]. In this way, the models built in this work for the discrimination of different cognitive levels could be optimised and combined with emotion evaluation sensors for the e-learning platforms design.

- **Finding the onset point of cognitive states**

  Analysing data's temporal evolution would allow us to determine the onset point of the cognitive state of a given subject. With this information, we could tell, for example, at what point in the lesson the student began to be interested, or on the other hand, which matter bored him. This analysis is promising not only for enhancing learning content and strategies but also to better understand the biological phenomena behind these events.

Bibliography

[1] M. S. Young, K. A. Brookhuis, C. D. Wickens, and P. A. Hancock. "State of science: mental workload in ergonomics." In: *Ergonomics* 58.1 (December 2014), pp. 1–17. DOI: 10.1080/00140139.2014.956151.

[2] M. Mazher, A. A. Aziz, A. S. Malik, and H. U. Amin. "An EEG-Based Cognitive Load Assessment in Multimedia Learning Using Feature Extraction and Partial Directed Coherence." In: *IEEE Access* 5 (2017), pp. 14819–14829. DOI: 10.1109/access.2017.2731784. URL: https://doi.org/10.1109/access.2017.2731784.

[3] A. C. Marinescu, S. Sharples, and A. C. Ritchie. "Physiological Parameter Response to Variation of Mental Workload." In: *Human Factors: The Journal of the Human Factors and Ergonomics* (September 2017). DOI: 10.1177/0018720817733101.

[4] V. Siritsa and A. Dolganov. "Overview of Functional Near-Infrared Spectroscopy Application Opportunities for Functional Diagnostics and Cognitive Assessment." In: *2020 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)*. Yekaterinburg, Russia, 2020, pp. 159–162. DOI: 10.1109/USBEREIT48449.2020.9117706.

[5] P. Zarjam, J. Epps, and N. H. Lovell. "Beyond Subjective Self-Rating: EEG Signal Classification of Cognitive Workload." In: *IEEE Transactions on Autonomous Mental Development* 7.4 (December 2015). DOI: 10.1109/TAMD.2015.2441960.

[6] U. Boehm, D. Matz, M. Gretton, K. C. S. Castro, M. Skinner, D. Strayer, and A. Heathcote. "Real-time prediction of short-timescale fuctuations in cognitive workload." In: *Cognitive Research: Principles and Implications* 6.30 (2021). DOI: 10.1186/s41235-021-00289-y.

[7] G. Young, L. Xavelina, and V. D. Hooper. "Assessment of Workload Using NASA Task Load Index in Perianesthesia Nursing." In: *Journal of PeriAnesthesia Nursing* 23.2 (May 2008). DOI: 10.1016/j.jopan.2008.01.008.

[8] L. T. Kohn, J. Corrigan, and M. S. Donaldson. *To Err is Human: Building a Safer Health System*. National Academies Press, 2000. ISBN: 0-309-06837-1.

[9] H. Maior, M. Pike, S. Sharples, and M. L. Wilson. "Examining the reliability of using fNIRS in realistic HCI settings for spatial and verbal tasks." In: *CHI'15 Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Volume 15.* 2015, pp. 3807–3816.

[10] M. B. Weinger and C. E. Englung. "Ergonomic and human factors affecting anesthetic vigilance and monitoring performance in the operating room environment." In: *Anesthesiology* 73 (November 1990), pp. 995–1021. DOI: 10.1097/00000542-199011000-00030.

[11] N. Kumar and J. Kumar. "Measurement of Cognitive Load in HCI Systems Using EEG Power Spectrum: An Experimental Study." In: *Procedia Computer Science* 84 (2016), pp. 70–78. DOI: 10.1016/j.procs.2016.04.068. URL: https://doi.org/10.1016/j.procs.2016.04.068.

[12] H. Maior, M. Pike, M. L.Wilson, and S. Sharples. "Continuous detection of workload overload." In: *Contemporary Ergonomics and Human Factors 2014.* Taylor &amp Francis, 2014, pp. 450–457. DOI: 10.1201/b16742-79. URL: https://doi.org/10.1201/b16742-79.

[13] S. Bench and H. Lench. "On the Function of Boredom." In: *Behavioral Sciences* 3.3 (Aug. 2013), pp. 459–472. DOI: 10.3390/bs3030459. URL: https://doi.org/10.3390/bs3030459.

[14] R. Ghali, S. Ouellet, C. Frasson, and and. "Using Electrophysiological Features in Cognitive Tasks: An Empirical Study." In: *International Journal of Information and Education Technology* 6.8 (2016), pp. 584–590. DOI: 10.7763/ijiet.2016.v6.756. URL: https://doi.org/10.7763/ijiet.2016.v6.756.

[15] C. J. Ho. "The effects of frustration on intellectual performance." In: *Science Education* 50.5 (Dec. 1966), pp. 457–460. DOI: 10.1002/sce.3730500509. URL: https://doi.org/10.1002/sce.3730500509.

[16] B. F. Jeronimus and O. M. Laceulle. "Frustration." In: *Encyclopedia of Personality and Individual Differences.* Springer International Publishing, 2017, pp. 1–5. DOI: 10.1007/978-3-319-28099-8_815-1. URL: https://doi.org/10.1007/978-3-319-28099-8_815-1.

[17] B. Bontchev. "Adaptation in Affective Video Games: A Literature Review." In: *Cybernetics and Information Technologies* 16.3 (August 2016). DOI: 10.1515/cait-2016-0032.

[18] C. Herff, D. Heger, O. Fortmann, J. Hennrich, F. Putze, and T. Schultz. "Mental workload during n-back task-quantified in the prefrontal cortex using fNIRS." In: *Frontiers in Human Neuroscience* 7 (January 2014). DOI: 10.3389/fnhum.2013.00935.

[19]   N. Naseer and K.-S. Hong. "fNIRS-based brain-computer interfaces: a review." In: *Frontiers in Human Neuroscience* 9 (Jan. 2015). DOI: 10.3389/fnhum.2015.00003. URL: https://doi.org/10.3389/fnhum.2015.00003.

[20]   Y. Li, K. K. Ang, and C. Guan. "Digital Signal Processing and Machine Learning." In: *Brain-Computer Interfaces*. Springer Berlin Heidelberg, 2009, pp. 305–330. DOI: 10.1007/978-3-642-02091-9_17. URL: https://doi.org/10.1007/978-3-642-02091-9_17.

[21]   G. Zaharija, P. Bogunović, and S. Mladenović. "Brain computer interface in enhanced learning system." In: *INTED2018 Proceedings*. IATED, Mar. 2018. DOI: 10.21125/inted.2018.1029. URL: https://doi.org/10.21125/inted.2018.1029.

