UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

Comparing EEG-neurofeedback visual modalities between screen-based and immersive head-mounted VR

Mariana Silva de Almeida

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

Dissertação orientada por:
Professora Gina Caetano, Faculdade de Ciências da Universidade de Lisboa
Professor Athanasios Vourvopoulos, Instituto Superior Técnico

2022

"Success is not final, failure is not fatal: it is the courage to continue that counts."

— Winston Churchill

Acknowledgments

First of all, I want to express my deepest gratitude to those who believed in me throughout the duration of this work, especially the ones who didn't let me give up.

To my supervisors: Profesora Gina Caetano and Professor Athanasios Vourvopoulos, thank you for all the meetings, all the e-mails, and for never making me feel unattended. I know for a fact that without your support and immense availability, I wouldn't have gotten this far.

À faculdade, à PRAXE e a todos os docentes que enriqueceram a minha vida de estudante universitária: obrigada. Aos meus amigos do GinPong-Squad, um obrigada por todas as épocas de exame, os trabalhos de grupo, as noitadas e os jantares em odivelas, mas principalmente obrigada pela a entreajuda e a vontade que sempre tivemos de nos vermos triunfar.

Um obrigada com o coração cheio, a toda a minha família. Ás minhas irmãs, por acreditarem em mim e não me deixarem saber o que é ser filha única. Aos meus tios Sara e Rui por fazerem parte da minha vida desde sempre. Obrigada à minha madrinha por mostrar sempre orgulho em mim e por estar sempre disponível para me ajudar. A vocês João e Kiko por serem realmente os meus amigos de infância. E a ti padrinho, que sei que irias adorar ouvir que entreguei a tese e ias fazer uma piada sobre não saberes nada sobre o assunto, mas ias estar na primeira fila a torcer por mim. Mãe e Pai, devo-vos todo o meu percurso académico, obrigada por serem o meu porto de abrigo e por me guiarem até aqui. Obrigada pelo apoio e amor incondicional, sem vocês nada disto seria possível.

E a ti Rodrigo, obrigada por seres o meu parceiro de todos os momentos, por aturares a montanha-russa de emoções que foi esta tese e por me relembrares do que realmente sou capaz.

Resumo

O Neurofeedback (NF) é uma técnica de *biofeedback* não invasiva que visa a auto-regulação da atividade cerebral. Usado juntamente com técnicas como a eletroencefalografia, por meio de uma interface cérebro-computador esta técnica é capaz de transmitir, em tempo real, os dados fisiológicos que estão a ser recolhidos. Estes dados podem ser apresentados de várias formas, sendo que estímulos visuais ou auditivos são as formas mais comuns. Esta técnica tem sido objeto de vários estudos durante várias décadas e foi popularizada na finais dos anos 50. Desde então, o NF tem vindo a ganhar popularidade por ser uma técnica não invasiva, que não utiliza atividade elétrica ou magnética externa, não recorre a produtos farmacológicos e não tem efeitos secundários conhecidos. Atualmente, o NF pode ser utilizado em, pelo menos três ramos principais: (i) a nível clínico, como um método terapêutico para normalizar a atividade cerebral de pacientes que sofram de condições associadas a atividade cerebral irregular (como por exemplo, défice de atenção e epilepsia), a fim de atenuar os sintomas. (ii) treino de *peak performance*, para melhorar o desempenho cognitivo de participantes saudáveis; (iii) como uma ferramenta experimental que permite explorar a causa de eventos neurais, como oscilações neuronais. Com o crescimento do número de estudos que envolvem NF, o interesse pela criação de protocolos experimentais *standard* tem vindo a surgir. Apesar de isto ainda não ter sido alcançado, existem alguns elementos chave que estão presentes em praticamente todas as experiências da NF: (i) os participantes (ii) aquisição de dados, (iii) pré-processamento de dados, (iv) seleção do tipo sinal de *feedback*, ou seja a forma que queremos que a ativade cerebral seja apresentada aos participantes e (v) extração da *feature* de interesse, ou seja o que queremos que seja modulado durante a experiência.

Apesar de todas as aplicações e o elevado número de estudos que argumentam sobre a utilidade desta técnica, existem ainda algumas incertezas acerca do qual será o método mais eficiente para o treino de neurofeedback. Por exemplo, ainda não existe nenhum protocolo pré-definido que refira como o *feedback* deve ser apresentado (através de estímulos visuais, auditivos ou ambos) Ainda assim, a forma como o estímulo é apresentado pode fortemente influenciar o protocolo de treino e, consequentemente, o resultado da experiência. Por esta causa, é necessário testar as várias formas possíveis de apresentar os estímulos de forma a que se possa fazer uma comparação.

Nesta tese, foi efetuada uma comparação sistemática entre dois tipos de modalidades sensoriais visuais: (i) estímulo visual em 2-D (ii) estímulo visual na forma de realidade virtual imersiva. Esta análise foi feita com o objetivo de comparar a eficácia que cada modalidade teria sobre os resultados do treino de NF. Foram utilizados dados de dois estudos anteriores: (i) **Estudo I** - no qual foi desenvolvido e implementado um sistema de NF e de forma a comparar a eficácia de duas modalidades sensoriais: visual e auditiva; **(ii) Estudo II** - no qual foram implementados algoritmos de conectividade funcional e posteriormente integrados num software *opensource* de processamento de EEG (OpenViBE),o qual foi utilizado num procedimento experimental numa experiência de NF em realidade virtual.

Ambos os protocolos experimentais tinham como alvo o aumento da banda alfa superior (*upper alpha*), medida no elétrodo Cz através de eletroencefalografia. A organização temporal dos protocolos

experimentais foi equivalente em ambos os estudos. Cada sessão de treino foi dividida em quatro partes, começando com uma aquisição de dados inicial (*pre-baseline*) de 4 minutos durante a qual, os participantes não iniciavam nenhuma tarefa. Esta fase permitia determinar a banda alfa individual (IAB) Depois de isto, iniciava-se a sessão de treino, que foi dividida em 5 *sets*, em que cada *set* continha 3 blocos que, individualmente, consistiam em aquisições de 1 minuto. Entre cada *set*, faziam-se pausas de 10 ou 15 minutos, a pausa era mais longa entre blocos. O tempo total de treino daria cerca de 37 minutos, por sessão. Em ambos os estudos, cada participante participou em 4 sessões que foram conduzidas em dias consecutivos, em aproximadamente na mesma altura do dia.

Os dados experimentais foram posteriormente divididos em dois grupos, consoante cada modalidade: o grupo de modalidade *Screen-Based* (4 F ; 4 M) e o grupo *Immersive-VR* (2 F ; 2 M). Foi escolhido e aplicado um critério para seleção de participantes considerados *non-learners*, ou seja participantes incapazes de regular a sua atividade cerebral durante a experiência. No grupo da modalidade *Screen-Based* foram encontrados quatro *non-learners*, enquanto que no grupo de *Immersive-VR* foi encontrado um. Foi aplicado um extenso protocolo de tratamento e limpeza de dados, destacando a Análise Independente de Componentes (ICA) como o principal método de remoção de artefactos. Este é baseado na divisão do sinal de EEG em componentes independentes, através de métodos estatísticos. Após esta separação é possível analisar componente a componente de forma a remover os mais ruidosos. Por fim, eficácia do treino foi medida através do cálculo da amplitude relativa da banda de interesse (UA), pela definição de índices que medem a capacidade de aprendizagem e aplicação de testes estatísticos.

Os resultados demonstram uma evolução positiva ao longo do treino, para ambas as modalidades, exclusivamente dentro de cada sessão de treino. Ou seja, em cada sessão, ambos os grupos mostraram a capacidade de aumentar a sua amplitude da banda alfa superior (*upper alpha*). No entanto, comparando sessões, os resultados mostraram ser inconclusivos e não apresentam indícios deste aumento. O efeito do treino não se limitouo à análise do intervalo de frequência de interesse, foram também analisadas as restantes bandas de frequência. Tais se revelaram, irrelevantes na banda Beta, mas mais evidentes nas bandas Theta e LA. Embora o tamanho da amostra não fosse suficiente para tirar conclusões estatísticas relevantes e embora seja necessária uma investigação mais aprofundada, o trabalho apresentado nesta tese mostrou que modalidade Immersive- foi mais eficaz em aumentar o parâmetro de feedback (RAUA) dentro das sessões.

A maior limitação a ser destacada neste estudo foi tamanho da amostra. Uma vez que um grupo continha dados de 8 sujeitos e o outro de 4 e após a análise das capacidades de aprendizagem, estes número ficou ainda mais reduzido (VR = 3; VIS = 4). Pelo que determinar de forma precisa o efeito do treino de NF tornou-se uma tarefa impraticável. Na verdade, um tamanho de amostra limitado reduz efetivamente o poder estatístico do estudo e aumenta a margem de erro. Dado isto, uma coisa a ter em conta em estudos futuros seria realizar mais sessões de treino com ambas as modalidades de sensoriais, para aumentar o tamanho da amostra.

Palavras-chave: Neurofeedback; Modalidade sensorial; Realidade Virtual Imersiva; Banda Alfa Superior Individual; EEG

Abstract

Neurofeedback (NF) can be defined as a form of biofeedback that trains subjects to have self-control over brain their functions, by providing real-time feedback of their own cerebral activity. This activity can be presented in various forms, with auditory and visual feedback being the most common. Recently, NF has been investigated as a potential treatment for various clinical conditions associated with abnormal brain activity or cognitive capacities. However, the greater research focus is not discussing how the feedback should be presented. The chosen modality for any NF training system may strongly influence the training protocol and consequently the outcome of the experiment.

In this thesis, a systematical comparison between two different type of visual modalities (Screen-Based vs. immersive-virtual reality (VR)) was performed with the goal to evaluate the effectiveness of each modality on the NF training results. Data from two previous studies, recorded on healthy participants, in protocols that targeted the increase in the upper alpha (UA) band power measured at the EEG electrode Cz was used. This was then divided into two modality groups: Screen-Based modality group (N = 8) and the Immersive-VR group (N = 4). An extensive data processing and cleaning protocol was applied to both groups and the training effectiveness was measured through band power calculation, the definition of learning ability indexes and the application of statistical tests. Results showed that, both groups had a generally positive training effect within sessions, however data regarding different sessions is inconclusive and does not show clear evidence of up-regulation of the target feature. Additionally, when only considering within-session evolution, only the Immersive-VR modality group was able to maintain an increasing trend in all sessions.

One of the main limitations of this study was the sample size, which was too small to determine the precise effect of NF training. Future work requires, not only an increase in sample size but also, the definition and incorporation of learning predictors that allow the pre-selection of subjects before the training sessions, in order to prevent high number of non-learners.

Keywords: Neurofeedback; feedback modality; virtual reality; individual upper alpha; EEG

Contents

List of Figures

List of Tables

Acronyms

ADHD attention deficit hyperactivity disorder. 1, 7, 10, 14, 46

BOLD blood-oxygen-level dependent signal. 8

CAVE-VR Cave Automatic Virtual Environment. 14

DMN Default Mode Network. 7

EC eyes closed. 20

EEG Electroencephalography. viii, 1–9, 11, 13, 18–23, 26, 29

EO eyes open. 20

FC Functional Connectivity. x, 13, 18, 19, 21

fMRI functional Magnetic Resonance Imaging. 1, 6, 8

fNIRS Functional Near Infrared Spectroscopy. 1

HMD Head-mounted display. 19, 22

HTF higher transition frequency. 20, 21, 26

IAB individual alpha band. 20, 26, 35

IAF individual alpha frequency. 3, 26

ICA Independent Component Analysis. viii, 5, 9, 24, 26

LA lower alpha. 41, 45, 46

LTF lower transition frequency. 20, 26

MEG Magnetoencephalography. 1, 8

NF Neurofeedback. iii, v, 1, 2, 6–11, 13, 14, 18–21, 28, 45, 46

NFT Neurofeedback training. 1, 8, 13, 14

NIRS Near Infrared Spectroscopy. 8

PSD power spectral density. 20, 27

PSPs postsynaptic potentials. 2

RA relative amplitude. 26, 27, 38

RAUA relative amplitude of the upper alpha. x, 21, 22, 28, 29, 33–35, 37, 44–46

SMR sensorimotor rhythm. 11

UA upper alpha. v, x, 1, 13, 14, 18, 19, 21, 35, 39, 44, 45

VR virtual reality. v, 1, 2, 12–14, 18, 19, 22, 29, 32, 34, 36, 39, 45, 46

Chapter 1

Introduction

1.1 Context and Motivation

Neurofeedback (NF) represents a form of monitoring brain activity, which by providing a closed feedback signal and measuring brain waves simultaneously, trains the participant to self-regulate the targetted brain activity (Marzbani et al., 2016). The feedback signal can be presented in various forms being auditory and visual feedback the most common. The neural activity may be acquired through Electroencephalography (EEG), Magnetoencephalography (MEG), functional Magnetic Resonance Imaging (fMRI) or Functional Near Infrared Spectroscopy (fNIRS).

Neurofeedback has been the object of various studies for several decades, and it was popularized in the late 1950s and early '60s through Joe Kamyia and Barry Sterman's work (**?**). Dr. Kamyia experiment aimed to train human subjects to discriminate alpha from non-alpha states (Frederick, 2012) and, by using a simple reward system, he discovered that people could learn to alter their brain activity. Dr Barry Sterman's ran several experiments to assess if cats could increase their sensory motor rhythm.

Since then, NF it has been growing in popularity, specifically among those who study methods for influencing brain activity, on account that, NF is a non-invasive technique that does not introduce any external electrical or magnetic activity, or pharmacological products into the brain and that there are no known side-effects (Niv et al., 2013). Recently, NF has been investigated as a potential treatment for some diseases or disorders, such as attention deficit hyperactivity disorder (ADHD) and epilepsy, as a way to normalize brain activity and enhance cognitive capacities (Enriquez-Geppert et al., 2017).

Despite the fact that a high number of studies have indicated NF to be beneficial for several clinical applications, in the current literature there is still some uncertainty about methodological factors that may affect the effectiveness of Neurofeedback training (NFT) procedures. One such example is the lack of guidance to how the feedback should be presented (visual feedback, auditory, combined, etc.). Some authors argue that the chosen modality for any NFT system may strongly influence the training protocol and consequently the outcome of the experiment. Despite this, only few studies within the last few years, have compared the effects of different feedback modalities (Bucho et al., 2019; Berhanu et al., 2019; Accoto et al., 2021).

In this thesis, we aim to compare the EEG-NFT responses to the standard 2-D visual and the immersive-virtual reality (VR) visual modalities. We will use data from two data sets previously recorded on healthy participants, in protocols that targeted the increase in the upper alpha (UA) power band: **i)** the first, being an EEG-NFT training protocol using visual and auditory modalities, and **ii)** the second, an EEG-NFT training protocol using VR. We aim to understand if the type of visual modality (2-D vs. immersive-VR) affects the learning outcome for EEG-based NF traning.

1.2 Thesis Outline

This thesis is divided into five chapters. The first and current chapter gives a theoretical introduction to concepts that are relevant to this study, i.e. EEG, NF and, VR as well as presents a summary of the pertinent literature. Chapter 2 gives an insight into the work methodology, starting with a description of the data, the training protocol, and feedback modality and ending with a full description of all methods used to clean and analyze the data. The results are shown in chapter 3 and discussed in chapter 4. Finally, the last chapter contains the overall conclusions as well as this work's limitations and possible subjects that could be addressed more in-depth in future studies.

1.3 Electroencephalography

EEG can be described as the electrical activity produced by human brain neurons firing, which is typically measured at the brain scalp (Hu and Zhang, 2019). Hans Berger, a German psychiatrist, who in 1924 recorded the very first EEG in the scalp of a human head, is considered to be the father of EEG (Louis et al., 2016; Schomer and Lopes da Silva, 2010). During an EEG recording, electrodes are positioned on the scalp to detect electrical signals that are produced as a result of the brain's synchronized neuronal activity. Each individual EEG electrode's voltage can be thought of as a change in the potential between the active electrode and the reference electrode over time.

1.3.1 Neurophysiology

Physiologically, EEG measures voltage fluctuations resulting from ionic current within the neurons and reflects correlated synaptic activity caused by post-synaptic potentials (Purves et al., 2018). These voltage fluctuations are measured by a set of electrodes placed along the scalp and given its low spatial resolution, the main utility of EEG is in the evaluation of dynamic cerebral functioning. (St. Louis et al., 2016). Furthermore, considering the fact that it is a non-invasive method, simple to use, and with a high temporal resolution, this makes it one of the most popular neuroimaging techniques.

The majority of the electrical activity recorded on the scalp is generated by pyramidal neurons (Louis et al., 2016), which are characterized by their apical and basal dendrites that are oriented perpendicular to the cortical surface (Murakami and Okada, 2006).

Electrical currents are generated, at the brain level, during two types of neural activity: action potentials and post-synaptic activity. The first, results from changes in the membrane potential of the neurons, which is caused by the flux of ions and the latter arises from the communication between two different neurons. Despite its large intracellular amplitude, action potentials, exhibit a low extracellular amplitude. Additionally they do not last long enough to produce changes that are measurable on the scalp. This means, that post-synaptic activity is responsible for most of the signal in EEG.

During synaptic activity, the electrical potential generated by a single neuron is too low to be detected by EEG (Nunez et al., 2006), consequently the activity recorded by the electrodes represents the summation of the inhibitory or excitatory postsynaptic potentials (PSPs). These potentials alone are too small, hence only a large number of PSPs that are simultaneous and in the same direction will generate EEG waves on the scalp (Kirschstein and Köhling, 2009). To summarize, the EEG signal represents the activity that arises from all PSPs that are synchronized and similarly oriented, as for cortical pyramidal neurons.

1.3.2 Brain Waves

During an EEG analysis it is possible to detect patterns of the electrical activity of the brain, usually referred to as **brain waves**. Electrical voltages in the brain scalp oscillate at a few microvolts, forming brain waves, which are characterized by their amplitudes and frequency (Abhang et al., 2016).

There are five widely recognized brain waves and their frequencies vary by which state (e.g., alert wakefulness, drowsiness, sleep) the individual is in. The main frequencies of human EEG waves are listed below, along with their characteristics.

- **Delta waves** δ (0.5–4 Hz): Physiologically seen in deep sleep or normally in children, and prominent in the frontal-central head regions. Pathologically, this type of rhythm can appear in awake states in cases of encephalopathy and focal brain dysfunction. (Nayak and Anilkumar, 2019)

- **Theta waves** θ (4–8 Hz) This rhythm can be seen in drowsiness or in the early stages of sleep being most notable in the fronto-central head regions. Enhanced frontal theta activity is presented during heightened emotional states in children or young adults. Focal theta activity during awake states can be a sign of focal cerebral dysfunction. (Nayak and Anilkumar, 2019)

- **Alpha waves** α (8–12 Hz): Present in normal awake EEG recordings in the occipital region, best identified when the eyes are closed or during mental relaxation and, on the other hand, it is usually attenuated by eye-opening or mental effort (Nayak and Anilkumar, 2019). Clinically, the slowing of the background alpha rhythm may be a sign of cerebral dysfunction.