[22]   G. Papanastasiou, A. Drigas, C. Skianis, and M. Lytras. "Brain computer interface based applications for training and rehabilitation of students with neurodevelopmental disorders. A literature review." In: *Heliyon* 6.9 (Sept. 2020), e04250. DOI: 10.1016/j.heliyon.2020.e04250. URL: https://doi.org/10.1016/j.heliyon.2020.e04250.

[23]   H. Aghajani, M. Garbey, and A. Omurtag. "Measuring Mental Workload with EEGfNIRS." In: *Frontiers in Human Neuroscience* 11 (July 2017). DOI: 10.3389/fnhum.2017.00359. URL: https://doi.org/10.3389/fnhum.2017.00359.

[24]   D. Afergan, E. M. Peck, E. T. Solovey, A. Jenkins, S. W. Hincks, and E. T. Brown. "Dynamic difficulty using brain metrics of workload." In: *CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. April 2014, pp. 3797–3806. DOI: 10.1145/2556288.2557230.

[25]   M. F. Pike, H. A. Maior, M. Porcheron, S. C. Sharples, and M. L. Wilson. "Measuring the effect of think aloud protocols on workload using fNIRS." In: *CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing*. April 2014, pp. 3807–3816. DOI: 10.1145/2556288.2556974.

[26]   E. T. Solovey, A.Girouard, K. Chauncey, L. M. Hirshfield, and A. Sassaroli. "Using fNIRS brain sensing in realistic HCI settings: experiments and guidelines." In: *UIST '09: Proceedings of the 22nd annual ACM symposium on User interface software and technology*. October 2009, pp. 157–166. DOI: 10.1145/1622176.1622207.

[27]   B. F. Yuksel, K. B. Oleson, L. Harrison, E. M. Peck, D. Afergan, R. Chang, and R. J. K. Jacob. "Learn Piano with BACh: An Adaptive Learning Interface that Adjusts Task Difficulty Based on Brain State." In: *CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. May 2016, 5372–5384. DOI: 10.1145/2858036.2858388.

[28]   V. Quaresima and M. Ferrari. "Functional Near-Infrared Spectroscopy (fNIRS) for Assessing Cerebral Cortex Function During Human Behavior in Natural/Social Situations: A Concise Review." In: *Organizational Research Methods* (July 2016), pp. 44–68. DOI: 10.1177/1094428116658959.

[29]   A. Baddeley. "Working Memory: Theories, Models, and Controversies." In: *Annual Review of Psychology* 63 (January 2012), pp. 1–29. DOI: 10.1146/annurev-psych-120710-100422.

[30]   H. A. Foy, P. Runham, and P. Chapman. "Prefrontal Cortex Activation and Young Driver Behaviour: A fNIRS Study." In: *PLOS ONE* (May 2016). DOI: 10.1371/journal.pone.0156512.

[31]   J. D. E. Gabrieli, R. A. Poldrack, and J. E. Desmond. "The role of left prefrontal cortex in language and memory." In: *Proceedings of the National Academy of Sciences of the United States of America. Volume 95, Issue 3*. February 1998, 906–913.

[32]   E. K. Miller and J. D. Cohen. "An Integrative Theory of Prefrontal Cortex Function." In: *Annual Review of Neuroscience* 24 (March 2001), pp. 167–202. DOI: 10.1146/annurev.neuro.24.1.167.

[33]   A. C. Guyton and J. E. Hall. *Textbook of Medical Physiology*. Eleventh. Elsevier Saunders, 2006. ISBN: 0-7216-0240-1.

[34]   L. S. Costanzo. *Physiology*. Sixth. Elsevier, 2018. ISBN: 978-0-323-47881-6.

[35]   M. Causse, Z. Chua, V. Peysakhovich, N. D. Campo, and N. Matton. "Mental workload and neural efficiency quantified in the prefrontal cortex using fNIRS." In: *Scientific Reports* 7.1 (July 2017). DOI: 10.1038/s41598-017-05378-x. URL: https://doi.org/10.1038/s41598-017-05378-x.

[36]   C. Herff, D. Heger, F. Putze, J. Hennrich, O. Fortmann, and T. Schultz. "Classification of mental tasks in the prefrontal cortex using fNIRS." In: *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, July 2013. DOI: 10.1109/embc.2013.6609962. URL: https://doi.org/10.1109/embc.2013.6609962.

[37]   W. Klimesch. "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis." In: *Brain Research Reviews* 29 (1999), pp. 169–195. DOI: 10.1016/s0165-0173(98)00056-3.

[38]   S. Sharma. 2020. URL: https://www.quora.com/What-are-brain-waves-Which-branch-of-science-deals-with-it-I-want-to-go-deep-into-this-topic.

[39]   D.Y.Gao and H.D.Sherali. *Advances in Applied Mathematics and Global Optimization*. Seventeen. Springer, 2009. ISBN: 978-0-387-75714-8. DOI: 10.1007/978-0-387-75714-8.

[40]   N.H.Liu, C. Chiang, and H. Chu. "Recognizing the Degree of Human Attention Using EEG Signals from Mobile Sensors." In: *sensors* 13 (2013), pp. 10273–10286. DOI: 10.3390/s130810273.

[41] C.-M. Chen, J.-Y. Wang, and C.-M. Yu. "Assessing the attention levels of students by using a novel attention aware system based on brainwave signals." In: *British Journal of Educational Technology* 48.2 (2015), pp. 348–369. DOI: 10.1111/bjet.12359. URL: https://doi.org/10.1111/bjet.12359.

[42] Y. Buskila, A. Bellot-Saez, and J. W. Morley. "Generating Brain Waves, the Power of Astrocytes." In: *Frontiers in Neuroscience* 13 (Oct. 2019). DOI: 10.3389/fnins.2019.01125. URL: https://doi.org/10.3389/fnins.2019.01125.

[43] E.P.Shaw, J.C.Rietschel, B.D.Hendershot, A.L.Pruziner, M.W.Miller, B.D.Hatfield, and R.J.Gentili. "Measurement of attentional reserve and mental effort for cognitive workload assessment under various task demands during dual-task walking." In: *Biological Psychology* 134 (2018), pp. 39–51. DOI: 10.1016/j.biopsycho.2018.01.009.

[44] R. F. Rojas, E. Debie, J. Fidock, M. Barlow, K. Kasmarik, S. Anavatti, M. Garratt, and H. Abbass. "Electroencephalographic Workload Indicators During Teleoperation of an Unmanned Aerial Vehicle Shepherding a Swarm of Unmanned Ground Vehicles in Contested Environments." In: *Frontiers in Neuroscience* 14 (Feb. 2020). DOI: 10.3389/fnins.2020.00040. URL: https://doi.org/10.3389/fnins.2020.00040.