  Some studies argue that there is an evident subdivision within the alpha band, into lower and upper alpha band, and that each sub-band reflects different cognitive processes. Klimesch et al. (1990) used the term individual alpha frequency (IAF), which presents the maximum power in the alpha band in individual subjects, and stated that lower and upper alpha were defined according to this frequency spectrum - lower alpha range can be defined as [IAF - 2, IAF] and upper alpha as [IAF, IAF + 2], but other definitions can also be found in the literature.

- **Beta waves** β (12–35 Hz): The most frequently seen rhythm in normal adults and children, it is most prominent in the frontal and parietal head regions. It usually occurs in an alert state or when individuals present direct attention toward external stimuli. It often increases in amplitude during drowsiness and right after you fall asleep. Most of the sedative medications increase the amplitude and quantity of beta activity in individuals.

- **Gamma waves** γ ($>$ 35 Hz): It is distributed throughout several cerebral structures and it participates in various cerebral functions, such as perception, attention, memory, consciousness, synaptic plasticity, and motor control (Amo et al., 2017).

1.3.3 Electrode Placement

To record an EEG signal, several electrodes are placed on the scalp. These electrodes act as a bridge between lead wires, which conduct electrical current through the flow of electrons, and the human tissues (Schomer and Lopes da Silva, 2010). Similar to what happens with other types of imaging techniques, a standard electrode positioning method was created in order to ensure consistency among studies.

The most common electrode placement system is the **international 10/20 system**, which consists of 21 electrodes plus one ground electrode in which a system of lines is created and the electrodes are

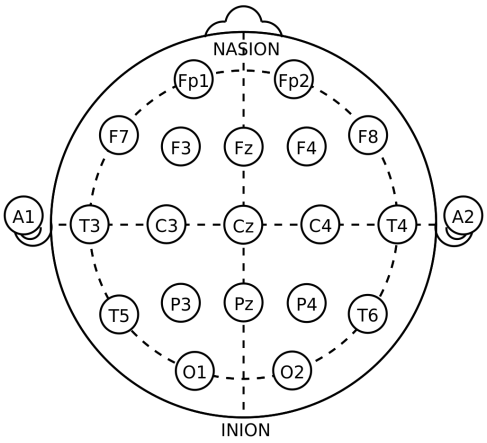

Figure 1.1: Electrode locations of International 10-20 system for EEG recordings (Commons, 2010).

placed at either 10% or 20% of the total front–back or right–left distance of the skull, which corresponds to the length of these lines (Figure 1.1). Each electrode is identified by letter and a number. The letters give out information about the area of the brain from which the electrode is reading the signal: frontalpolar (Fp), frontal (F), temporal (T), parietal (P), occipital (O), and central (C). Additionally the "A", that can also be indicated as "M" (mastoid process) refers to the prominent bone process usually found behind the ear. Even numbers refer to the right brain hemisphere, whereas odd numbers refer to the left brain hemisphere and these increase with their distance from the midline. Finally, electrode sites labeled with "z" (zero) refer to the sagittal midline of the brain.

In order to diminish ground-related noise, EEG uses online reference electrodes. This means that the signal in each electrode is given by the difference between the electric potential in its location and the signal of a reference electrode, such as Fz, Pz, Cz, mastoids, or earlobes. However, after the data has been recorded it is possible to change its reference, i.e. to express the voltage in each electrode with respect to another reference - this technique is called re-referencing. The optimal reference always depends on the signal of interest, however, the two most commonly used re-referencing techniques include the average mastoids or the average of all scalp channels (Leuchs et al., 2019).

1.3.4 Artifacts

Like any other kind of signal, cerebral activity recorded by an EEG system is often affected by other electrical activity of non-neural origin, these are called **artifacts**. On both clinical and research applications, artifacts are a major obstacle in the interpretation of any EEG signal (Seok et al., 2021). The most common origin of an EEG artifact is usually internal, this means that it is caused by physiologic functions, some examples include:

- **Ocular Artifacts**: These artifacts are seen in nearly every conscious individual during an EEG and are crucial to correctly identify different stages of sleep. Ocular artifacts are noticeable in the frontal region of the head and are a result of eye movements or blinks, in which the eye acts as an electric dipole. In fact, the cornea is electropositive relatively to the retina and that will generate a difference in the current potential that can be measured in both the horizontal and vertical plane (Tatum et al., 2011).

  Vertical eye movements are most commonly presented in form of a blink. During blinks, the eyes create an upwards movement (ie, Bell phenomenon), which causes the cornea (that acts like

a positive pole) to move closer to the frontopolar electrodes (Fp1-Fp2), resulting in a downward deflection (Figure 1.2). Logically during the opposite movement, an upward deflection will be generated (Louis et al., 2016). Saccades are described by the small movements made while reading or even larger movements made while gazing around a room and in an EEG recording they may appear as small and rapid deflections in the frontal areas. Finally, lateral eye movements affect mostly electrodes F7 and F8.

- **Muscle Artifacts:** Another common artifact during EEG recordings is caused by any type of myogenic (muscle) movement. The temporalis and frontalis muscles are the main source for this type of artifact by participating in simple movements like swallowing or the related movement of the tongue. Muscle artifacts are seen as high-frequency and fast bursts of activity (Tatum et al., 2011), as depicted in Figure 1.3.

- **Cardiac Artifacts:** Also referred to as pulse artifact, this type of artifact arises from the effect that cardiac activity has on an EEG signal. Despite the amplitude of an ECG (Electrocardiogram) signal being low when observed on the scalp, some electrodes might detect heartbeats which would be seen as rhythmic pattern distortion overlapping the EEG signal (Bitbrain, 2020).

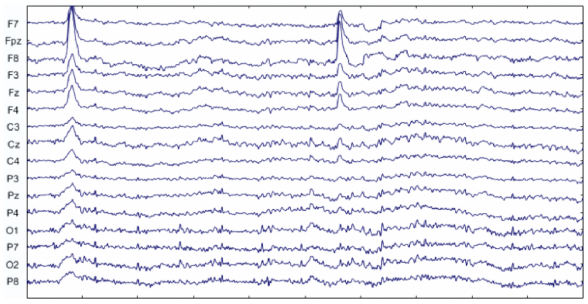

Figure 1.2: Representation of an EEG recording time series, per channel. The high amplitude voltage peaks present in frontal electrodes (F7, Fpz, F8) correspond to blink artifacts (Bitbrain, 2020).

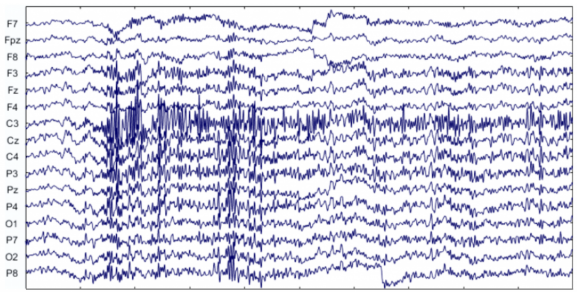

Figure 1.3: Representation of an EEG recording time series, per channel. The high-frequency portions present in the signal occur during jaw movements (Bitbrain, 2020).

The EEG recording can also be contaminated by external factors, i.e., by non-physiological artifacts. Mostcommon non-physiological artifacts include noise generated by monitoring devices, movement of the cables connecting the electrodes and the amplification system or incorrect placement/bad contact of reference channels (Bitbrain, 2020).

Considering this, it is mandatory to apply methods that effectively detect and eliminate any artifacts from EEG recordings. To this day, several methods have been developed and can be split into two categories: **i)** the estimation of the artifactual signals using reference channels and, **ii)** the other through decomposition of the EEG signal into other domains (Jiang et al., 2019). Some examples of these techniques include Regression Methods, Wavelet Transform (WT), Blind Source Separation (BSS), and Independent Component Analysis (ICA). Figure 1.4, represents a chart of the percentage of the number of artifact removal methods referred in literature in the past five years. It shows that the BSS-based methods and especially ICA are the most commonly used algorithms.

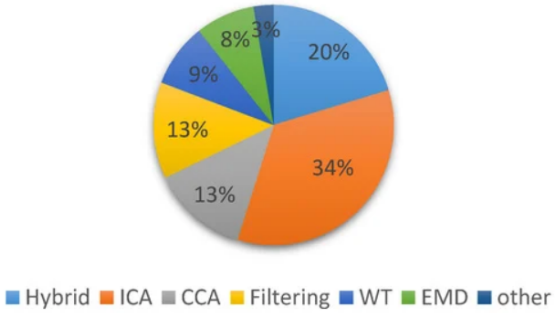

Figure 1.4: Percentage of the number of artifact removal methods referred in literature in the past five years; Independent Component Analysis (ICA), Canonical Correlation Analysis (CCA), Wavelet Transform (WT), Empirical mode decomposition (EMD) (Jiang et al., 2019).

1.4 Neurofeedback

NF can be defined as a biofeedback technique that helps individuals to gain control over their brain activity, by making them respond to a display of their own brainwaves or other electrical activity of the nervous system in real-time (Niv et al., 2013; Ros et al., 2014). The first NF experiment took place in mid-1960s through Joe Kamyia's work (?), which showed that voluntary control of human brain oscillations was possible with sensory feedback from a brain-computer interface (BCI). Since then, this technique has gained in popularity and to this day, it is seen not only as a method for cognitive enhancement in healthy subjects but also as a therapeutic tool (Enriquez-Geppert et al., 2017).

1.4.1 Neurofeedback Mechanisms

Neurofeedback is based on two facts, the first is that the state of the brain is reflected in parameters of any EEG recording, and the second fact is, that the human brain is capable of "learning" new cognitive states (Kropotov et al., 2010). However, to this day, one current limitation in the literature is the lack of research and consensus about the underlying mechanisms of neurofeedback, i.e. there are still some arguments on how this technique really work.

Niv et al. (2013), published a review that summarizes different perspectives about the mechanisms underlying NF. First of all, the authors define **neuroplasticity** as the ability of the brain to reorganize and form new synaptic connections, mainly in response to learning or experience. The authors suggested that NF may be effective by enhancing the strength between synapses through repeated firing, this means that by causing a long-lasting increase in signal transmission between two neurons, NF will induce brain plasticity. This supports the fact that NF experiments generally involve several training sessions. Furthermore, several studies show evidence of this premise by taking advantage of NF protocols to make electrical changes in the brain, together with cognitive improvements (Ros et al., 2014; Kober et al., 2017). Niv et al. (2013), suggest that this factor could be investigated with a fMRI connectivity analysis before and after treatment with NF.

Secondly, it is proposed that NF may be strengthening and regulating connectivity within and between networks of neurons present in the cortex. Several studies present the brain as a model, i.e. a network model that is characterized by high-density local connectivity of neurons that are hierarchically connected with surrounding neurons. For example, in the cortex, there is dense connectivity inside each cluster of neurons but this connectivity is diminished between clusters. Some studies suggest that psychopathologies, such as Alzheimer's disease or schizophrenia could be a consequence of

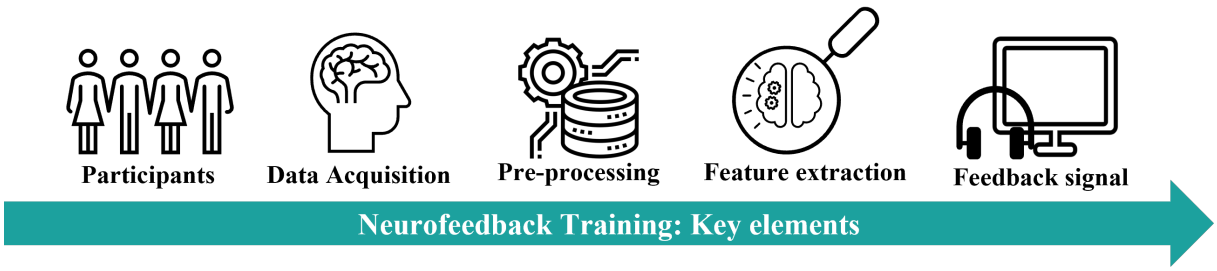

Figure 1.5: Neurofeedback experiment - The key elements.

dysfunctions in these networks (Zoefel et al., 2011; Menon et al., 2011). If this was true, NF methods would be effective in improving various symptoms of psychopathologies. This premise is supported by some studies that show that NF is capable of tuning brain pathological oscillations towards homeostasis (Ros et al., 2014; Marzbani et al., 2016)

Finally, in this review three neurocognitive networks are identified as responsible for brain self-regulation and are suggested to be regulated by NF: *(i)* the Default Mode Network (DMN), which comprises several portions of the cortex and serves as a regulatory function of the brain, being usually strengthened by meditation, *(ii)* the Central Executive Network (CEN), which is involved in all kinds of executive functions, and *(iii)* the Salience Network (SN), that manages emotional, somatic and autonomic information. Deficits found in any of these three networks are commonly associated with pathologies that are characterized by abnormal oscillatory brain activity (e.g. Alzheimer's disease, autism or ADHD). The fact that these networks oscillate at low cortical frequencies is the bridge that connects them to NF methods, considering that such frequencies are used in recent neurofeedback protocols. Hence, NF may be directly affecting these networks serving as a potential successful therapeutic tool.

1.4.2 EEG neurofeedback experiment

Neurofeedack research is evolving and with that comes the creation of several experimental protocols. Although these experiments present some differences, generally there are some key elements that are coincident in almost every NF experiment. In this section, based on the work by Enriquez-Geppert et al. (2017) , five of those elements are identified and further explored: **(i)** participants (learner) **(ii)** data acquisition, **(iii)** pre-processing, **(v)** feature extraction **(iv)** feedback signal.

Participants

Any NF experiment requires participants, whose brain's activity is measured and processed. The feature of interest is extracted from the acquired data and is then transformed so it can be presented to the participants in a form of instant feedback that allows them to modulate their own brain activity. Unfortunately, not all participants respond in the same way to a NF training session, some individuals are not as capable to learn how to regulate their brain patterns, i.e.they present less learning ability, or some might even be identified as non-learners (participants who cannot achieve self-regulation of their own brain activity) (Nan et al., 2015). The learning ability of a group of participants in a NF experiment will highly affect the success of the session and because of this, few recent studies (Ninaus et al., 2015; Reichert et al., 2015; Witte et al., 2013; Weber et al., 2011) have been conducted with the purpose of identifying factors that predict the success of the training. Such factors include the characteristics of EEG

activity in the resting state, along with specific neuroanatomical structures' volume and concentration or even intrinsic aspects related to the brain's structural properties.

Learners vs. Non-learners

In the current literature, there are several studies on the effectiveness of neurofeedback however, there is a lack of information and data to explain the failure of neurofeedback, and therefore it is important to explore the learning ability as a factor that affects neurofeedback results. Reiner et al. (2018), reflect on the need for finding methods that can successfully differentiate learners and non-learners, since this is the most relevant way to measure the impact of the learning process on the effectiveness of neurofeedback.

Some studies recognize different types of non-learners identification criteria. Okumura et al. (2019), explores predictive training assessment indices that could predict learners, prior to training. Learners were classified based on cognitive and neurophysiological measures taken during a series of Stroop tasks. Wan et al. (2014), found higher alpha activity at a resting state before training a significant predictor for better learning indices for alpha NFT. Nan et al. (2015), aimed to predict the learning ability in training that involved beta/theta ratio training. They encountered that low beta in eyes-open resting state measured before NF and the beta-1 amplitude in the first training block could predict the learning ability across training sessions.

Data acquisition

Acquiring data for a NF experiment can be achieved by using any type of technique that records brain signals. Some examples include EEG, MEG, fMRI and NIRS.

EEG was the first modality to be used in a NF environment and characteristics such as high temporal resolution, inexpensiveness, safety, and straight forward usability make it, to this day, the most popular method for recording electrophysiological brain activity in research.

Similarly to EEG, MEG is an optimal technique for real-time feedback of brain processes, that measures the magnetic fields generated by electric currents in the brain (Enriquez-Geppert et al., 2017). It has a high temporal and a good spatial resolution and comparatively to EEG is less sensitive to signal distortion caused by the conductivity of the head tissues. Despite this, the fact that this technique requires extensive funding and is not in any way portable makes it a less popular data acquisition method than EEG in NF experiments.

Besides those two, high spatial-resolution fMRI and the NIRS are two additional ways of collecting data for NF experiments. The fMRI technique provides a spatial resolution of the order of the millimeter, and the blood-oxygen-level dependent signal (BOLD) method enables to estimate the feedback signal successful for the successful regulation of brain activity by fMRI-NF. Unlike NF withEEG, fMRI-NF protocols do not require an extensive number of training sessions (Thibault et al., 2018).

Near Infrared Spectroscopy (NIRS), similarly to fMRI, measures hemodynamic changes in hemoglobin, associated with neural activity, with the help of a portable cap that may have up to 50 channels. To detect changes in the attenuation of radiated light, it uses infrared emitting diodes and light detectors (Mihara et al., 2012). Furthermore, NIRS is relatively resistant to subject motion and requires less time for attachment without the use of paste, resulting in fewer limitations. While functional NIRS is less expensive, more artifact-resistant, and more portable than fMRI, its spatial resolution is lower, in the order of centimeters (Thibault et al., 2018).

To summarize, nowadays neurofeedback employs a variety of imaging techniques to guide voluntary control over both electromagnetic and hemodynamic changes in brain activity and each imaging modality has been used in various neurofeedback protocols that target different brain signals (Thibault et al., 2016).

Pre-processing

As described in Section *1.3.4*, all cerebral activity recordings can be affected by **artifacts**, which included EEG. Thus, to improve the efficiency of any neurofeedback system, strategies for eliminating artifacts from the signal must be developed. Enriquez-Geppert et al. (2017) mentions in his review that not treating eye artifacts in a NF training environment may result in participants falsely learning to modulate their eye movements rather than their brain activity.

As it is shown in Figure 1.4, ICA is the most popular method for artifact removal in current literature, and this is also true for NF training. This method includes extracting and separating statistically the components present in the signal in question. ICA algorithms emerged during the 1980s and ten years later became a very popular method not only, in biomedical signal analysis but also, in the financial and economic fields (Vullings et al., 2009). Figure 1.6 illustrates the ICA analysis procedure and its ability to filter out artifacts from the signal. Let's say we have an EEG recording with several channels and name it matrix X, when we apply the ICA the algorithm finds a decomposing matrix (W) that isolates the data and transforms it into a sum of temporally independent and spatially fixed components (Jung et al., 2000). After this, it is possible to look at each independent component and filter those which are thought to be artifactual, and in the end, the algorithm subtracts those chosen components and forms the new and corrected EEG data. ICA can easily isolate variations of potential due to blinks, eye movements or muscle contractions. However, the algorithm also finds components that do not fit into any of these categories and because of this any ICA analysis must avoid not only, under-correction but also the over-correction, which may lead to the loss of important data information.

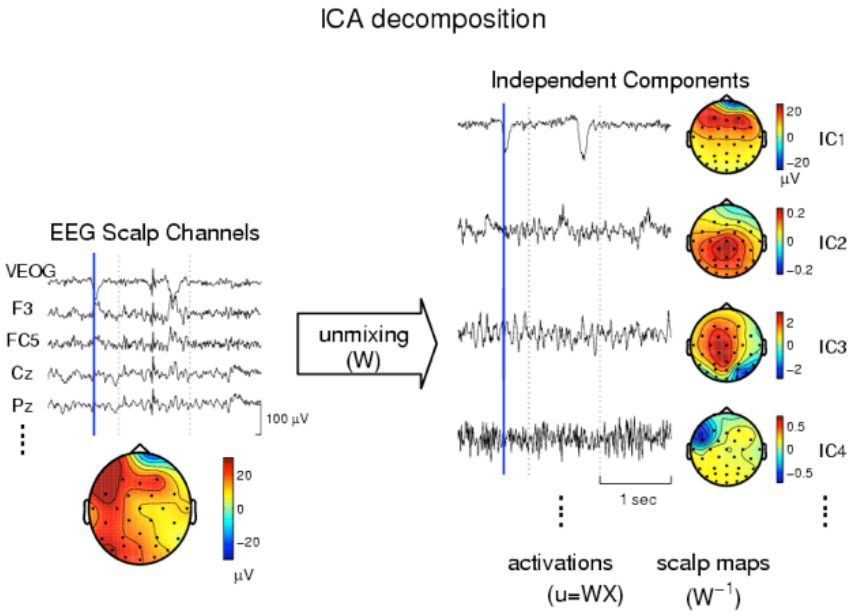

Figure 1.6: schematic illustration of the procedure

Feature extraction

Feature extraction is also a key element of NF experiments, and consists of selecting, from brain activity, the desired feature to be modulated and consequently, extracting it. Most therapeutic NF training protocols aim to modulate one specific frequency band, like theta, alpha, and beta which are linked to particular cognitive processes (Enriquez-Geppert et al., 2017; Marzbani et al., 2016) or the brain activity in the desired frequency range can also be determined by relative relation between two frequency bands (e.g., theta/beta ratio training).

Extracting the target feature is the last step before the feedback parameter is computed. After this, the computed feedback parameter is converted into a sensory stimulus that is fed to the participant via the chosen feedback modality.

Feedback signal

This element involves the conversion of the target feature into a sensory stimulus that is then introduced and presented to the subject (learner). The feedback is most commonly obtained through external stimulis, visual and auditory being the most common modalities (Davelaar, 2018). While the stimulis are being presented, the learner becomes aware of the changes taking place throughout the session and is then able not only, to modulate the desired brain patterns but also to memorize the current neural state.

Enriquez-Geppert et al. (2017) highlights the fact that they aren't many studies addressing different feedback modalities in the current literature, despite the fact that many authors argue the chosen feedback modality can strongly influence the outcome of any NF experiment (Accoto et al., 2021; Bucho et al., 2019; Huster et al., 2014). Due to these uncertainties in the literature, the choosing of the feedback modality is usually based on the participants' characteristics and the type of study being developed.

1.4.3 Neurofeedback Applications

Enriquez-Geppert et al. (2017) review's identifies three main fields of application for neurofeedback: (i) clinical field, as a therapeutic tool specifically for conditions related to deviating brain activity in patients, (ii) peak-performance training field and (iii) investigation field, as a tool to study the relation between brain neural events and cognitive functions. Evidently, NF can be applied in various situations apart from the ones mentioned above, but for the current work, the most relevant application is without a doubt at a clinical level, examples of this are explored below.

ADHD

Attention deficit disorder is one of the most common neurobehavioral disorders, it is usually chronic and with notable symptoms that can persist into adulthood (Wilens and Spencer, 2010). Currently accepted and more often applied treatments for this condition consist of pharmacotherapy combined with behavioral therapy, both have been proven to present serious limitations in this context (Enriquez-Geppert et al., 2019).

Thibault et al. (2016) points out that treatment for ADHD is the most researched application of neurofeedback. Studies show that slower brain activity (theta) and less beta activity is associated with the diagnosis of this disorder (Marzbani et al., 2016). Because of this, standard training protocols with ADHD attempt to decrease theta and/or increase beta power. Additionally, some studies also take advantage of NF to help diagnose this condition (Gnecchi et al., 2007).

Epilepsy

Epilepsy is a neurobiological disorder characterized by sudden bursts of electrical activity in the brain, called seizures. Treatment of epilepsy through neurofeedback was perhaps the first clinical application to be registered (Egner and Sterman, 2006). To this day, NF is known to effectively and reliably reduce epileptic seizures by means of sensorimotor rhythm (SMR) (12-15 Hz) training. For instance, Sterman et al. (1974) and Lubar and Bahler (1976) both conducted an SMR neurofeedback protocol and successfully were able to reduce seizure frequency.

Schizophrenia

Schizophrenia is a severe long-term mental health condition. It severely affects and interferes with a person's ability to think clearly, manage emotions, make decisions and relate to others (Marzbani et al., 2016). Current treatment methods have proved to not be sufficient to treat all symptoms and can also cause problems such as medication adherence, metabolic side effects, and comorbid mood disorders (Surmeli et al., 2012). This shows great necessity to explore better therapy methods for this condition. To this day, NF has also been proven to help improve the life of patient's with schizophrenia, for example, Surmeli et al. (2012), studied the effect of quantitative EEG-guided neurofeedback (NF) treatment in a population with this condition and out of 48 participants 47 showed clinical improvement after NF treatment.

1.5 Virtual Reality (VR)

virtual reality (VR) can be defined as ”an artificial environment which is experienced through sensory stimuli (as sights and sounds) provided by a computer and in which one’s actions partially determine what happens in the environment” (vr, 2022). Sherman and B (2019), identifies different key elements in experiencing VR: (1) the **people involved** not only in designing and implementing the environments but also the participants who experience them; (2) the **virtual world**, which is is defined as an imaginary space often manifested through a computer interface; (3) **immersion**, the idea of being physically present in a non-physical world and (4) **interactivity**, for any VR system to seem genuine, while the user is interacting with the environment, it should interact back by responding to the user’s actions. In fact, since the concept of VR, as we know today was first formed in 1968, several types of different formats have been created: (i) Non-Immersive VR; (ii) Semi-Immersive VR; (iii) Immersive VR; (iv) Augmented Reality; (v) Collaborative VR.

In the scope of this work, Immersive VR is the most relevant for the current work. The concept of ”immersion” refers to intense participation in an activity on an intellectual and emotional level and in VR technology is a measurable feature used in several research works (Accoto et al., 2021). There are two types of Immersive VR:

- Cave Automatic Virtual Environments (CAVE), in which the user is put in a room, usually squared, where projectors may be used on the floor and ceiling to create a highly immersive environment (Figure 1.7).
- Head-mounted display VR, used with the classic VR glasses and can be complemented with headphones and produce the immersive feeling of being in a simulated world (Figure 1.8).

Because the brain recognizes the virtual world as real, the ability to train and acquire knowledge or abilities in various contexts or environments is possible. For this reason, when it comes to the development of several skills, virtual environments are allegedly more successful than other digital methods.

Figure 1.7: Example of a CAVE VR environment (ST Engineering Antycip, 2022).

Figure 1.8: Example of a Virtual Reality Headset (Kiyoshi Ota, 2016)

1.5.1 Neurofeedback in Virtual-Reality

The improvement of software and hardware of computer leads to developing and improving VR technology and its applications. Contrary to popular belief, applications for virtual reality go way be-

yond media and entertainment, and nowadays this technology is being used in fields such as healthcare, aerospace, military, sports, and education.

More recently, VR started growing in popularity among neurofeedback researchers, who started to hypothesize if the illusions induced by a VR environment, would activate the brain areas of interest and therefore enhance the training procedure Accoto et al. (2021). Vourvopoulos et al. (2019), explored the effect of a VR EEG-neurofeedback could have in helping chronic stroke patients in motor recovery and movement re-learning. Accoto et al. (2021) investigated the effects of vividness on neurofeedback training with a CAVE-VR environment and concluded that increased performance was related to higher vividness during training.

Any NF training protocol aims to modulate a specific brain feature and it is plausible to assume that this modulation would be better achieved if the brain was being fed with a realistic feedback signal. Even though there are some limitations in the current literature on how the feedback should be presented, considering VR technology is based on the creation of close-to-reality environments, it ought to be considered as a strong candidate for a suitable feedback modality.

1.6 Previous works

This section presents a brief contextualization of the work carried out in this area, specifically studies that conduct NF experiments that either compared two different modalities or used VR as feedback. Tables 1.1, 1.2 and 1.3 summarize all the information presented below.

The present work has as reference **two previous main works**. Firstly, Bucho et al. (2019) developed and implemented a NF-training with the goal to **compare the effectiveness of feedback provided via two different sensory modalities**, visual and auditory respectively. An EEG-based NF protocol was implemented, targeting the individual UA band and working memory enhancement. The sample size consisted of 16 healthy participants, hence 8 individuals for each modality. The results revealed that both groups showed significant improvements in training sessions, but no significant improvements regarding working memory nor differences between groups. Hence, regarding the comparison of sensory modalities, the authors indicate that the sample size was small and that further investigation is required. The second most relevant work is by Berhanu et al. (2019), which consisted in implementing/validating a real-time computation algorithm for functional-connectivity-based EEG-NF while using immersive-VR for feedback, also compared to the standard UA-based EEG-NF protocol implemented by Bucho et al. (2019) but adapted to immersive-VR. The aim was to modulate the amplitude of the weighted-node-degree Functional Connectivity (FC) in the individual UA band and for the electrode Cz. The sample consisted of 8 healthy subjects, 4 who performed the standard UA EEG-NF protocol with immersive-VR, and 4 who performed the functional-connectivity-based EEG-NF experiment. The results showed an increase of the upper alpha and FC along training and that alpha training has an effect in FC. Moreover, a VR environment was revealed to be efficient as a stimulus delivery mechanism. Yet again, it is a very small sample size to draw definitive conclusions from.

To date, only a few studies compared different types of feedback modalities. Fernández et al. (2016), compared the efficacy of visual versus auditory modalities in a NF study with disabled children. The sample consisted of 20 learning-disabled children with an abnormally high theta/alpha ratio. During training, the target was to reduce the theta/alpha ratio. The results showed that auditory stimuli were more effective than visual.

Although they did not compare different types of modalities, Plerou et al. (2017) and Nazer et al. (2018), conducted relevant studies that evaluated the effectiveness of NFT. The first evaluated the brain

activity and activation differences between participants enrolled in NFT and controls and then inferred that the effect of NF therapy was significant. The latter aimed to determine the effects of NF training on verbal and visual memory and results showed that this type of procedure can be used to improve the memory of individuals.

Regarding the effect of UA training, Kober et al. (2017) investigated the effects of UA-based NFT on electrical brain activity and cognitive functions in stroke survivors and observed positive effects on memory functions and cortical "normalization" in a stroke patient. Additionally, Zoefel et al. (2011) aimed to evaluate the trainability of UA and its effects on cognitive performance. The sample consisted of 14 participants who were trained in five sessions within 1 week. Eleven out of fourteen participants showed significant training success. This means that, the UA was increased and the enhancement of cognitive performance was significantly larger for the NF group than for the control group.

Similarly to Berhanu et al. (2019), few authors have used an immersive VR environment as feedback modality. Yan et al. (2008) used VR to create a piece of immersive feedback information, with results suggesting that the attention of subjects had been strengthened after 20 training sessions, thus hypothesizing that a NF system could be an effective tool for treating children with ADHD.

Additionally, the authors studied the role of alpha oscillations in attentional control with a sample of 22 participants. Half of the individuals were presented with a 3D-VR environment while the other half received feedback in a 2D-VR environment. Results showed a larger learning rate in the 3D environment compared with the 2D group.

Finally, Accoto et al. (2021) investigated the effects of vividness in a Cave Automatic Virtual Environment (CAVE-VR) on neurofeedback training outcome and assessed the effect on working memory performance, with a sample of 21 participants, subdivided into 3 groups. Each group performed - in one of three versions of the same living room, presented at different levels of vividness in a CAVE-VR environment. Results showed that the most vivid feedback corresponded to a higher increase in neurofeedback performance and had a more positive effect on the subjects' motivation, concentration and reduced boredom and also improved working memory performance.

1.7 Objectives

In this thesis, we aim to explore if the type of visual modality (Screen-Based vs. Immersive-VR) affects the effectiveness of a NF training results. The main motivation for this thesis is: (i) the fact that NF has been growing in popularity, not only as a tool to better understand the brain functions, but also as potential method to treat few conditions related with stress, anxiety or difficulty to focus; (ii) the lack of studies in current literature that compare different types of feedback modalities and consequently their impact of the training effectiveness. This was achieved by:

- Using two data sets from previous studies and applying pre-processing tools and artifact removal algorithms to it.
- Systematically comparing the use of Immersive-VR with conventional visual reinforcement signals by identifying the training effect for both groups of subjects.

Table 1.1: Summary of two main previous studies.

Title (Authors)	Aim	Participants	Stimuli	Target feature	Results/Conclusions
Comparison of visual and auditory Modalities for upper-alpha eeg-neurofeedback **Bucho, Teresa et al. (2019)**	Compare visual and auditory NF signals in an Upper Alpha-NF protocol	16 healthy participants (8 for each modality)	Sensory modalities (visual and auditory)	upper-alpha in electrode Cz	Both groups were able to up-regulate UA in electrode Cz + the auditory sensory modality is just as effective as the visual (small sample)
Connectivity-based EEG-neurofeedback in VR **Berhanu, Cristiano (2019)**	1- Implement/Validate a real time computation algorithm functional-connectivity based NF; 2- Implement a VR environment for feedback; 3 - Compare NFT based on individual UA or FC in UA for a VR environment	8 healthy participants (8 for each experiment)	VR (three different paradigms; particle changing size, speed, proximity)	UA and UA-WND in electrode Cz	No evidence of an ability to self-regulate the amplitude of the upper alpha No cognitive (working memory) enhancements (small sample)

Table 1.2: Summary of Visual/Auditory feedback modalities studies.

Title (Authors)	Aim	Participants	Stimuli	Target feature	Results/Conclusions
Neurofeedback in Learning Disabled Children: Visual versus Auditory Reinforcement **Fernandez, Thalia et al. (2015)**	Comparing the efficacy of visual, vs. auditory reinforcers as a tool to treat learning disabled children	20 LD children with an abnormally high theta/alpha ratio	Visual (eyes open; white square) and Auditory (eyes closed; 500Hz tone)	EEG recording, targeting theta/alpha ratio	Auditory stimuli were more effective (in reducing theta/alpha ratio) than visual stimuli
EEG Analysis of the Neurofeedback Training Effect in Algorithmic Thinking **Plerou, A., et al. (2017)**	Evaluate the brain activity and activation differences between NFT participants and controls	182 participants (91 per group)	10 algorithmic tasks	Enhancement of alpha-theta ratio in the C4 region and SMR-low Beta ratio enhancement in the P4 region	Effect of real NFT was significant
Effectiveness of neurofeedback training on verbal memory, visual memory and self-efficacy in students **Mohammad Nazer et al. (2018)**	Determine the effects of neurofeedback training on verbal and visual memory and self-efficacy	26 students	Visual (displays the individual's different brainwaves as shapes)	Increase in the sensorimotor response (SMR) wave	NF can be used to improve the memory of individuals in various contexts

Table 1.3: Summary of studies with VR as a feedback modality.

Title (Authors)	Aim	Participants	Stimuli	Target feature	Results/Conclusions
Upper Alpha Based Neurofeedback Training in Chronic Stroke: Brain Plasticity Processes and Cognitive Effects **Kober, et al. (2017)**	Investigate the effects of UA NFT on electrical brain activity and cognitive functions in stroke survivors	2 chronic stroke patients Healthy elderly control group (N=24)	Visual feedback (vertically moving bars)	EEG-based NF protocol targeting the UA	Positive effects on memory functions and cortical "normalization" in a stroke patient
The Effect of Neurofeedback Training in CAVE-VR for Enhancing Working Memory **Floriana Accoto et al. (2018)**	Investigate the effects of vividness in a CAVE-VR on NFT outcome and assess the effect on working memory performance	21 participants (15M, 6F)	VR with 3 versions of the same living room at different levels of vividness	Increasing the amplitude of the brain activity in UA band	Most vivid = higher increase in NF performance, more positive effect on user's motivation, concentration and reduced boredom, improved working memory performance
Frontal Alpha Oscillations and Attentional Control: A Virtual Reality Neurofeedback Study **Berger et al.(2018)**	Role of alpha oscillations in attentional control	22 participants	50% feedback on alpha amplitude in a 3D VR environment and the other 50% received feedback in a 2D environment	Increase their level of alpha amplitude	Larger learning rates in the 3D compared with the 2D group

Chapter 2

Methods

This chapter is divided into two sections, the first half covers the characteristics of the data retrieved from previous works (Bucho et al., 2019; Berhanu et al., 2019), starting with details about the participants, followed by the equipment and signal acquisition elements and finally, the neurofeedback training protocol. The second section includes a detailed description of the data pre-processing and data analysis steps, that allowed the comparison of the effectiveness of the two different visual feedback modalities.

2.1 Data and Training Description

2.1.1 Participants

Study I (Visual vs Auditory)

As previously described in Table 1.1, the aim of this study was to compare the effectiveness of two sensory modalities (visual and auditory) in the context of an EEG-based NF protocol that aimed at the increase of the individual UA. The experiment counted with a total of 16 healthy participants, between 19 and 28 years old and who also completed self-assessment health-related surveys. Two groups were created by randomly allocating participants while trying to keep an equal number of females and males in both groups, which were named after the modality being tested - the visual group and the auditory group. The demographic information is presented in Table 2.1 as, adapted from Bucho et al. (2019).

Group	Visual	Auditory
Age (years; mean $\pm$ SD)	22.5 $\pm$ 2.73	22.88 $\pm$ 1.25
Gender (F/M)	4 F ; 4 M	5 F ; 3 M
Dominant Hand (L/R)	0 L ; 8 R	0 L ; 8 R

Table 2.1: Demographic characteristics of the visual and the auditory groups, respectively (adapted from Bucho et al. (2019)) .

Study II (Head-Mounted VR)

Within the scope of this study, as stated in Table 1.1, a pipeline for an EEG-NF experiment in virtual reality was designed and implemented. In total, 8 healthy subjects took part in this study, with ages ranging from 18 to 50 years. The aim of the experiment was to study EEG-NF learning outcome for both the individual UA and the estimated weighted-node-degree FC for the imaginary part of coherency

in the UA. Study participants were randomly distributed between the two groups, with demographic information described in Table 2.2.

Group	Alpha	FC
Age (years; mean $\pm$ SD)	34 ± 12	27.5 ± 5.5
Gender (F/M)	2 F ; 2 M	1 F ; 3 M

Table 2.2: Demographic characteristics of UA and FC groups (adapted from Berhanu et al. (2019)).