[45] S. Puma, N. Matton, P. Paubel, E. Raufaste, and R. El-Yahoubi. "Using theta and alpha band power to assess cognitive workload in multitasking environments." In: *International Journal of Psychophysiology* 123 (2018), pp. 111–120. DOI: 10.1016/j.ijpsycho.2017.10.004.

[46] Y. Moriguchi and K. Hiraki. "Prefrontal cortex and executive function in young children:a review of NIRS studies." In: *Frontiers in Human Neuroscience* 7 (2013), p. 867. DOI: 10.3389/fnhum.2013.00867.

[47] The Office of Learning and Teaching, 2004. Melbourne: Department of Education and Training; OECD, 2010. Nature of Learning, Paris. URL: http://www.ibe.unesco.org/en/geqaf/annexes/technical-notes/most-influential-theories-learning.

[48] L. Embree and W. James. "The Principles of Psychology." In: *Philosophy and Phenomenological Research* 44.1 (1983), p. 124. DOI: 10.2307/2107586. URL: https://doi.org/10.2307/2107586.

[49] J. A. DEUTSCH. "Perception and Communication." In: *Nature* 182.4649 (Dec. 1958), pp. 1572–1572. DOI: 10.1038/1821572a0. URL: https://doi.org/10.1038/1821572a0.

[50] T. P. Zanto and A. Gazzaley. "Chapter 20 - Aging of the frontal lobe." In: *Handbook of Clinical Neurology* 163 (2019), pp. 369–389. DOI: 10.1016/B978-0-12-804281-6.00020-3.

[51] F. D'Errico, M. Paciello, and L. Cerniglia. "When emotions enhance students' engagement in e-learning processes." en. In: *Journal of e-Learning and Knowledge Society* Vol 12 (2016), No 4 (2016): Journal of e–Learning and Knowledge Society. DOI: 10.20368/1971-8829/1144. URL: http://www.je-lks.org/ojs/index.php/Je-LKS_EN/article/view/1144.

[52] R. E. Mayer. "Searching for the role of emotions in e-learning." In: *Learning and Instruction* 70 (Dec. 2020), p. 101213. DOI: 10.1016/j.learninstruc.2019.05.010. URL: https://doi.org/10.1016/j.learninstruc.2019.05.010.

[53] M. C. Duffy, S. P. Lajoie, R. Pekrun, and K. Lachapelle. "Emotions in medical education: Examining the validity of the Medical Emotion Scale (MES) across authentic medical learning environments." In: *Learning and Instruction* 70 (Dec. 2020), p. 101150. DOI: 10.1016/j.learninstruc.2018.07.001. URL: https://doi.org/10.1016/j.learninstruc.2018.07.001.

[54] H. Järvenoja and S. Järvelä. "How students describe the sources of their emotional and motivational experiences during the learning process: A qualitative approach." In: *Learning and Instruction* 15.5 (Oct. 2005), pp. 465–480. DOI: 10.1016/j.learninstruc.2005.07.012. URL: https://doi.org/10.1016/j.learninstruc.2005.07.012.

[55] J. M. Harley, S. P. Lajoie, C. Frasson, and N. C. Hall. "Developing Emotion-Aware, Advanced Learning Technologies: A Taxonomy of Approaches and Features." In: *International Journal of Artificial Intelligence in Education* 27.2 (Dec. 2016), pp. 268–297. DOI: 10.1007/s40593-016-0126-8. URL: https://doi.org/10.1007/s40593-016-0126-8.

[56] A Villringer. "Non-invasive optical spectroscopy and imaging of human brain function." In: *Trends in Neurosciences* 20.10 (Oct. 1997), pp. 435–442. DOI: 10.1016/s0166-2236(97)01132-6. URL: https://doi.org/10.1016/s0166-2236(97)01132-6.

[57] F. Herold, P. Wiegel, F. Scholkmann, and N. Müller. "Applications of Functional Near-Infrared Spectroscopy (fNIRS) Neuroimaging in Exercise–Cognition Science: A Systematic, Methodology-Focused Review." In: *Journal of Clinical Medicine* 7.12 (Nov. 2018), p. 466. DOI: 10.3390/jcm7120466. URL: https://doi.org/10.3390/jcm7120466.

[58] P. Pinti, I. Tachtsidis, A. Hamilton, J. Hirsch, C. Aichelburg, S. Gilbert, and P. W. Burgess. "The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience." In: *Annals of the New York Academy of Sciences* 1464.1 (Mar. 2020), pp. 5–29. DOI: 10.1111/nyas.13948. URL: https://doi.org/10.1111/nyas.13948.

[59] F. Scholkmann, S. Kleiser, A. J. Metz, R. Zimmermann, J. M. Pavia, U. Wolf, and M. Wolf. "A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology." In: *NeuroImage* 85 (Jan. 2014), pp. 6–27. DOI: 10.1016/j.neuroimage.2013.05.004. URL: https://doi.org/10.1016/j.neuroimage.2013.05.004.

[60] B. Bracken, E. Festa, H.-M. Sun, C. Leather, and G. Strangman. "Validation of the fNIRS Pioneer™, a Portable, Durable, Rugged functional Near-Infrared Spectroscopy (fNIRS) Device." In: *Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies*. SCITEPRESS - Science and Technology Publications, 2019. DOI: 10.5220/0007471405210531. URL: https://doi.org/10.5220/0007471405210531.

[61] D. Embrey, C. Blackett, P. Marsden, and J. Peachey. *Development of a Human Cognitive Workload Assessment Tool*. Tech. rep. 2006.

[62] I. Hoang, M. Ranchet, R. Derollepot, F. Moreau, and L. Paire-Ficout. "Measuring the Cognitive Workload During Dual-Task Walking in Young Adults: A Combination of Neurophysiological and Subjective Measures." In: *Frontiers in Human Neuroscience* 14 (Nov. 2020). DOI: 10.3389/fnhum.2020.592532. URL: https://doi.org/10.3389/fnhum.2020.592532.

[63] Özgür Örün and Y. Akbulut. "Effect of multitasking, physical environment and electroencephalography use on cognitive load and retention." In: *Computers in Human Behavior* 92 (Mar. 2019), pp. 216–229. DOI: 10.1016/j.chb.2018.11.027. URL: https://doi.org/10.1016/j.chb.2018.11.027.

[64] N. Cowan. "Chapter 20 What are the differences between long-term, short-term, and working memory?" In: *Progress in Brain Research*. Elsevier, 2008, pp. 323–338. DOI: 10.1016/s0079-6123(07)00020-9. URL: https://doi.org/10.1016/s0079-6123(07)00020-9.