2.1.2 Equipment and Signal Acquisition

Bucho et al. (2019) and Berhanu et al. (2019) used similar protocols and equipment for signal acquisition. For both, the signal acquisition took place in Neurolab of the Evolutionary Systems and Biomedical Engineering Lab (LaSEEB), of the Institute for Systems and Robotics (ISR), at Instituto Superior Técnico (IST). The room provided the necessary sound and light conditions for an EEG acquisition and experiment. Additionally, the EEG amplifier LiveAmp (Brain Products GmbH, Gilching, Germany) was used for EEG acquisitions, the open source software OpenViBE (Renard et al., 2010) was utilized for the EEG-NF protocol and the display of the visual, auditory, and VR feedbacks was performed with the software Unity. The signals were sampled at 500 Hz from the ActiCap's system with 32ch active electrodes (Figure 2.1), with the ground electrode located at the forehead and the reference placed over the left mastoid (channel TP9).

In the study Bucho et al. (2019), the feedback was presented on a computer monitor from which the participants sat one meter away while using headphones.

In Berhanu et al. (2019) study, Oculus Rift Virtual reality Head-mounted display (HMD) were used for feedback presentation.

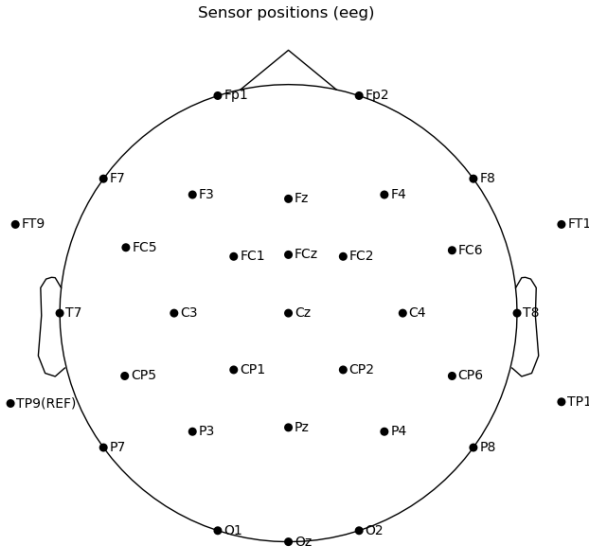

Figure 2.1: Electrode locations for the experimental protocols.

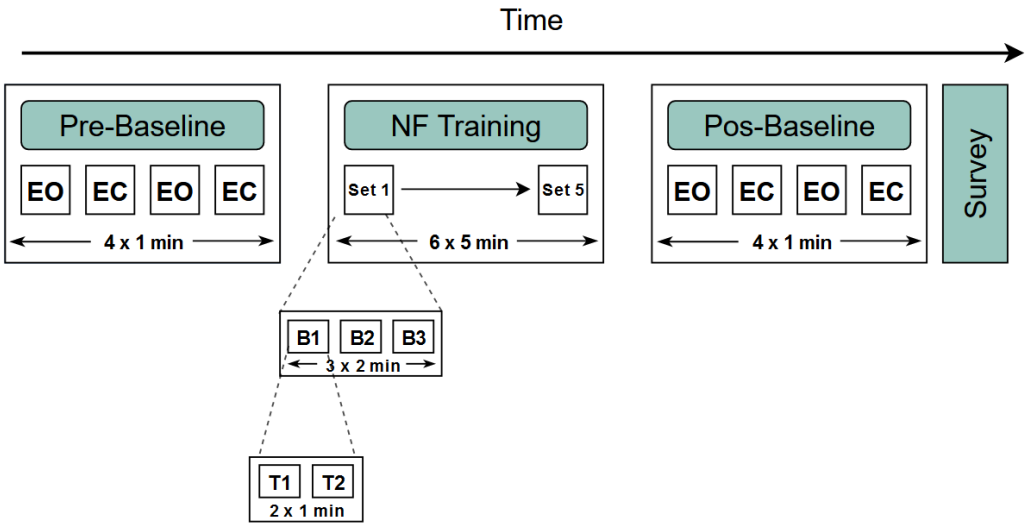

Figure 2.2: Temporal organization of a training session. Between trials there was a 10 seconds long pause, 15 seconds long between blocks and at least 15 seconds between sets (EO = Eyes Open ; EC = Eyes Closed).

2.1.3 Neurofeedback Training protocol

2.1.3.1 Session Design

The temporal organization of the training session is equivalent for both studies (Bucho et al., 2019; Berhanu et al., 2019) and it is illustrated in Figure 2.2. Each training session was split into four parts, it started with a 4-minute baseline acquisition, that consisted of alternate 1-minute periods with eyes open and eyes closed, during which the subjects did not perform any task. This pre-baseline period was fundamental, since it was used to calculate the minimum, maximum and threshold values of the feedback parameter and to determine the individual alpha band (IAB), which are both described in 2.1.3.2. After this, the NF training session began, which was divided into 5 sets, with each set containing 3 blocks, that individually, consisted of two 1-minute trials. Between each trial, there were pauses of 10 long or 15 seconds if it was between blocks. Between sets, the subjects were given a longer break of at least 15 seconds long. At the end of the EEG-NF, a pos-baseline period was acquired The total training time was around 37 minutes, per session. In the end, similarly to what was done before, a baseline was acquired and participants were asked to fill a mental state survey.

In both studies, each participant took part in 4 sessions that were conducted on consecutive days, at approximately the same time.

2.1.3.2 Feedback parameters

As it is referred in 2.1.3.1, the IAB is calculated during the pre-baseline period and it is measured again during the pos-baseline period of the last session to identify possible changes. This variable can be defined as the **maximum power value** in the EEG frequency spectrum between the lowest and higher transition frequencies, which are called lower transition frequency (LTF) and higher transition frequency (HTF). These boundary values are obtained through the intersection in a power spectral density (PSD) between eyes open (EO) and eyes closed (EC) signals (see figure 2.3)

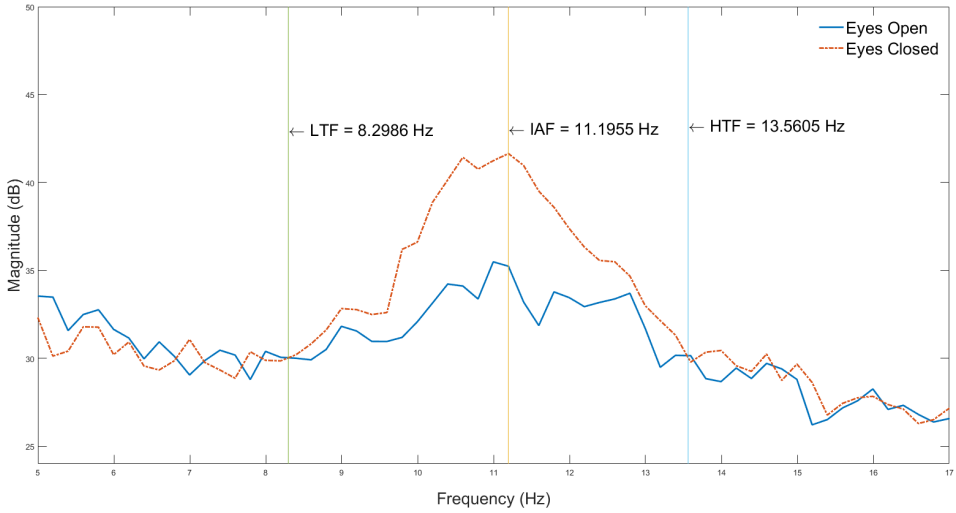

Figure 2.3: Illustration of the IAF, LTF, and HTF measurement (adapted from Bucho et al. (2019)).

In Bucho et al. (2019), the feature used for EEG-NF training was the relative amplitude of the upper alpha (RAUA), with the UA defined from the individual alpha peak to higher transition frequency (HTF). In Berhanu et al. (2019) the alpha group has the RAUA feature, whereas for the FC group the feature consisted on the weighted-node-degree for the imaginary part of coherency in the individual UA band.

This parameter was explicitly defined by Wan et al. (2014) and it was adapted for the UA band. The RAUA is determined by the sum of amplitude spectra in the UA band divided by the total sum of amplitude spectra when considering the signal from value 4 Hz to value 30 Hz, as defined by:

$$RAUA = \frac{\sum_{k=IAF/\Delta f}^{HTF/\Delta f} X(k)}{HTF - IAF} \Big/ \frac{\sum_{k=4/\Delta f}^{30/\Delta f} X(k)}{30-4} \tag{2.1}$$

where $X(k)$ is the amplitude spectrum at frequency k and Δf the frequency resolution. This parameter was, in both studies, computed in real-time within the OpenViBE software.

2.1.3.3 Display of feedback

Bucho et al. (2019) presents two different feedback displays: auditory and visual, however only the visual feedback is relevant for the present work. The visual feedback was presented in a computer monitor and it consisted on a sphere, that varied in size and was located over a horizon background (see Figure 2.4). The radius of the sphere changed linearly with the feedback parameter (RAUA) that was collected at the electrode location Cz. Likewise, the sphere's color altered between red and white, whether RAUA was below or above a predefined threshold, respectively. During the training session, participants were told to maintain the sphere as large as possible and keep it white for as long as possible.

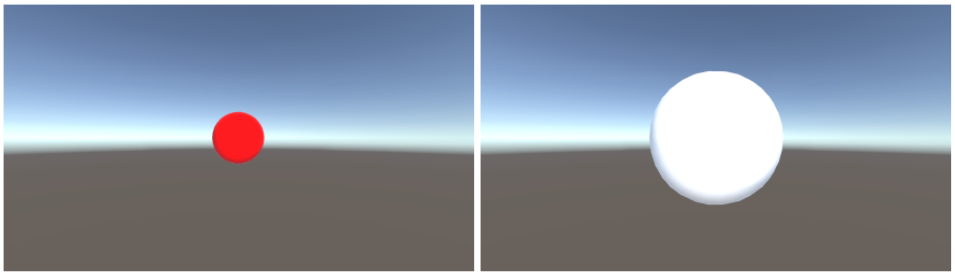

Figure 2.4: Visual feedback display from Bucho et al. (2019).

Berhanu et al. (2019), delivered the feedback using VR-HMD in the form of a spherical group of particles that changed size, rotated, and moved closer to the subject (see Figure 2.5), these movements corresponded to a distinct paradigm for every training session. For the first paradigm, participants were told to increase the size of a rotating sphere that was standing in front of a planet. The sphere equaled the size of the planet when $\boxed{RAUA = 1}$. For the Rotation Speed paradigm, the sphere maintained a fixed size, and subjects were instructed to increase its velocity. Finally, in the last paradigm, the sphere started further back in the virtual world and the participants aimed to get the particle as close as possible and whenever they reached the center of the particle system, they were returned to the initial position.

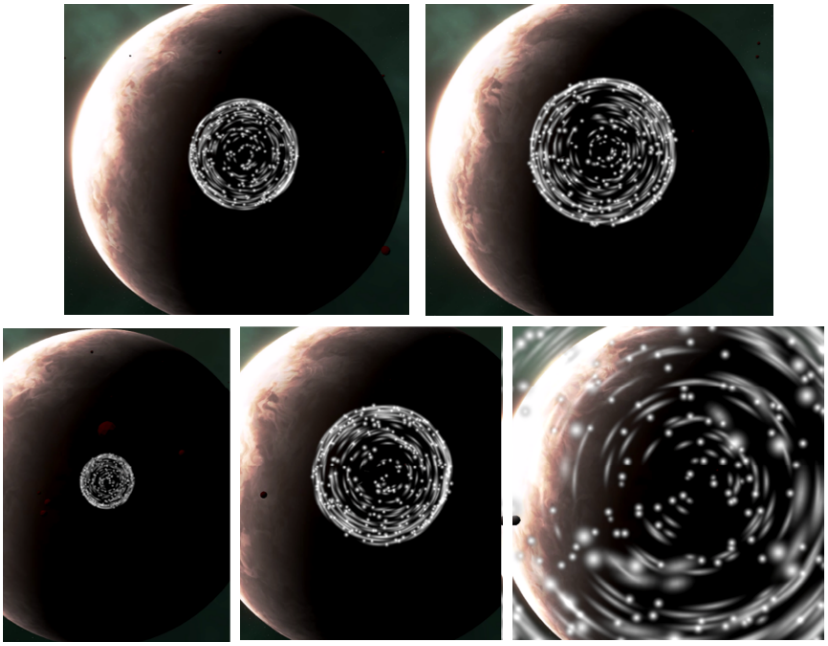

Figure 2.5: Feedback display in Berhanu et al. (2019). Top: the sphere of particles changed in size, Bottom: sphere gets closer to the participant.

2.2 Data Analysis

Throughout this section, all the processing, as well as analysis steps, will be described in detail, along with screenshots of the code created for this purpose. All the steps of data processing, which go from extracting the raw data to data cleaning, along with analysis, were executed in a developed *Python* script, with the help of **MNE-Python** library (Gramfort et al., 2013) . This script was built so that it could take any EEG data file (in .gdf format) and run all of the stages involved in the data processing.

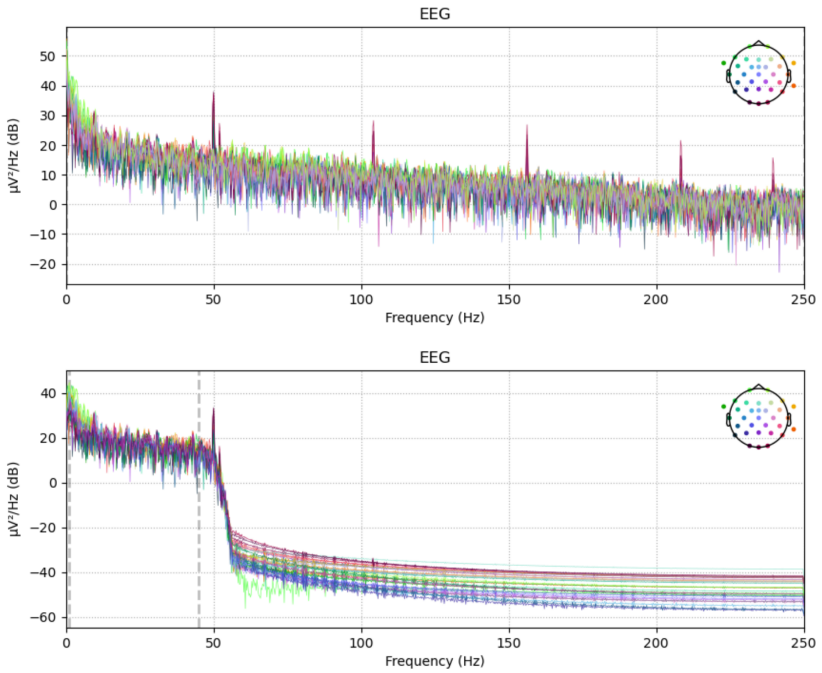

Figure 2.6: Frequency spectrum across channels before and after filtering, of a selected data file. Each channel is represented by a different color, as shown in the topographic image.

2.2.1 Data cleaning

In Bucho et al. (2019), it was pointed out that explicit and direct artifact removal was not performed, since it aimed at being closer to the true experimental conditions, however, if it had been it could have had a big impact on the study conclusions. With this in mind, all the data extracted from previous work was in its raw form. In other words, data processing was performed from scratch which aims to improve results.

Data Loading was the first step, each file was uploaded and read individually. After this, the EEG montage was collected in order to associate each channel with its correct name, additionally, an image that showed each channel location on the scalp was created (Figure 2.1).

Filtering is a crucial part of any data cleaning procedure, the idea is to remove any electrical noise or artifacts from the signal. Brainwaves occur at various frequencies that can go up to 200 Hz, however in an EEG spectrum low frequencies are the most identifiable. Usually, EEG filters are made to reject both very low and high frequency activities, with high-pass filters being applied at 0.1 or 1 Hz and low-pass filters above 40 or 50 HZ. In the current work, a band-pass filter of 1-45 Hz was applied to all data files. Figure 2.6 shows the frequency spectrum that was plotted before and after applying the band-pass filter.

After filtering, the data was **re-referenced**. In this study, the reference electrode was channel TP9, which means that all channels were expressed as the difference in electrical potential to this channel. This referencing is usually performed during recording, however offline referencing can also be achieved, which is called re-referencing. For this study, a re-reference to average was applied, which means that the voltage at each channel is now with respect to the average of all channels.

Artifact removal

The importance of artifact removal was already covered in section 1.4.2, therefore this section outlines how the artifacts present in this data were handled. At this point, the data was already filtered and while this is a key step of preprocessing it is not sufficient to remove artifacts that arise from eye movement, muscle activity or other external noise sources (Klug and Gramann, 2021). As described in 1.4.2, ICA is a very popular method to address this type of artifacts since it allows for isolation of independent components in the signal and consequently their removal (Chaumon et al., 2015). Considering this, ICA was applied to all data files with the goal to find and remove any non-cerebral activity from the signal. This was accomplished by using the **MNE-Python** function mne.ICA(), which requires as input not only the number of components (n_components) that is passed to the algorithm during fitting (ranged between 1 and the number of existing channels) but also, the fit method. In order to maintain coherence throughout the analysis the same number was applied to all files when running ICA: **n_components = 19**, additionally the chosen fit method was FastICA.

This algorithm was applied to the filtered raw data and after this a series of images were generated. The next step was to reject the components thought to be artifactual, which is accomplished by a careful examination of this images. Figure 2.7 and 2.8 show the time series and the scalp topography of each individual component, after visualizing them and learn to interpret them it is relatively simple to find what components should be removed. The process of removing components from the signal was kept as conservative as possible, i.e. components were rejected only when there was a very clear noise pattern (e.g. ocular movements or heart beats). Then, the chosen components were subtracted from the original raw data and a new cleaned data file was created.

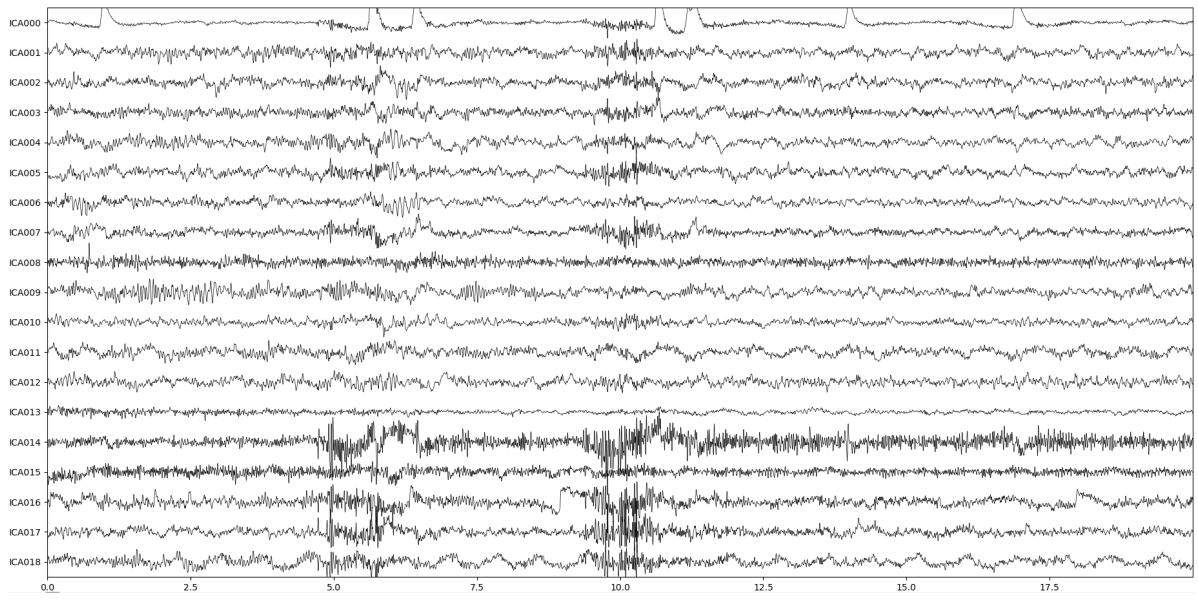

Figure 2.7: Time series of each component. Plot generated from a data file from Bucho et al. (2019) study (Sub14, Session 3)

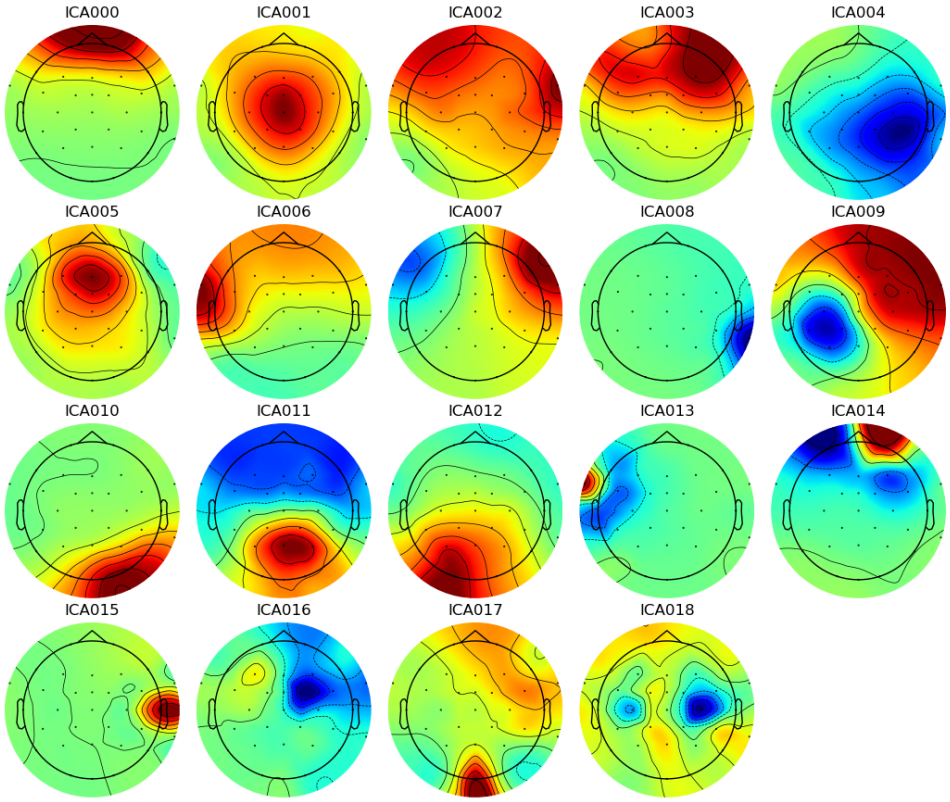

Figure 2.8: Scalp topography of each component. Plot generated from a data file from Bucho et al. (2019) study (Sub14, Session 3).

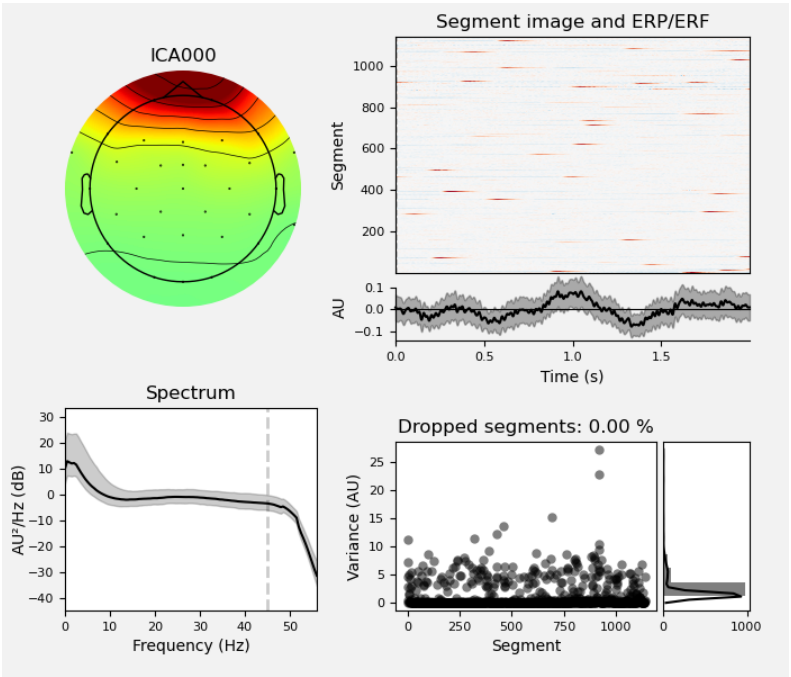

Figure 2.9: Individual component properties, which include - topographic epochs image, ERP/ERF, power spectrum, and epoch variance (Sub14, Session 3).

Epoching

Epoching can be defined as extracting specific time-windows from a continuous signal. From epoching we obtain epochs, which are chunks of data with equal duration, that usually represent experimental events, like the response in a certain time window after the onset of the stimulus. Because of this, epoching EEG data is a recurrent technique used in data analysis. As it was described in section 2.1.3.1, the experimental sessions consisted of 5 sets, each set containing 6 trials in them (two trials per block). After filtering, the data was epoched in two different ways: **(1)** Epoching by sets - which resulted in 5 epochs and **(2)** Epoching by trials, which resulted in 30 epochs.

Inside an epochs object, the data comes in the form of an array, that contains the shape - [epoch, channel, time]. For both studies, each epoch was created based on the markers present on the EEG data and using the **MNE-Python** function mne.Epochs(). The data was epoched before and after artifact removal and the epochs were stored in .fif format files and then used for analysis.

With the epochs generated from the new data it was possible to assess the cleaning effect through topomaps of the power spectral density across epochs in Figure 2.10.

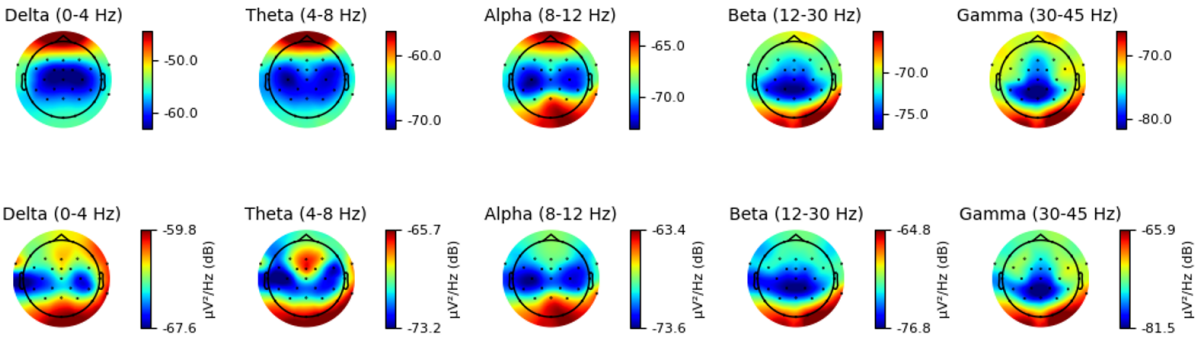

Figure 2.10: Topomaps of the power spectral density across epochs before (Top) and after (Bottom) applying ICA.

2.2.2 Data processing

After cleaning and epoching the data, the next step consisted in performing a time-frequency analysis, which provides an overview of the spectral influence of the NF training. To achieve this, it was necessary to decompose the signal into frequency bands of interest, which were adapted to each subject's IAB. Considering that LTF = IAF - 2 Hz and HTF = IAF + 2 Hz:

- Theta: 4Hz to LTF
- Lower Alpha (LA): LTF to IAF
- Upper Alpha (UA): IAF to HTF
- Beta: HTF to 30Hz

Each subject data file was uploaded and the previously created trial epochs were put together to form all five sets. By using trial epochs instead of set epochs, it was guaranteed that each set only contained data that corresponded to training periods (all pauses were removed).

Afterward, in order to calculate the relative amplitude (RA) of each band, equation 2.1 was adapted for all frequencies of interest. For this to work, a python function was created and implemented.

This function started by computing an estimate of the PSD using the Multitaper method, which unlike Welch's method does not require a time window to be selected, and proceeded to calculate the selected band power by averaging the PSD in the selected set and channel and normalizing within the frequency range (see figure B.5). Each band RA was computed for all training data, for every subject and stored in a .csv file.

2.2.3 Assessing Training Effect

Calculating the power of each band assesses the questions: 1) How well did each participant was able to up-regulate the target feature, with respect to each feedback modality. 2) Is this regulation limited to the target frequency band. Then, to assess the training effect and similarly to what was done in previous studies (Berhanu et al., 2019; Bucho et al., 2019), three **learning indices** that have in consideration the variation of a feedback parameter were defined:

- **IntraS:** calculates, for each session (i), the slope (m_i) obtained by linear regression of the learning parameter along all 5 sets in that session, while averaging across sessions:

$$IntraS = \frac{\sum_{i=1}^{nsessions} m_i}{nsessions} \tag{2.2}$$

- **IntraA1:** calculates, for each session (i), the mean difference between the mean of the first set and the mean of each set (j), while averaging across sessions:

$$IntraA1 = \frac{\sum_{i=1}^{nsessions} \sum_{j=2}^{nsets} (\bar{set}_j - \bar{set}_1)_i}{nsessions(nsets-1)} \tag{2.3}$$

- **IntraA2:** calculates, for each session (i), the difference between the mean of the last and the first sets, always relative to the mean of the first, while averaging across sessions:

$$IntraA2 = \frac{\sum_{i=1}^{nsessions} (\frac{\bar{set}_5 - \bar{set}_1}{\bar{set}_1})_i}{nsessions} \tag{2.4}$$

- **InterS**: calculates the slope (m) of the evolution along the four sessions:

$$InterS = m \tag{2.5}$$

- **InterA1**: calculates the difference between the mean of the last two sessions and the mean of the first two, while averaging across the first two sessions:

$$InterA1 = \frac{(\bar{S}_4 + \bar{S}_5) - (\bar{S}_1 + \bar{S}_2)}{\bar{S}_1 + \bar{S}_2} \tag{2.6}$$

- **InterA2**: calculates the difference between the means of the last two sets of the last session and the means of the first two sets of the first session, relative to the first two:

$$InterA2 = \frac{(\bar{set}_4 + \bar{set}_5)_{S4} - (\bar{set}_1 + \bar{set}_2)_{S1}}{(\bar{set}_1 + \bar{set}_2)_{S1}} \tag{2.7}$$

Identification of non-learners

As it is mentioned in 1.4.2 there is no clear definition that separates learners from non-learners. However, practically all studies involving NF, take into account the fact that there are some participants who are not capable of self regulating brain activity. As revealed by Reiner et al. (2018), there is an evident necessity to differentiate between learners and non-learners, since this is the only way to quantify the effect of the analysis of the learning process on the effectiveness of neurofeedback. Because of this and based on what was taken into account by Bucho et al. (2019), **IntraS** was chosen as a criteria for identifying non-learners. Meaning, whenever this index presents a negative value the participant is classified as non-learner. Evidently, from this section on, learners will be subjects whose slope (m_i) obtained by linear regression of the evolution of the feedback feature (RAUA) is positive.

2.2.4 Statistical Analysis

Visual inspection is not sufficient to fully understand and analyze the results obtained. A statistical approach can provide valuable and quantifiable information about the data. For this reason, a statistical analysis in order to evaluate differences between modalities was performed.

The first step in any statistical analysis is to assess the sample size. Parametric tests can analyze only continuous data and require sample data big enough to approximate normality. Whenever the sample size requirements are not met and there is no information on the normality of the data, non-parametric tests must be used. This work consisted of two different data samples, both were small (4 subjects in the Immersive-VR group and 8 subjects in the Screen-Based) and did not grant normality. Because of this, a non-parametric Wilcoxon Signed-Rank test was used.

The Wilcoxon Signed-Rank test is the non-parametric equivalent to the Student's t-test. It examines the differences between one pair (a, b) of data that are non-normally distributed. The null hypothesis is that the distribution of $a - b$ has a median equal to zero, i.e. there are no significant differences between sample a and b. The test results must be compared to our significance value: when the p-value is less than or equal to your significance level the null hypothesis is rejected.

In the present work, the Wilcoxon Signed-Rank test was used to compare sets within the same session and evaluate differences between learning indexes.

Chapter 3

Results

In this chapter, the results are presented and divided into four sections. First, the target feature is explored for both modalities by means of the individual evolution of the RAUA within and across modalities and also by a boxplot representation of this feature's distribution (3.1). Then in the next section, the remaining EEG bands are explored mostly by showing the group median evolution (3.2). The results concerning the learning indexes are shown in the third section through tables and boxplot distributions (3.3). To simplify the results, the subjects from each group had a letter and number assigned to it: Screen-Based group - subjects V1 to V8 and Immersive-VR group - subjects A1 to A4.

3.1 Target Feature Evolution within sessions

In figures 3.1 3.2, it is presented the evolution of the RAUA at Cz along sets and for all sessions, for each participant in Screen-Based and Immersive-VR modalities, respectively. Each blue point represents the mean of each set in that session.

Regarding the Screen-Based modality, we observe several different distributions for each subject: some show a direct growing trend, while others show the opposite (figure 3.1). When analyzing the data per individual subject, subject V1 shows a positive trend within each session, however the RAUA of the last session is substantially lower than the first. Similarly to subject V1, subject V4 shows a growing trend over all sessions. Subjects V2 and V5 show a considerable decrease in the target feature over time, as well as subject V8 who of all subjects presents the higher values of RAUA and shows a more sharp decrease in the data. V3 and V6 values remain roughly constant, although the latter shows a slight increase only during the first two sessions. Finally, subject V7 shows in the first two sessions a gentle decrease, however that changes in sessions 3 and 4, show an evident growth. For the Immersive-VR group, it is clear that there is an increasing tendency for all subjects, with the exception of subject A4. In general, the data from the Screen-Based modality group appears to have more fluctuations within sessions and consequently, are harder to take visual conclusions.

Figure 3.3 provides an overall vision of the target feature evolution throughout different sessions, for all participants. To achieve this, the mean RAUA for each session was calculated for subjects in the Screen-Based modality (left) and the Immersive-VR (right) groups. In the first image, subjects V6 and V8 show a noticeable drop from the first session to the second and third. On the other hand, subject V4 reveals a slight increase starting in session 2. In the Immersive-VR group, it seems that only subject A3 had a positive increment through sessions. Besides that, subject A1 shows a very sharp drop when compared with subjects A2 and A4.

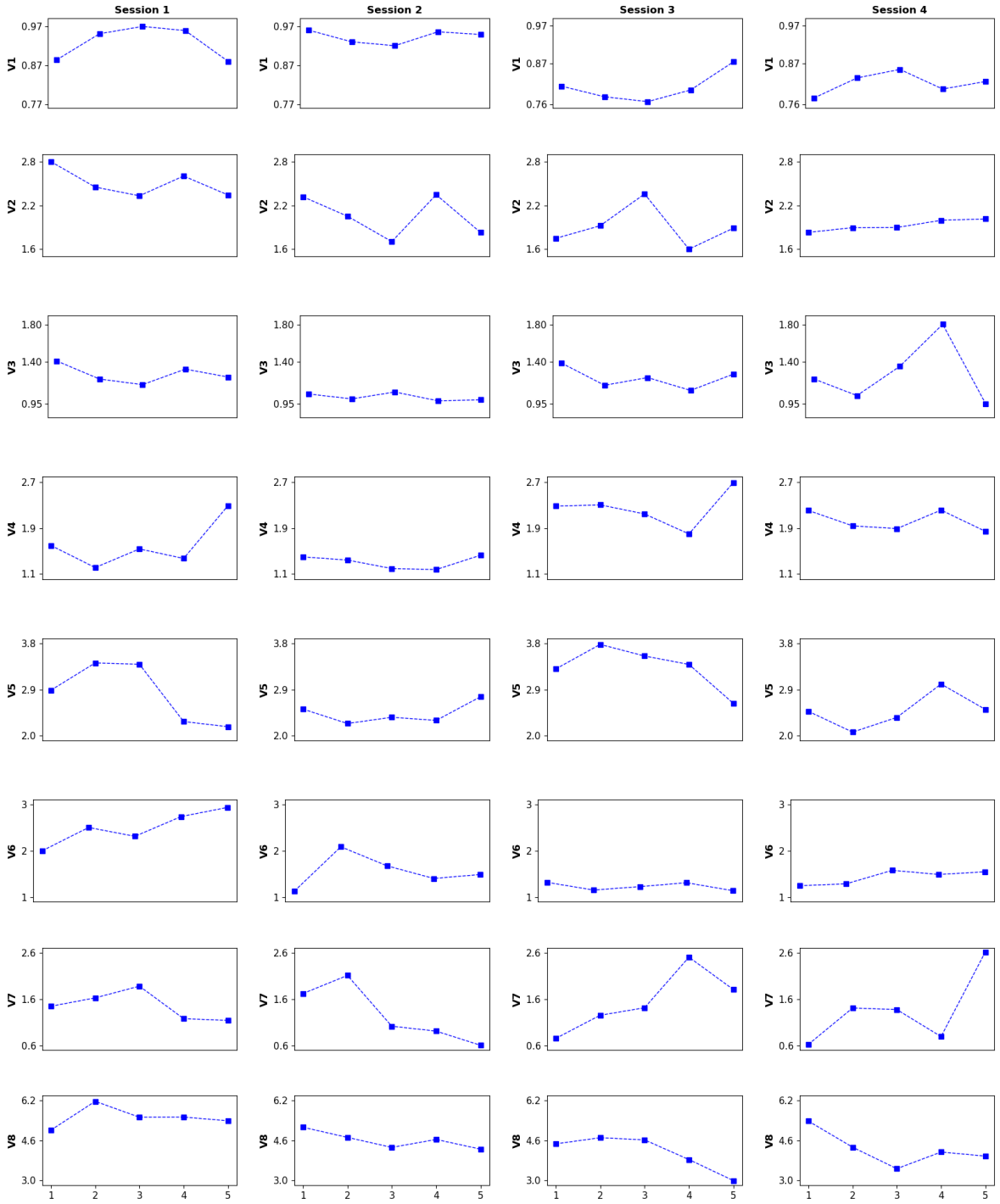

Figure 3.1: Evolution of RAUA at Cz, for each participant (V1 to V8), along sets (1 to 5) in all four sessions for the Screen-Based modality. Each point represents the mean RAUA of that set.