[65] S. Midha, H. A. Maior, M. L. Wilson, and S. Sharples. "Measuring Mental Workload Variations in Office Work Tasks using fNIRS." In: *International Journal of Human-Computer Studies* 147 (Mar. 2021), p. 102580. DOI: 10.1016/j.ijhcs.2020.102580. URL: https://doi.org/10.1016/j.ijhcs.2020.102580.

[66] F. Dehais, A. Lafont, R. Roy, and S. Fairclough. "A Neuroergonomics Approach to Mental Workload, Engagement and Human Performance." In: *Frontiers in Neuroscience* 14 (Apr. 2020). DOI: 10.3389/fnins.2020.00268. URL: https://doi.org/10.3389/fnins.2020.00268.

[67] M. Akay, ed. *Wiley Encyclopedia of Biomedical Engineering*. John Wiley & Sons, Inc., Apr. 2006. DOI: 10.1002/9780471740360. URL: https://doi.org/10.1002/9780471740360.

[68] G. D. Flumeri, G. Borghini, P. Aricò, N. Sciaraffa, P. Lanzi, S. Pozzi, V. Vignali, C. Lantieri, A. Bichicchi, A. Simone, and F. Babiloni. "EEG-Based Mental Workload Neurometric to Evaluate the Impact of Different Traffic and Road Conditions in Real Driving Settings." In: *Frontiers in Human Neuroscience* 12 (Dec. 2018). DOI: 10.3389/fnhum.2018.00509. URL: https://doi.org/10.3389/fnhum.2018.00509.

[69] R. W. Thatcher, D. North, and C. Bive. "EEG and intelligence: Relations between EEG coherence, EEG phase delay and power." In: *Clin. Neurophysiol.* 116.9 (2005), pp. 2129–2141. DOI: 10.1016/j.clinph.2005.04.026.

[70] S. Noshadi, V. Abootalebi, M. T. Sadeghi, and M. S. Shahvazian. "Selection of an efficient feature space for EEG-based mental task discrimination." In: *Biocybern. Biomed. Eng.* 34.3 (2014), pp. 159–168. DOI: 10.1016/j.bbe.2014.03.004.

[71] H. Qu, Y. Shan, Y. Liu, L. Pang, Z. Fan, J. Zhang, and X. Wanyan. "Mental Workload Classification Method Based on EEG Independent Component Features." In: *Applied Sciences* 10.9 (Apr. 2020), p. 3036. DOI: 10.3390/app10093036. URL: https://doi.org/10.3390/app10093036.

[72] M. Plechawska-Wójcik, M. Tokovarov, M. Kaczorowska, and D. Zapała. "A Three-Class Classification of Cognitive Workload Based on EEG Spectral Data." In: *Applied Sciences* 9.24 (Dec. 2019), p. 5340. DOI: 10.3390/app9245340. URL: https://doi.org/10.3390/app9245340.

[73] L. Zhiwei and S. Minfen. "Classification of mental task EEG signals using wavelet packet entropy and SVM." In: *in Proc. 8th Int. Conf. Electron. Meas.Instrum.* 2007, pp. 3906–3909. DOI: 10.1109/ICEMI.2007.4351064.

[74] C. Lin and M. Hsieh. "Classification of mental task from EEG data using neural networks based on particle swarm optimization." In: *Neurocomputing* 72.4 (2009), pp. 1121–1130. DOI: 10.1016/j.neucom.2008.02.017.

[75] G. Rodríguez-Bermúdez, P. J. García-Laencina, J. Roca-González, and J. Roca-Dorda. "Efficient feature selection and linear discrimination of EEG signals." In: *Neurocomputing* 115 (2013), pp. 161–165. DOI: 10.1016/j.neucom.2013.01.001.

[76] S. Karkare, G. Saha, and J. Bhattacharya. "Investigating long-range correlation properties in EEG during complex cognitive tasks." In: *Chaos, Solitons Fractals* 42.4 (2009), pp. 2067–2073. DOI: 10.1016/j.chaos.2009.03.148.

[77] H. U. A. et al. "Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques." In: *Chaos, Solitons Fractals* 38.1 (2015), pp. 139–149. DOI: 10.1007/s13246-015-0333-x.

[78] C. Beal and C. Galan. *EEG estimates of engagement and cognitive workload predict math problem solving outcomes.* presented at User Modeling And Personalization Conference, Montreal, Canada. 2012.

[79] R. H. Stevens, T. Galloway, and C. Berka. "EEG-Related Changes in Cognitive Workload, Engagement and Distraction as Students Acquire Problem Solving Skills." In: *User Modeling 2007*. Springer Berlin Heidelberg, pp. 187–196. DOI: 10.1007/978-3-540-73078-1_22. URL: https://doi.org/10.1007/978-3-540-73078-1_22.

[80] R. A. Sottilare. "Using Learner Data to Influence Performance during Adaptive Tutoring Experiences." en. In: (2014). DOI: 10.13140/2.1.3380.4801. URL: http://rgdoi.net/10.13140/2.1.3380.4801.

[81] M. S. Caywood, D. M. Roberts, J. B. Colombe, H. S. Greenwald, and M. Z. Weiland. "Gaussian Process Regression for Predictive But Interpretable Machine Learning Models: An Example of Predicting Mental Workload across Tasks." In: *Frontiers in Human Neuroscience* 10 (Jan. 2017). DOI: 10.3389/fnhum.2016.00647. URL: https://doi.org/10.3389/fnhum.2016.00647.

[82] T. M. Mitchell, R. Hutchinson, R. S. Niculescu, F. Pereira, X. Wang, M. Just, and S. Newman. "Learning to Decode Cognitive States from Brain Images." In: *Machine Learning* 57.1/2 (Oct. 2004), pp. 145–175. DOI: 10.1023/b:mach.0000035475.85309.1b. URL: https://doi.org/10.1023/b:mach.0000035475.85309.1b.

[83] T. Csipo, A. Lipecz, P. Mukli, D. Bahadli, O. Abdulhussein, C. D. Owens, S. Tarantini, R. A. Hand, V. Yabluchanska, J. M. Kellawan, F. Sorond, J. A. James, A. Csiszar, Z. I. Ungvari, and A. Yabluchanskiy. "Increased cognitive workload evokes greater neurovascular coupling responses in healthy young adults." In: *PLOS ONE* 16.5 (May 2021). Ed. by S. Ogoh, e0250043. DOI: 10.1371/journal.pone.0250043. URL: https://doi.org/10.1371/journal.pone.0250043.

[84] S. M. Jaeggi, R. Seewer, A. C. Nirkko, D. Eckstein, G. Schroth, R. Groner, and K. Gutbrod. "Does excessive memory load attenuate activation in the prefrontal cortex? Load-dependent processing in single and dual tasks: functional magnetic resonance imaging study." In: *NeuroImage* 19.2 (June 2003), pp. 210–225. DOI: 10.1016/s1053-8119(03)00098-3. URL: https://doi.org/10.1016/s1053-8119(03)00098-3.