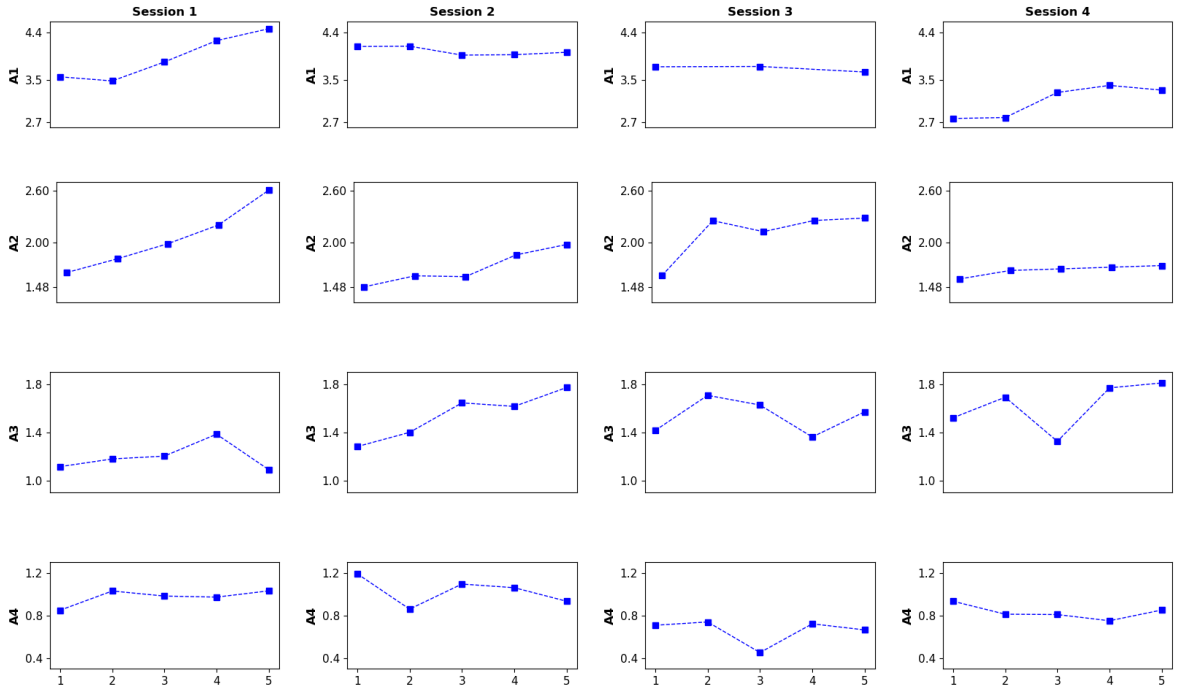

Figure 3.2: Evolution of RAUA at Cz, for each participant (A1 to A4), along sets (1 to 5) in all four sessions for the Immersive-VR modality. Each point represents the mean RAUA of that set.

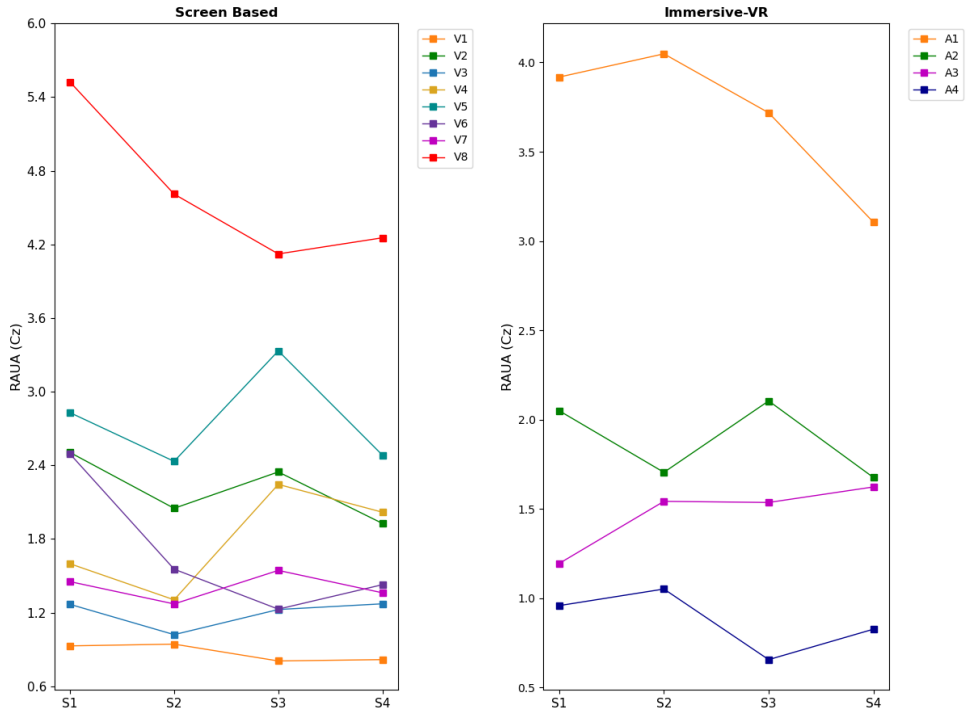

Figure 3.3: Evolution of RAUA at Cz for each subject, across sessions for both modalities. Each point represents the mean RAUA for each session (S).

Figure 3.4 illustrates, for each session, the distribution of values of the target feature at Cz, across all participants for both modality groups. In the first session, the medians of the Screen-Based modality are consistently higher than the Immersive-VR. However, only the median of the Immersive-VR group increases during the session as well as the maximum value. Looking at the range of the boxes it is

possible to tell that the distribution of the values from the first session to the second differs significantly. Moreover, while the median for the Screen-Based group stays approximately constant, the Immersive-VR group median continues to increase throughout the session and this time presents higher values. However, Session 3 does not show any particular trend regarding the median values, nevertheless, the maximum values seem to decrease for both modalities. During the last session, the median behavior does not look very different from the previous session, however, it is possible to say that the medians for both modalities present very similar values.

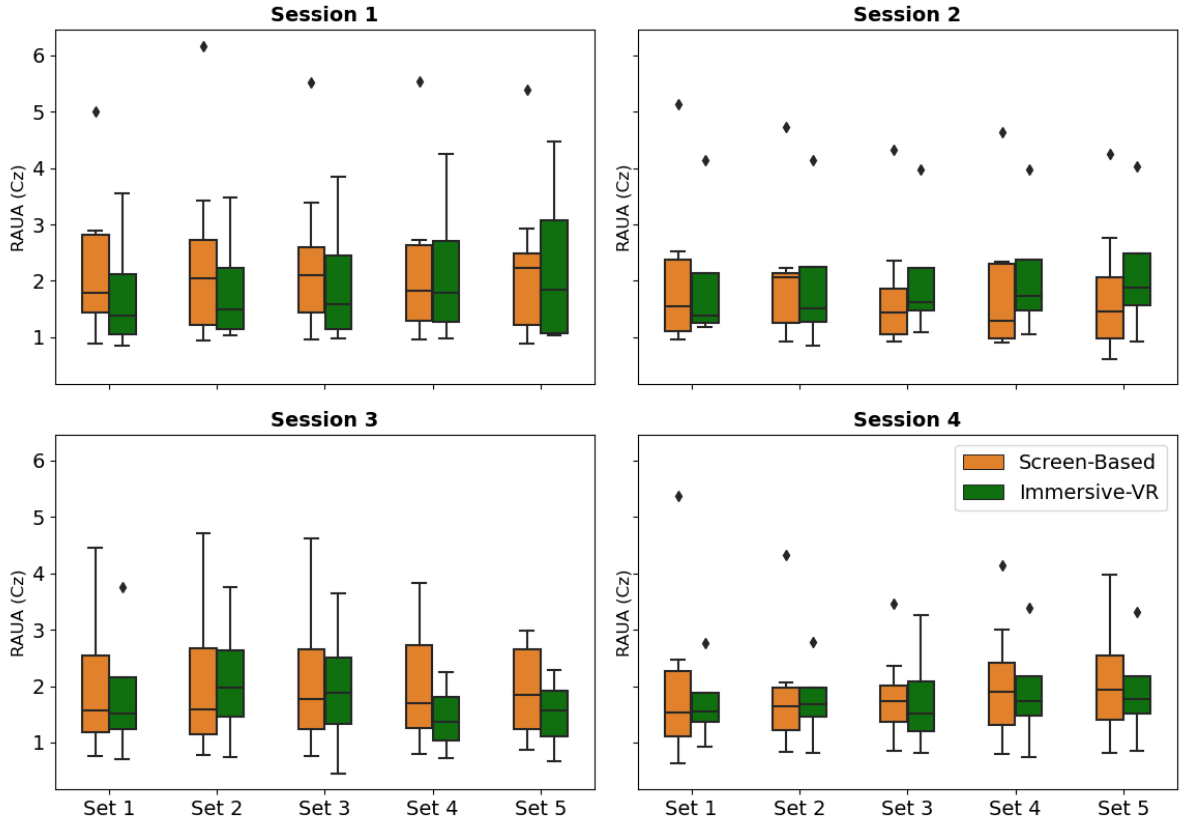

Figure 3.4: Boxplot representation of the distribution of RAUA at Cz along sets and sessions across all participants from both modalities.

The distribution of values of the target frequency band across learners is represented in Figure 3.5 and further analyzed on tables 3.1, 3.2, 3.3 and 3.4. When considering only the learners we see that there is a major shift in the data from the Immersive-VR group, for instance, the median for the last session of each set increased by 41% (S1), 84% (S2), 19% (S3) and 29% (S4). Furthermore, the descriptive tables confirm that in the Immersive-VR group there is a positive growth in the relative amplitude of upper alpha in all sessions. Additionally, one thing to notice is that IQR values are much lower in the Screen-Based modality group, i.e. the data for this group has a smaller spread and consequent less variability. However, this variability seems to decrease throughout sessions for the Immersive-VR group, by 50 %.

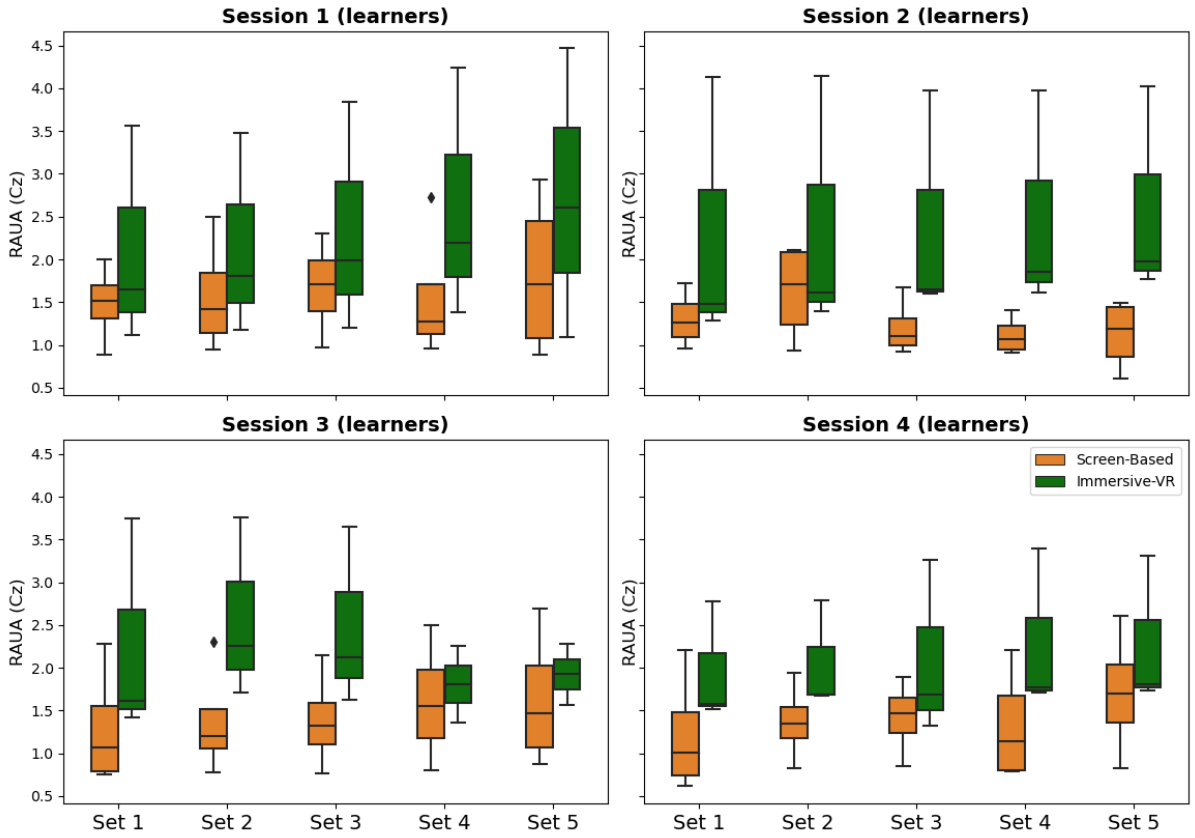

Figure 3.5: Boxplot representation of the distribution of RAUA at Cz along sets and sessions across participants classified as learners from both modalities.

Session 1

	Screen-Based (learners)				Immersive-VR (learners)			
	Min	**Max**	**IQR**	**Median**	**Min**	**Max**	**IQR**	**Median**
Set 1	0.884	1.999	0.391	1.519	1.114	3.559	1.222	1.648
Set 2	0.951	2.498	0.699	1.416	1.1783	3.481	1.151	1.808
Set 3	0.970	2.308	0.591	1.705	1.202	3.839	1.318	1.983
Set 4	0.959	2.733	0.586	1.272	1.383	4.243	1.430	2.200
Set 5	0.880	2.929	1.374	1.714	1.090	4.468	1.689	2.606

Table 3.1: Descriptive statistical values of the distribution of RAUA for learners in Session 1.

Session 2

	Screen-Based (learners)				Immersive-VR (learners)			
	Min	**Max**	**IQR**	**Median**	**Min**	**Max**	**IQR**	**Median**
Set 1	0.960	1.719	0.388	1.261	1.281	4.132	1.426	1.482
Set 2	0.930	2.107	0.851	1.711	1.399	4.139	1.370	1.610
Set 3	0.920	1.671	0.320	1.102	1.600	3.968	1.185	1.643
Set 4	0.909	1.401	0.286	1.064	1.614	3.977	1.181	1.853
Set 5	0.604	1.485	0.576	1.186	1.770	4.023	1.126	1.974

Table 3.2: Descriptive statistical values of the distribution of RAUA for learners in Session 2.

Session 3

	Screen-Based (learners)				Immersive-VR (learners)			
	Min	**Max**	**IQR**	**Median**	**Min**	**Max**	**IQR**	**Median**
Set 1	0.753	2.283	0.763	1.062	1.415	3.749	1.167	1.614
Set 2	0.780	2.304	0.455	1.201	1.706	3.755	1.024	2.250
Set 3	0.767	2.146	0.484	1.318	1.626	3.650	1.012	2.123
Set 4	0.798	2.500	0.790	1.553	1.361	3.650	1.144	2.253
Set 5	0.874	2.694	0.955	1.471	1.569	3.650	1.040	2.280

Table 3.3: Descriptive statistical values of the distribution of RAUA for learners in Session 3.

Session 4

	Screen-Based (learners)				Immersive-VR (learners)			
	Min	**Max**	**IQR**	**Median**	**Min**	**Max**	**IQR**	**Median**
Set 1	0.624	2.207	0.747	1.011	1.520	2.771	0.625	1.574
Set 2	0.831	1.938	0.368	1.346	1.673	2.788	0.558	1.690
Set 3	0.854	1.889	0.411	1.474	1.324	3.263	0.969	1.690
Set 4	0.792	2.212	0.870	1.144	1.710	3.396	0.843	1.768
Set 5	0.821	2.611	0.672	1.693	1.729	3.307	0.789	1.809

Table 3.4: Descriptive statistical values of the distribution of RAUA for learners in Session 4.

Figure 3.6 gives us a clear view of the disparities between modality groups and also between learners and non-learners by showing the medians' evolution over the four training sessions. There is a noticeable positive growth tendency, which is more evident for the Immersive-VR modality. However, despite the fact that some sessions do not show a constant upwards tendency, the median of the last set is always higher than the first set of that session, for both modalities with all subjects included. Additionally, when comparing the medians' first value in each session, there is an evident increase, exclusively in the Immersive-VR group. Removing the data from the non-learners had more impact in the Screen-Based modality data. In fact, this group's medians significantly decreased in all sessions. On the other hand, the data from the Immersive-VR group presents higher values and even more evident growth in each session when only considering the learners.

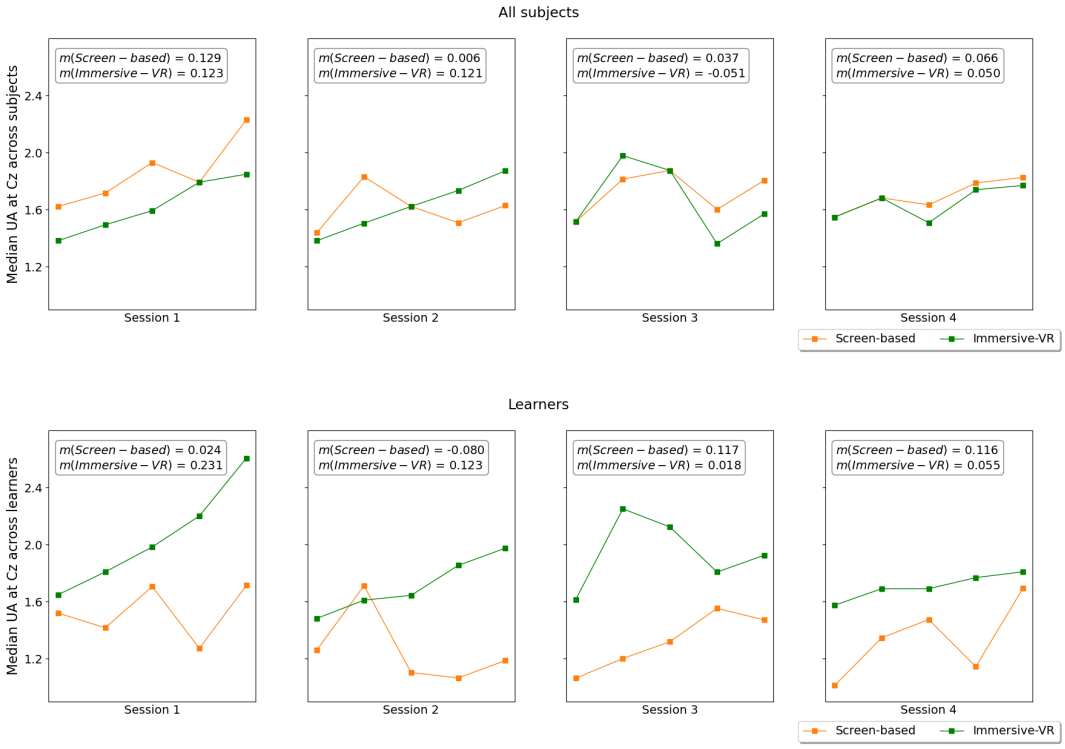

Figure 3.6: Evolution of the group median relative amplitude of upper alpha at Cz with and without non-learners.

Tables 3.5 and 3.9 show the results of the statistical tests performed on the data, which are described in section 2.2.4.