[85] S. Howard, H. Burianová, J. Ehrich, L. Kervin, A. Calleia, E. Barkus, J. Carmody, and S. Humphry. "Behavioral and fMRI evidence of the differing cognitive load of domain-specific assessments." In: *Neuroscience* 297 (June 2015), pp. 38–46. DOI: 10.1016/j.neuroscience.2015.03.047. URL: https://doi.org/10.1016/j.neuroscience.2015.03.047.

[86] R. McKendrick and A. Harwood. "Cognitive Workload and Workload Transitions Elicit Curvilinear Hemodynamics During Spatial Working Memory." In: *Frontiers in Human Neuroscience* 13 (Nov. 2019). DOI: 10.3389/fnhum.2019.00405. URL: https://doi.org/10.3389/fnhum.2019.00405.

89

[87]   S. Rubio, E. Diaz, J. Martin, and J. M. Puente. "Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods." In: *Applied Psychology* 53.1 (Jan. 2004), pp. 61–86. DOI: 10.1111/j.1464-0597.2004.00161.x. URL: https://doi.org/10.1111/j.1464-0597.2004.00161.x.

[88]   K. Knaepen, U. Marusic, S. Crea, C. D. R. Guerrero, N. Vitiello, N. Pattyn, O. Mairesse, D. Lefeber, and R. Meeusen. "Psychophysiological response to cognitive workload during symmetrical, asymmetrical and dual-task walking." In: *Human Movement Science* 40 (Apr. 2015), pp. 248–263. DOI: 10.1016/j.humov.2015.01.001. URL: https://doi.org/10.1016/j.humov.2015.01.001.

[89]   J. Vogels, V. Demberg, and J. Kray. "The Index of Cognitive Activity as a Measure of Cognitive Processing Load in Dual Task Settings." In: *Frontiers in Psychology* 9 (Nov. 2018). DOI: 10.3389/fpsyg.2018.02276. URL: https://doi.org/10.3389/fpsyg.2018.02276.

[90]   S. Marshall. "The Index of Cognitive Activity: measuring cognitive workload." In: *Proceedings of the IEEE 7th Conference on Human Factors and Power Plants.* IEEE. DOI: 10.1109/hfpp.2002.1042860. URL: https://doi.org/10.1109/hfpp.2002.1042860.

[91]   V. Demberg, A. Sayeed, A. Mahr, and C. Müller. "Measuring linguistically-induced cognitive load during driving using the ConTRe task." In: *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13.* ACM Press, 2013. DOI: 10.1145/2516540.2516546. URL: https://doi.org/10.1145/2516540.2516546.

[92]   S. D. Power, A. Kushki, and T. Chau. "Intersession Consistency of Single-Trial Classification of the Prefrontal Response to Mental Arithmetic and the No-Control State by NIRS." In: *PLoS ONE* 7.7 (July 2012). Ed. by Y. Fan, e37791. DOI: 10.1371/journal.pone.0037791. URL: https://doi.org/10.1371/journal.pone.0037791.

[93]   J. Shin, A. von Lühmann, D.-W. Kim, J. Mehnert, H.-J. Hwang, and K.-R. Müller. "Simultaneous acquisition of EEG and NIRS during cognitive tasks for an open access dataset." In: *Scientific Data* 5.1 (Feb. 2018). DOI: 10.1038/sdata.2018.3. URL: https://doi.org/10.1038/sdata.2018.3.

[94]   C. Cepeda, R. Tonet, D. N. Osorio, H. P. Silva, E. Battegay, M. Cheetham, and H. Gamboa. "Latent: A Flexible Data Collection Tool to Research Human Behavior in the Context of Web Navigation." In: *IEEE Access* 7 (2019), pp. 77659–77673. DOI: 10.1109/ACCESS.2019.2916996.

[95]   J. N. Acharya, A. Hani, J. Cheek, P. Thirumala, and T. N. Tsuchida. "American Clinical Neurophysiology Society Guideline 2." In: *Journal of Clinical Neurophysiology* 33.4 (Aug. 2016), pp. 308–311. DOI: 10.1097/wnp.0000000000000316. URL: https://doi.org/10.1097/wnp.0000000000000316.

[96] J. Peirce, J. R. Gray, S. Simpson, M. MacAskill, R. Höchenberger, H. Sogo, E. Kastman, and J. K. Lindeløv. "PsychoPy2: Experiments in behavior made easy." In: *Behavior Research Methods* 51.1 (Feb. 2019), pp. 195–203. DOI: 10.3758/s13428-018-01193-y. URL: https://doi.org/10.3758/s13428-018-01193-y.

[97] C. S. Bediz, A. Oniz, C. Guducu, E. U. Demirci, H. Ogut, E. Gunay, C. Cetinkaya, and M. Ozgoren. "Acute Supramaximal Exercise Increases the Brain Oxygenation in Relation to Cognitive Workload." In: *Frontiers in Human Neuroscience* 10 (Apr. 2016). DOI: 10.3389/fnhum.2016.00174. URL: https://doi.org/10.3389/fnhum.2016.00174.

[98] A Gevins. "High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice." In: *Cerebral Cortex* 7.4 (June 1997), pp. 374–385. DOI: 10.1093/cercor/7.4.374. URL: https://doi.org/10.1093/cercor/7.4.374.

[99] M. E. Smith and A. Gevins. "Neurophysiologic monitoring of mental workload and fatigue during operation of a flight simulator." In: *Biomonitoring for Physiological and Cognitive Performance during Military Operations*. Ed. by J. A. Caldwell and N. J. Wesensten. SPIE, May 2005. DOI: 10.1117/12.602181. URL: https://doi.org/10.1117/12.602181.

[100] H. Ayaz, P. A. Shewokis, S. Bunce, K. Izzetoglu, B. Willems, and B. Onaral. "Optical brain monitoring for operator training and mental workload assessment." In: *NeuroImage* 59.1 (Jan. 2012), pp. 36–47. DOI: 10.1016/j.neuroimage.2011.06.023. URL: https://doi.org/10.1016/j.neuroimage.2011.06.023.

[101] H. Sato, N. Yahata, T. Funane, R. Takizawa, T. Katura, H. Atsumori, Y. Nishimura, A. Kinoshita, M. Kiguchi, H. Koizumi, M. Fukuda, and K. Kasai. "A NIRS–fMRI investigation of prefrontal cortex activity during a working memory task." In: *NeuroImage* 83 (Dec. 2013), pp. 158–173. DOI: 10.1016/j.neuroimage.2013.06.043. URL: https://doi.org/10.1016/j.neuroimage.2013.06.043.

[102] F. A. Fishburn, M. E. Norr, A. V. Medvedev, and C. J. Vaidya. "Sensitivity of fNIRS to cognitive state and load." In: *Frontiers in Human Neuroscience* 8 (2014). DOI: 10.3389/fnhum.2014.00076. URL: https://doi.org/10.3389/fnhum.2014.00076.

[103] K. Mandrick, V. Peysakhovich, F. Rémy, E. Lepron, and M. Causse. "Neural and psychophysiological correlates of human performance under stress and high mental workload." In: *Biological Psychology* 121 (Dec. 2016), pp. 62–73. DOI: 10.1016/j.biopsycho.2016.10.002. URL: https://doi.org/10.1016/j.biopsycho.2016.10.002.

[104] URL: file:///C:/Users/inesi/Downloads/18DigitalSignalProcessingandMachineLearning.pdf.

[105] URL: https://github.com/KaunilD/mes2hb#license.

[106] W. J. Conover. *Practical nonparametric statistics*. John Wiley Sons, Inc., 1999, p. 350.

[107] L. E. Ismail and W. Karwowski. "Applications of EEG indices for the quantification of human cognitive performance: A systematic review and bibliometric analysis." In: *PLOS ONE* 15.12 (Dec. 2020). Ed. by W. Mumtaz, e0242857. DOI: 10.1371/journal.pone.0242857. URL: https://doi.org/10.1371/journal.pone.0242857.

[108] J. Kim and E. Andre. "Emotion recognition based on physiological changes in music listening." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30.12 (Dec. 2008), pp. 2067–2083. DOI: 10.1109/tpami.2008.26. URL: https://doi.org/10.1109/tpami.2008.26.

[109] Z. A. Bakar, P. Long, and R. Muhlberger. "The effects of dynamic aesthetic interfaces on students' emotion and performance." In: *2011 International Conference on User Science and Engineering (i-USEr)*. IEEE, Nov. 2011. DOI: 10.1109/iuser.2011.6150527. URL: https://doi.org/10.1109/iuser.2011.6150527.

[110] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. "Scikit-learn: Machine Learning in Python." In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[111] URL: https://datascience.stackexchange.com/questions/74465/how-to-understand-anova-f-for-feature-selection-in-python-sklearn-selectkbest-w.

[112] W. Jean. *Mistakes in Applying Univariate Feature Selection Methods*. 2020. URL: https://towardsdatascience.com/mistakes-in-applying-univariate-feature-selection-methods-34c43ce8b93d.

[113] M. Barandas, D. Folgado, L. Fernandes, S. Santos, M. Abreu, P. Bota, H. Liu, T. Schultz, and H. Gamboa. "TSFEL: Time Series Feature Extraction Library." In: *SoftwareX* 11 (Jan. 2020), p. 100456. DOI: 10.1016/j.softx.2020.100456. URL: https://doi.org/10.1016/j.softx.2020.100456.

[114] URL: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.linregress.html.

[115] URL: https://www.itl.nist.gov/div898/handbook/eda/section3/eda35b.htm.

[116] J. Alzubi, A. Nayyar, and A. Kumar. "Machine Learning from Theory to Algorithms: An Overview." In: *Journal of Physics: Conference Series* 1142 (Nov. 2018), p. 012012. DOI: 10.1088/1742-6596/1142/1/012012. URL: https://doi.org/10.1088/1742-6596/1142/1/012012.

[117] M. I. Jordan and T. M. Mitchell. "Machine learning: Trends, perspectives, and prospects." In: *Science* 349.6245 (July 2015), pp. 255–260. DOI: 10.1126/science.aaa8415. URL: https://doi.org/10.1126/science.aaa8415.

[118] K. P. Murphy. *Machine Learning: A Probabilistic Perspective*. 1991, pp. 1–25. ISBN: 9780262018029. URL: http://link.springer.com/chapter/10.1007/978-94-011-3532-0_2.

[119] S. Misra and H. Li. "Noninvasive fracture characterization based on the classification of sonic wave travel times." In: *Machine Learning for Subsurface Characterization*. Elsevier, 2020, pp. 243–287. DOI: 10.1016/b978-0-12-817736-5.00009-0. URL: https://doi.org/10.1016/b978-0-12-817736-5.00009-0.

[120] H. W. Dong, C. Mills2, R. T. Knight, and J. W. Y. Ka. "Detection of mind wandering using EEG: Within and across individuals." In: *Plos One* 16 (Nov. 2018), p. 7. DOI: 10.17605/OSF.IO/DFWBQ.. URL: https://doi.org/10.1088/1742-6596/1142/1/012012.

[121] H. M and S. M.N. "A Review on Evaluation Metrics for Data Classification Evaluations." In: *International Journal of Data Mining & Knowledge Management Process* 5.2 (Mar. 2015), pp. 01–11. DOI: 10.5121/ijdkp.2015.5201. URL: https://doi.org/10.5121/ijdkp.2015.5201.

[122] A. Chakure. *Random Forest Regression*. 2019. URL: https://medium.com/swlh/random-forest-and-its-implementation-71824ced454f.

[123] *Difference between Random Forest and Extremely Randomized Trees*. 2016. URL: https://stats.stackexchange.com/questions/175523/difference-between-random-forest-and-extremely-randomized-trees.

[124] S. Wu. *3 Best metrics to evaluate Regression Model?* 2020. URL: https://towardsdatascience.com/what-are-the-best-metrics-to-evaluate-your-regression-model-418ca481755b.

[125] G. Dolce and H. Waldeier. "Spectral and multivariate analysis of EEG changes during mental activity in man." In: *Electroencephalography and Clinical Neurophysiology* 36 (Jan. 1974), pp. 577–584. DOI: 10.1016/0013-4694(74)90224-7. URL: https://doi.org/10.1016/0013-4694(74)90224-7.

[126] B. Dunst, M. Benedek, E. Jauk, S. Bergner, K. Koschutnig, M. Sommer, A. Ischebeck, B. Spinath, M. Arendasy, M. Bühner, H. Freudenthaler, and A. C. Neubauer. "Neural efficiency as a function of task demands." In: *Intelligence* 42 (Jan. 2014), pp. 22–30. DOI: 10.1016/j.intell.2013.09.005. URL: https://doi.org/10.1016/j.intell.2013.09.005.

A

Supplementary Material

This appendix contains the mean waves from fNIRS and the EEG band powers that are related to Chapter 6.

A.1 Mean Waves from Functional Near Infrared Spectroscopy

Figure A.1 presents the mean oxyHb and deoxyHb concentrations (from fNIRS) from the position F8 of the 10-20 system from the 8 subjects in the subtraction task and in the rest period. Figures A.2 and A.3 show the mean oxyHb and deoxyHb concentrations (from fNIRS) from the position F8 of the 10-20 system from the 8 subjects for the four n-back conditions.