In the table 3.5 the results from the Wilcoxon Signed-Rank test that compares the RAUA of the first and last set of each session for all participants are shown. There is no significant p-value, which means that the null hypothesis must not be rejected, considering that the null hypothesis is that there are no significant differences between the pair of sample distributions.

	Screen-Based	Immersive-VR
Session 1	1.0	0.25
Session 2	0.383	0.625
Session 3	0.844	0.25
Session 4	0.945	0.25

Table 3.5: p-values resulted from the Wilcoxon Signed-Rank test that compares the RAUA of the first and last set of each session for all participants.

3.2 EEG bands evolution within sessions

In order to assess the training effect aside from the target feature, the relative amplitude of the remaining frequency bands (Lower-Alpha, Beta and Theta) was calculated at the target electrode Cz. Figures 3.8, 3.9 and 3.7, similarly to figure 3.6, show the median relative amplitude for each band during the four training sessions. Note that, as it is described in section 2.2.2 the frequency bands were adapted to each subject, based on their IAB.

Starting with the Theta band, we see that, contrary to what happens with the UA band, the relative amplitude appears to have, for the most part, a negative evolution (slopes below zero) for both modali-

ties. However, this is not true for the Screen-Based modality during session 3. Additionally, we see that removing the data from the non-learners did not have such an evident impact as it did in figure 3.6. Nonetheless, it is noticeable that the data drifts slightly apart, with the exception of session 1. Regarding the evolution of the Lower Alpha band, it becomes apparent that the modality groups present an opposite behavior to one another, in fact while the Immersive-VR modality group appears to have a generally positive trend, the Screen-Based shows a decrease in the relative amplitude for all sessions. Moreover, the removal of the non-learners had a significantly greater impact on the Immersive-VR group. Lastly, the results for the Beta band show that the values for the learner of the Screen-Based group are at least 30% higher than the Immersive-VR group, for all sessions, which does not happen with any other frequency band. Furthermore, we see that during the first two sessions, while the Screen-based modality learners show an increase in amplitude by 16 %, the Immersive-VR shows a decrease by 12 %, and they switch behaviors during the last two sessions.

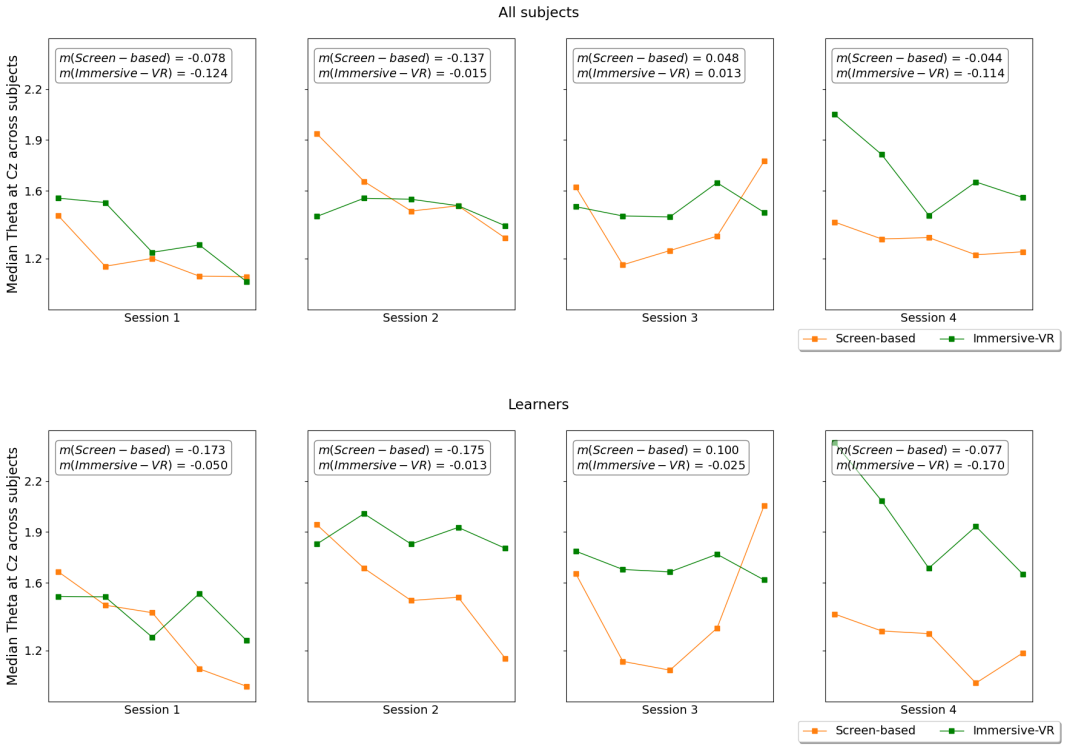

Figure 3.7: Evolution of the group median relative amplitude of **theta** at Cz with and without non-learners.

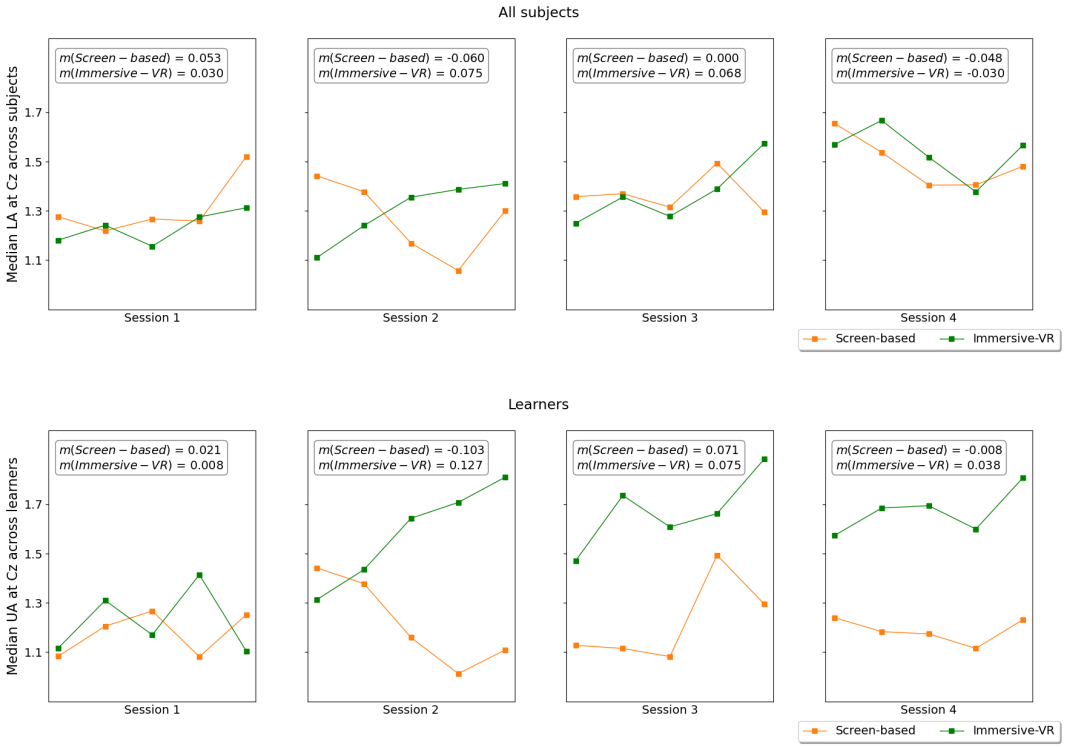

Figure 3.8: Evolution of the group median relative amplitude of **lower alpha** at Cz with and without non-learners.

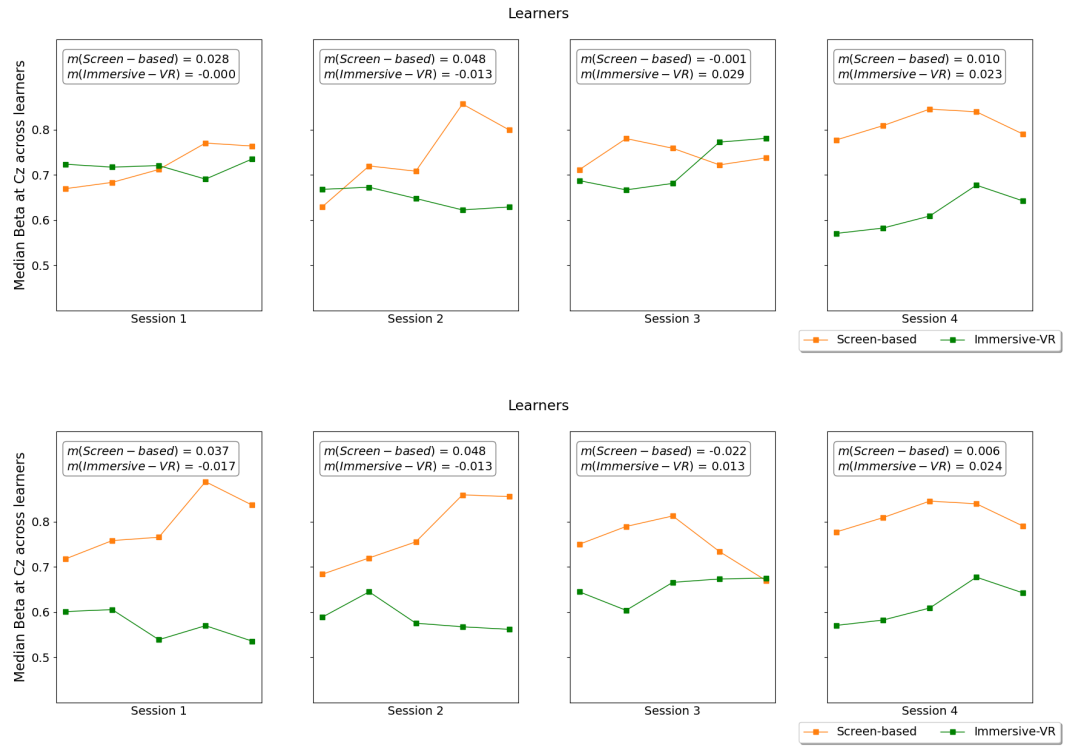

Figure 3.9: Evolution of the group median relative amplitude of **beta** at Cz with and without non-learners.

Tables 3.6 and 3.7, show the results of the Spearman correlation computed between the set number and the RAUA of that set for each session, for all participants and learners only, respectively. Significant p values show a relevant correlation between set number and the RAUA. The table only shows significant values for the Theta band, in the Screen-Based modality group. By looking at those results,

one can assume that the RA of Theta has a negative correlation with set number, i.e. throughout Session 1 and Session 2 learners of the Screen-Based group evidently decrease their band power in the Theta frequency range. Additionally, we see that for target frequency band there is no significant correlation with the set number.

Screen-Based								
	Theta		**LA**		**UA**		**Beta**	
	correlation	p value	correlation	p value	correlation	p value	correlation	p value
Session 1	-0.302	0.058	0.021	0.895	-0.047	0.771	0.142	0.380
Session 2	-0.349	0.027	-0.119	0.463	-0.132	0.418	0.260	0.105
Session 3	0.091	0.572	0.029	0.859	0.037	0.822	0.031	0.851
Session 4	-0.150	0.355	-0.039	0.807	0.135	0.407	0.018	0.910

Immersive-VR								
	Theta		**LA**		**UA**		**Beta**	
	correlation	p value	correlation	p value	correlation	p value	correlation	p value
Session 1	-0.466	0.038	0.012	0.959	0.153	0.519	0.024	0.918
Session 2	-0.147	0.536	0.276	0.239	0.165	0.485	-0.153	0.519
Session 3	-0.132	0.579	-0.082	0.728	-0.009	0.969	0.088	0.709
Session 4	-0.343	0.138	0.110	0.643	0.233	0.323	0.245	0.297

Table 3.6: Spearman correlation between set number and mean RA of the studied frequency bands, for all participants.

Screen-Based (learners)								
	Theta		**LA**		**UA**		**Beta**	
	correlation	p value	correlation	p value	correlation	p value	correlation	p value
Session 1	-0.472	0.036	-0.061	0.797	0.031	0.898	0.276	0.239
Session 2	-0.497	0.026	-0.208	0.378	-0.264	0.261	0.374	0.104
Session 3	0.135	0.571	0.086	0.719	0.215	0.364	-0.080	0.738
Session 4	-0.178	0.453	0.049	0.837	0.264	0.261	-0.006	0.980

Immersive-VR (learners)								
	Theta		**LA**		**UA**		**Beta**	
	correlation	p value	correlation	p value	correlation	p value	correlation	p value
Session 1	-0.349	0.202	0.033	0.908	0.207	0.458	-0.153	0.587
Session 2	-0.185	0.508	0.436	0.104	0.338	0.218	-0.349	0.202
Session 3	-0.169	0.546	0.027	0.923	0.016	0.954	0.082	0.772
Session 4	-0.415	0.124	0.218	0.435	0.458	0.086	0.426	0.114

Table 3.7: Spearman correlation between set number and mean RA of the studied frequency bands, for participants classified as learners.

3.3 Learning Indexes

As previously mentioned in section 2.2.3, the index **IntraS** was calculated and then used for identifying and then classifying subjects as either learners or non-learners. Table 3.8 shows the index values calculated for the target frequency band for all subjects of both groups, with subjects considered non-learners highlighted in red. In the Screen-Based group, four subjects were classified as non-learners (V2, V3, V5 and V8) and one subject (V4) in the Immersive-VR group.

	Screen-Based								Immersive-VR			
	V1	**V2**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**	**A1**	**A2**	**A3**	**A4**
IntraS	0.005	-0.025	-0.011	0.033	-0.069	0.068	0.056	-0.213	0.085	0.130	0.049	-0.034
IntraA1	0.020	-0.097	-0.087	-0.098	-0.032	0.317	0.340	-0.456	0.185	0.379	0.176	-0.06
IntraA2	0.030	-0.048	-0.121	0.118	-0.084	0.646	0.931	-0.172	0.099	0.356	0.165	-0.038
InterS	-0.047	-0.188	0.021	0.220	-0.014	-0.352	≈ 0	-0.430	-0.277	-0.072	0.128	-0.079
InterA1	-0.133	-0.159	0.091	0.469	0.105	-0.344	0.067	-0.173	-0.143	≈ 0	0.154	-0.256
InterA2	-0.116	-0.236	0.050	0.447	-0.125	-0.326	0.109	-0.275	-0.048	-0.444	-0.169	-0.147

Table 3.8: Indexes for the UA band at location Cz for all subjects of both groups.

In figure 3.10, the distribution of each learning index across all participants is represented for the four frequency bands. As it is mentioned in previous sections, if these indexes have positive values, it means that there is a general increase in the measured features. We see that for the Immersive-VR group, the median of the learning indexes is above zero for all bands with the exception of theta. We also see that the target frequency band (UA), not only has the highest medians ($Intr\bar{a}A1$ = 0.18; $Intr\bar{a}A2$= 0.13; $Intr\bar{a}S$= 0.07;), but also shows more data variance. What happens in the other modality group is quite different. The indexes relative to the UA band show median values very close to zero or even below (Intra A1). Theta is the band that shows a higher variance, and Beta has both the highest median and lowest variance. Figure 3.11, shows the distribution of these same exact learning indexes but restricted to the learners. A clear effect is noticed, more evidently in the Screen-Based modality group where the median of all indexes goes from being very close to zero to growing to a positive value. This is expected, since non-learners were classified based on the signal on the index IntraS and this group presented a considerably higher number of non-learners.

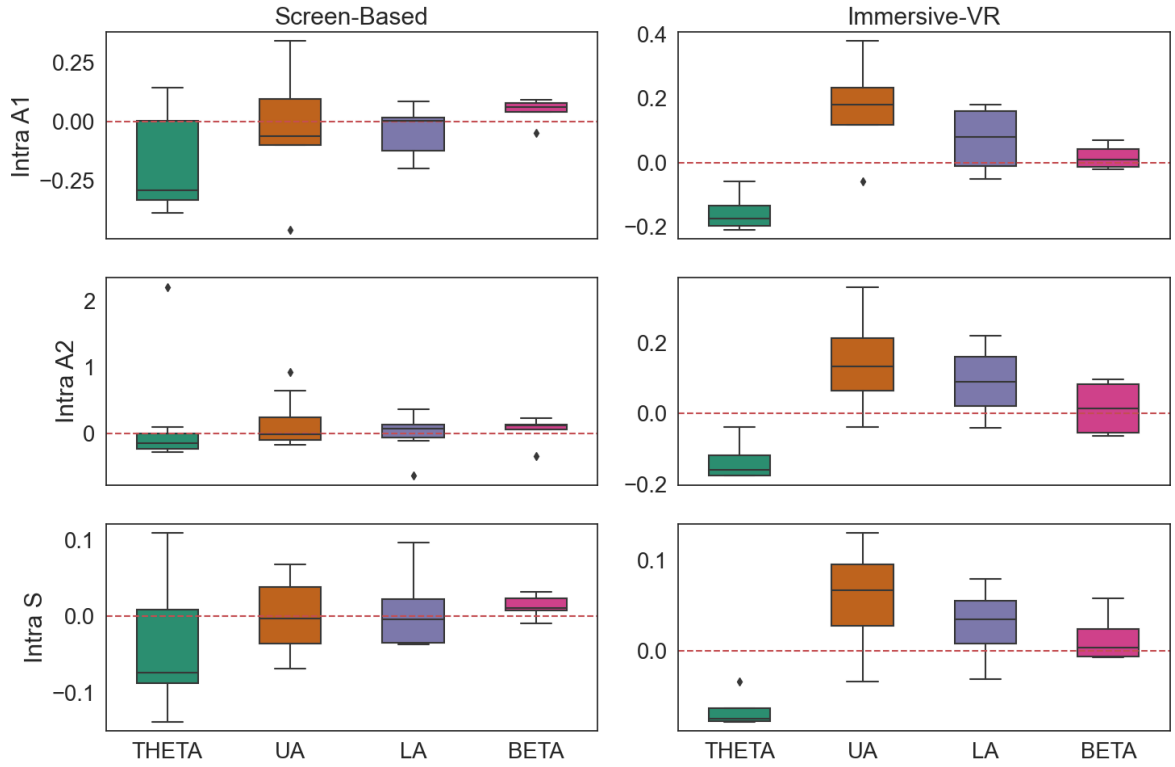

Figure 3.10: Distribution across all participants of learning indexes corresponding to within session evolution, at Cz for all bands, for each modality.