A.2 Electroencephalography Band Powers

Tables A.1 and A.2 show the results of the comparison between EEG band powers for the subtraction task and for the resting period and Table A.3 presents the respective p-values, for statistical relevance assessment.

Table A.1: **EEG bands power for the subtraction and the resting period (F7).** The EEG signals were acquired in the position F7 of the 10-20 system (left side). The unit of power is μV^2.

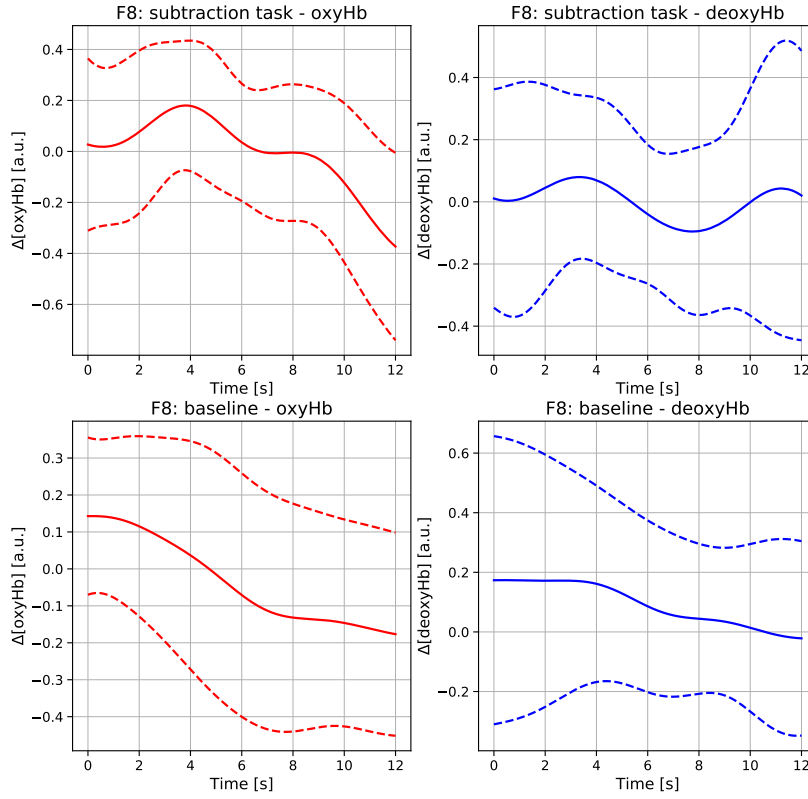| Power	Subtraction	Rest
Alpha	2.92 ± 1.62	3.92 ± 2.48
Beta	7.27 ± 5.76	4.82 ± 4.38
Delta	45.83 ± 50.67	78.58 ± 68.68
Gamma	0.59 ± 0.78	0.61 ± 0.98
Theta	9.03 ± 8.16	8.90 ± 7.65

Figure A.1: **Mean oxyHb and deoxyHb concentrations of the 8 subjects in the subtraction task and in the rest period (F8).** On the left side of the figure, the oxyHb signals and on the right side the deoxyHb signals are represented. In the same figure, the signals corresponding to the subtraction task, at the top, and, at the bottom, the signals corresponding to the subjects' rest period are represented. These signals were measured at position F8 of the 10-20 system. The dashed lines represent the standard deviation.

Tables A.4 and A.5 show the results of the EEG bands powers for the multiple levels of the n-back task and Tables A.6 and A.7 show the respective p-values, for statistical relevance assessment.
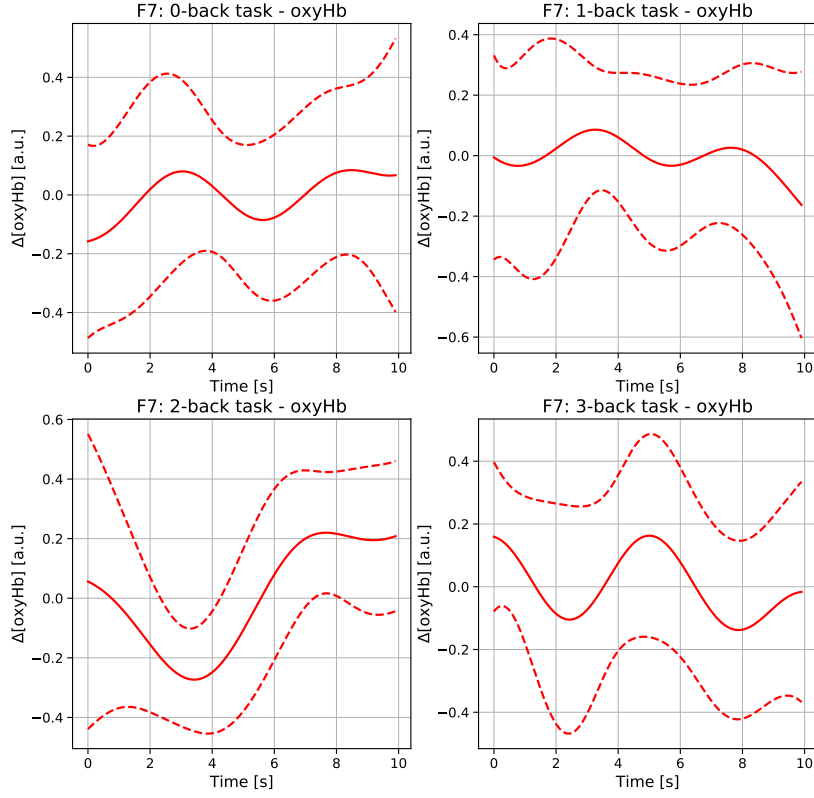
Figure A.2: **Mean oxyHb concentrations for 0/1/2 and 3-back tasks (F7).** The signals are the mean waves of oxyHb signals (fNIRS) for 0/1/2 and 3-back tasks at position F7 of the 10-20 system. The dashed lines represent the standard deviation.

Table A.2: **EEG bands power for the subtraction and the resting period (F8).** The EEG signals were acquired in the position F8 of the 10-20 system (right side). The unit of power is μV^2.

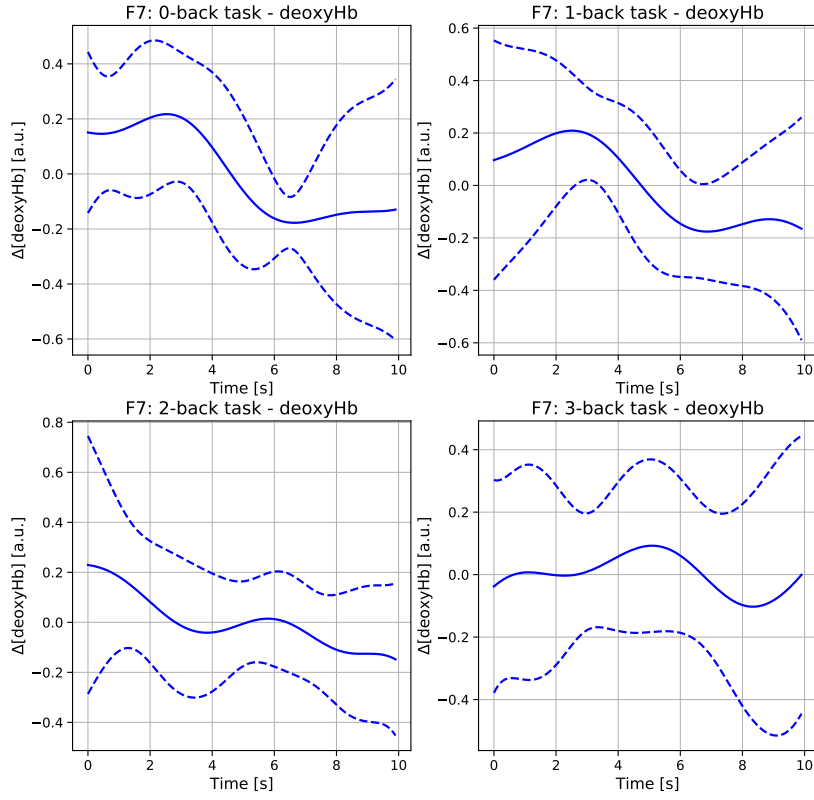| Power	Subtraction	Rest
Alpha	3.78 ± 3.12	3.21 ± 1.21
Beta	7.85 ± 8.47	4.28 ± 3.82
Delta	40.22 ± 47.53	78.11 ± 70.61
Gamma	0.33 ± 0.36	0.472 ± 0.49
Theta	11.41 ± 10.06	7.78 ± 6.88

Figure A.3: **Mean deoxyHb concentrations for 0/1/2 and 3-back tasks (F7).** The signals are the mean waves of deoxyHb signals (fNIRS) for 0/1/2 and 3-back tasks at position F7 of the 10-20 system. The dashed lines represent the standard deviation.

Table A.3: **EEG bands powers p-values for subtraction and resting period.**

Variables	P-Value
Alpha Subtraction & Alpha Rest - F7	0.08
Beta Subtraction & Beta Rest - F7	0.22
Delta Subtraction & Delta Rest - F7	0.58
Gamma Subtraction & Gamma Rest - F7	0.03
Theta Subtraction & Theta Rest - F7	0.47
Alpha Subtraction & Theta Subtraction - F7	0.01
Alpha Rest & Theta Rest - F7	0.16
Alpha Subtraction & Alpha Rest - F8	1.0
Beta Subtraction & Beta Rest - F8	0.47
Delta Subtraction & Delta Rest - F8	0.37
Gamma Subtraction & Gamma Rest - F8	0.03
Theta Subtraction & Theta Rest - F8	0.58
Alpha Subtraction & Theta Subtraction - F8	0.01
Alpha Rest & Theta Rest - F8	0.16

Table A.4: **Comparison between EEG bands powers during 0/1/2 and 3-back tasks (F7).** The EEG signals were acquired in the position F7 of the 10-20 system (left side).

Power	0-back	1-back	2-back	3-back
Alpha	0.39 ± 0.24	0.35 ± 0.22	0.24 ± 0.10	0.32 ± 0.21
Beta	0.45 ± 0.23	0.65 ± 0.29	0.57 ± 0.43	0.71 ± 0.55
Delta	6.83 ± 7.43	5.15 ± 6.67	4.20 ± 5.06	4.86 ± 4.17
Gamma	0.033 ± 0.032	0.045 ± 0.040	0.040 ± 0.030	0.038 ± 0.028
Theta	0.80 ± 0.40	0.94 ± 0.72	0.85 ± 0.46	1.10 ± 1.13

Table A.5: **Comparison between EEG bands powers during 0/1/2 and 3-back tasks (F8).** The EEG signals were acquired in the position F8 of the 10-20 system (right side).

Power	0-back	1-back	2-back	3-back
Alpha	0.28 ± 0.14	0.25 ± 0.10	0.21 ± 0.08	0.28 ± 0.13
Beta	0.55 ± 0.37	0.64 ± 0.48	0.64 ± ± 0.70	0.90 ± 1.10
Delta	4.11 ± 4.50	3.50 ± 3.05	3.47 ± 2.93	4.48 ± 4.31
Gamma	0.042 ± 0.025	0.054 ± 0.033	0.035 ± 0.029	0.071 ± 0.086
Theta	0.99 ± 0.57	0.98 ± 0.61	0.79 ± 0.5	1.36 ± 1.22

Table A.6: **EEG bands powers p-values for 0/1/2 and 3-back tasks (F7).** The EEG signals were acquired in the position F7 of the 10-20 system (left side).

Variables - F7	P-Value
Alpha 0-back & Alpha 1-back	0.59
Beta 0-back & Beta 1-back	0.31
Delta 0-back & Delta 1-back	0.64
Gamma 0-back & Gamma 1-back	0.84
Theta 0-back & Theta 1-back	0.61
Alpha 0-back & Theta 0-back	0.007
Alpha 1-back & Theta 1-back	0.04
Alpha 2-back & Alpha 3-back	0.46
Beta 2-back & Beta 3-back	0.73
Delta 2-back & Delta 3-back	0.18
Gamma 2-back & Gamma 3-back	0.73
Theta 2-back & Theta 3-back	0.87
Alpha 2-back & Theta 2-back	0.008
Alpha 3-back & Theta 3-back	0.008

Table A.7: **EEG bands powers p-values for 0/1/2 and 3-back tasks (F8).** The EEG signals were acquired in the position F8 of the 10-20 system (right side).

Variables - F8	P-Value
Alpha 0-back & Alpha 1-back	0.32
Beta 0-back & Beta 1-back	0.40
Delta 0-back & Delta 1-back	0.84
Gamma 0-back & Gamma 1-back	0.40
Theta 0-back & Theta 1-back	0.39
Alpha 0-back & Theta 0-back	0.04
Alpha 1-back & Theta 1-back	0.04
Alpha 2-back & Alpha 3-back	0.15
Beta 2-back & Beta 3-back	0.50
Delta 2-back & Delta 3-back	0.73
Gamma 2-back & Gamma 3-back	0.18
Theta 2-back & Theta 3-back	0.11
Alpha 2-back & Theta 2-back	0.01
Alpha 3-back & Theta 3-back	0.007