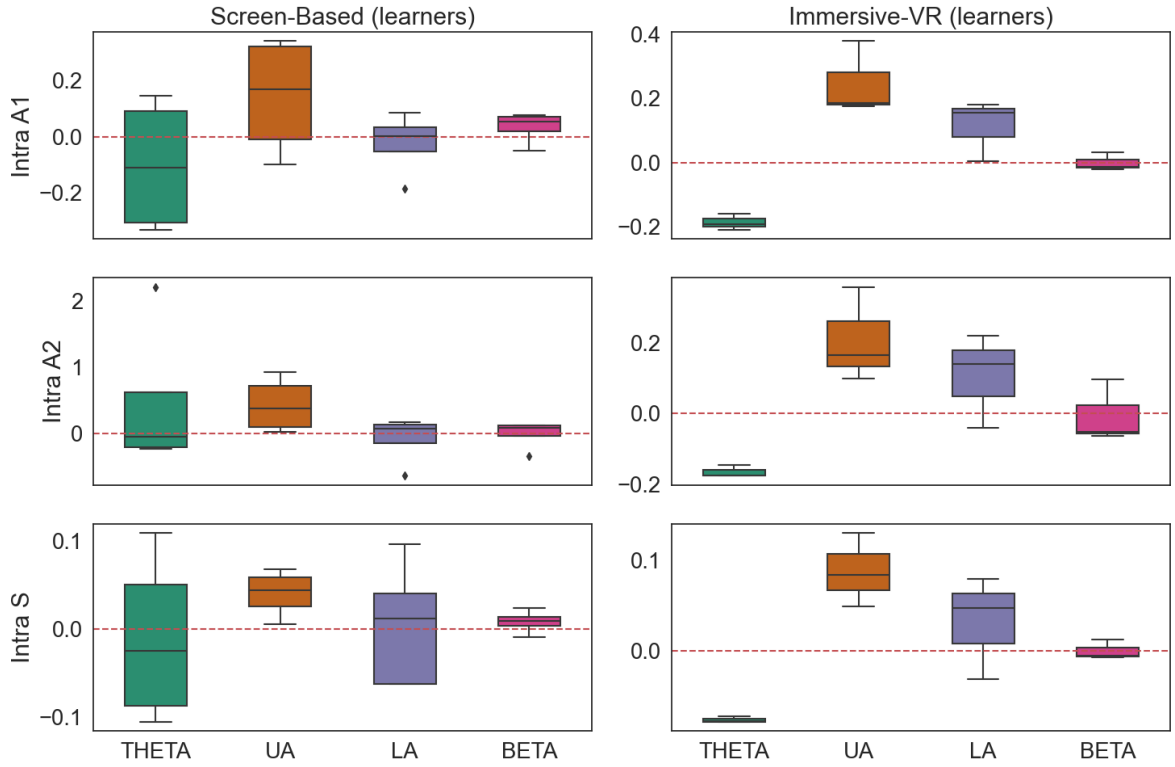

Figure 3.11: Distribution across participants classified as learners of learning indexes corresponding to within session evolution, at Cz for all bands, for each modality.

Across session indexes are represented in Figures 3.12 and 3.13. These appear to have an irregular

distribution among participants and modality groups, which makes it harder to take visual conclusions from. Contrarily to what happens with the indexes calculated for within sessions results, the target frequency band is the most negative of all band, while LA behaves contrarily by presenting the most positive median.

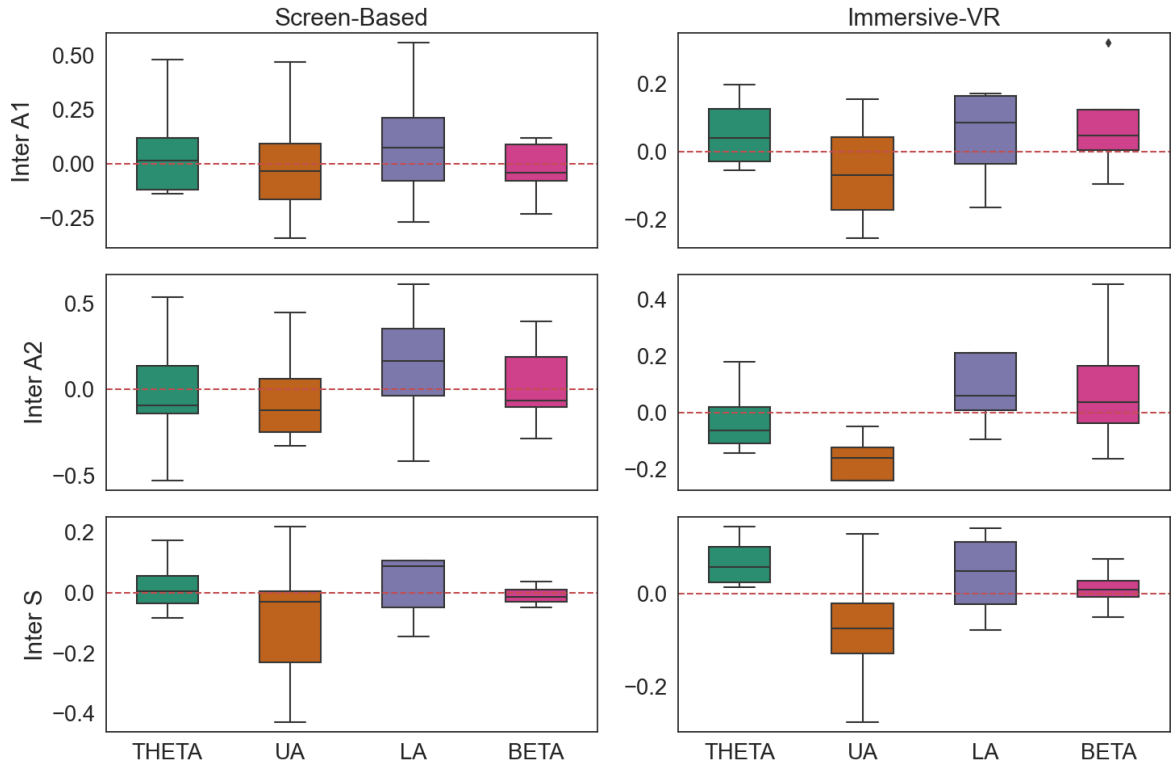

Figure 3.12: Distribution across all participants of learning indexes corresponding to the evolution between sessions, at Cz for all bands, for each modality.

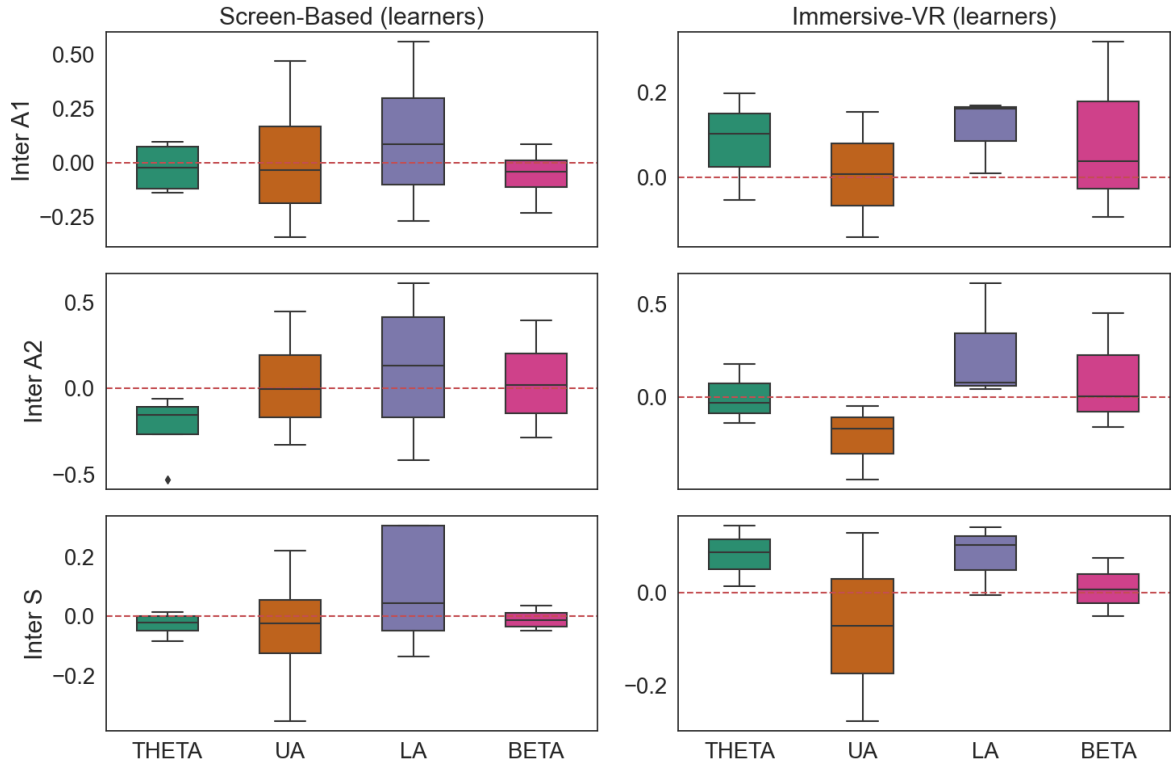

Figure 3.13: Distribuition across participants classified as learners, of learning indexes corresponding to the evolution between sessions, at Cz for all bands, for each modality.

The table below displays the results of the Wilcoxon Signed-Rank test for the distribution of each learning index for all bands. For this test, the null hypothesis would be that the median value for each index is zero. The table shows significant two p-values (< 0.05), in the IntraS and Intra A1 calculated for the Beta frequency range for the Screen-Based modality. Additionally, we see that for the Immersive-VR group, in each band there is no difference in the p-value, in fact, all values happen to be significantly higher than 0.05, which indicates strong evidence for the null hypothesis. However, similarly to what is indicated before, the sample size for the Immersive-VR is quite small, which may affect the results.

	Screen-Based				Immersive-VR			
	Theta	**LA**	**UA**	**Beta**	**Theta**	**LA**	**UA**	**Beta**
IntraS	0.3125	0.742	0.629	**0.039**	0.125	0.375	0.125	0.625
IntraA1	**0.078**	0.641	0.727	**0.039**	0.125	0.375	0.125	0.625
IntraA2	0.383	0.547	0.422	0.195	0.125	0.375	0.125	0.625

Table 3.9: p-values resulted from the Wilcoxon Signed-Rank statistical test for the distribution of the learning indexes for all participants. The null hypothesis would be that the median value of the distribution for each frequency band is zero, with 5% level of significance. Relevant pvalues (< 0.05) are highlighted in color.

	Screen-Based				Immersive-VR			
	Theta	**LA**	**UA**	**Beta**	**Theta**	**LA**	**UA**	**Beta**
InterS	0.844	0.641	0.902	0.547	0.125	0.625	0.8125	0.625
InterA1	0.844	0.547	0.727	0.945	0.625	0.625	0.6875	0.625
InterA2	0.844	0.547	0.844	0.844	0.875	0.625	1.0	0.625

Table 3.10: p-values resulted from the Wilcoxon Signed-Rank statistical test for the distribution of the learning indexes for all participants. The null hypothesis would be that the median value of the distribution for each frequency band is zero, with 5% level of significance.

Chapter 4

Discussion

In this chapter the discussion is presented, divided into two parts: (i) training effect on target feature; (ii) training effect on other frequency bands.

4.1 Training Effect on Target Feature

In general, a training effect was observed for both modalities, more evidently within each session. As Figure 3.6 and Tables 3.1, 3.2, 3.3 and 3.4 demonstrate the median of the target feature corresponding to the last set is consistently higher than the first set of that session, for both modalities. This shows that, in each training session subjects from both groups were able to increase their RAUA willingly. Despite this, the Wilcoxon Signed-Rank test comparing the first and last set of each session did not show any significant difference (Table 3.5). It is, however, important to take into account that this statistical test was performed with a very small data sample (n = 4 for the Immersive-VR; n = 8 for the Screen-Based), and if the assumptions of any statistical test are not met, then the results could be unreliable. Similarly, the statistical test performed on the across sessions indexes (IntraS, Intra A1 and A2), in the UA range showed no significant difference (p value > 0.05), see Table 3.9. Furthermore, results from the Spearman correlation between the set number and mean RAUA across all the corresponding sets of the four sessions (Table 3.6) showed no significance in the UA band.

Comparing with within session performance, in between session results are harder to draw conclusions from. As revealed by the Wilcoxon Signed-Rank test for these learning indexes, InterS, Inter A1 or A2, none were found to be significantly different from zero (Table 3.10). Moreover, results from Table 3.8 and Figure 3.12, show that these learning indices seem to have an irregular distribution among participants and modality groups, which gives very little information about the evolution of the feedback parameter over different the four different sessions. One thing to take into account is that, the metrics chosen to evaluate across sessions results are the same used within session analysis. But in reality, because different session were recorded on different days, one can assume that the participants might have had their training abilities affected by external factors such as their mood or level of tiredness on that day. This suggests that, metrics to compare results from different sessions could have been further investigated. Furthermore, even from session to session most experimental conditions stay constant, i.e. session design, display of feedback, equipment etc., there may exist differences between recordings that be overlooked even after normalization. Another thing to take into account is inter and intra-subject variability, for example Fasanya et al. (2013) highlights the need of taking into account participants' unique characteristics, given the limited number of training procedures that support training and analysis individualization. Finally, Gruzelier (2014) argues that in order to increase the odds of success in

a NF training experience, the number of sessions advised is eight to ten, which is double the amount used in this study (four), this may also have affected the across session training performance.

In the Screen-Based modality group 50% of the subjects were classified as non learners (V2, V3 V4, V8), while only one was considered in the Immersive-VR group (A4). Visually, these subjects appear to have a general decrease in the RAUA within each session. However in Table 3.8, we can see that some values, despite their sign, are relatively close to zero, i.e. the learning slope might be negative but that could be disregarded since the value is ≈ 0.01. The Wilcoxon Signed-Rank test confirms this, by not finding values of IntraS, in the UA band significantly different from zero (*Table 3.9*). For this reason, the criteria used to classify participants (sign of IntraS) as learners or non-learners, could have been further investigated. Regarding IntraA1 and IntraA2, Table 3.9 shows that non-learners consistently have negative values for these indexes as well. This confirms that in some way, learners present an evolution during training within each session. During the analysis of the group median RAUA (Figure 3.6, we notice that the data from the non-learners had more impact in the Screen-Based modality data, which is logical since this group has a higher number of non-learners. In fact, Figures 3.10 and 3.11, show this exactly - when the non-learners are removed there is noticeably more evident shift in the interquartil range of the UA box for the Screen-Based modality group than for the Immersive-VR group. In addition, 3.3 shows that in both modality groups, subjects that were classified as non-learners show considerably higher values of RAUA, which may suggest that greater values are more difficult to increase.

4.2 Training Effect on Other Frequency Bands

The evolution of the relative amplitude of other frequency bands as well as their learning indexes was assessed in order to evaluate training specificity. Furthermore, the statistical tests were extended to frequency bands other than UA and the results are discussed below. Regarding the relative amplitude of the theta band, it appears to have, for the most part, a negative progression for both modalities, in contrast to what happens with the UA band. In fact, in this frequency range, the within session learning indexes were the most negative of all bands, in both modalities. Figures 3.10 and 3.11 highlight the fact that the theta band has a strong negative correlation with the target frequency band, by showing the boxes of these two bands always on opposite sides. Additionally, the Spearman test found a significant negative correlation in the first 2 sessions of the Screen-Based modality group as well as in the first session of the Immersive-VR group, which loses this significance after the non learners are removed. In fact, decreased theta activity is believed to be an indicator of cognitive performance (Cao et al., 2022; Hanslmayr et al., 2005). Similarly to what happened with the UA band, no significant differences from zero were seen for none of the learning indexes in any group. Additionally, no notable correlation was found between the set number and the relative amplitude of this band. Looking at the distribution of the within session indexes, there is not a noticeable difference between this band and UA, both present similar median values and distribution ranges. As for InterS, InterA1 and InterA2, we see that for both groups, while the target frequency band is mostly negative, LA stays consistently above zero, which is more evident among learners.

Concerning the Beta band, we see that the values are significantly higher for the Screen-Based group. In fact, the Wilcoxon Signed-Rank test found that IntraS and IntraA1 were significantly different from zero for this modality group. But looking at the other results, one can assume that this band was not affected by training neither within, nor across sessions.

Chapter 5

Conclusions

To this day, NF training has been successfully used to not only, normalize irregular brain activity related with several neurological and psychiatric conditions (such as epilepsy, schizophrenia or ADHD), but also to enhance cognitive function (such as working memory) in healthy individuals. In the current literature, the majority of NF studies focus on data acquisition and pre-processing as well as learning strategies. In fact, very few studies have compared the effect that different feedback modalities have on training success. The current work consisted on a systematical analysis on the influence of two different type of visual modalities (2-D vs. immersive-VR) in the effectiveness of the NF training results. Data from two previous studies, recorded on healthy participants, in protocols that targeted the increase in the upper alpha (UA) power band was used. An extensive data processing and cleaning protocol was applied and the training effectiveness was measured through band power calculation, definition of learning ability indexes and application of statistical tests.

Results show that a general training effect was observed for both modalities, exclusively within sessions. In each session, both groups showed the ability to increase their RAUA. However, across sessions, results are inconclusive and do not show clear evidence of up-regulation of the target feature. The training effect aside from the target frequency was also assessed and these effects were irrelevant in the Beta band but quite evident in Theta and LA bands. Although the sample size wasn't sufficient to take relevant statistical conclusions when considering only within-session evolution, only the results from the Immersive-VR group showed an increase in the relative amplitude of upper alpha in all sessions. Whereas further investigation is required, the work presented in this thesis showed that the Immersive-VR modality was more effective in increasing the feedback parameter (RAUA) within sessions.

5.1 Limitations and Future Work

Evidently, this study major limitation was the sample size. One group had data from 8 subjects while the other had 4 and after analysing learning abilities this number was count down even more (VR = 3; VIS = 4). Because of this, determining the precise effect of NF training was an unfeasible task. In fact, a limited sample size actually reduces the study's statistical power and increases the margin of error, which decreases the power of the study. Given this, one thing to take into account in future work would be conducting more training sessions with both feedback modalities, to increase the sample size.

As is the case, the learning ability of the experimental group of participants in a NF training session highly affect the effectiveness of training. In the present work, there were non-learners in both modality groups, more so in the Screen-Based group, this together with sample size precluded the analysis and

therefore the results. To prevent this, future studies could incorporate a pre-selection of subjects before the actual training. This could consist on one trial training session in which pre-defined predictors of training success would be measured.

Subject variability was not taken into account on this study. The elements of feedback display, data analysis and processing remained unchanged for all participants. In the future, to improve effectiveness of training, the experimental protocol could be adapted to each participant regarding the feedback display and even individualization in frequency ranges before and after analysis.

Finally, the data cleaning protocol applied in this study was extensive and most certainly had impact in the results. One thing to be added to future work could be assessing the effect of the data cleaning in the effectiveness of training. Questions such as ”How does it affect the learning ability?” and ”Does it have the same effect on both modalities?” would be interesting to answer.

Appendices

Appendix A

Learning indexes

Table A.1: All band-specific learning indexes for all the participants of the Screen-based modality group.

Subject	Band	IntraS	IntraA1	IntraA2	InterA1	InterA2	InterS
V1	UA	0,005	0,020	0,030	-0,133	-0,116	-0,047
V2		-0,025	-0,097	-0,048	-0,159	-0,236	-0,188
V3		-0,011	-0,087	-0,121	0,091	0,050	0,022
V5		-0,069	-0,032	-0,085	0,105	-0,125	-0,014
V6		0,068	0,317	0,646	-0,344	-0,326	-0,352
V7		0,056	0,340	0,931	0,067	0,109	0,000
V8		-0,213	-0,456	-0,172	-0,173	-0,275	-0,430
V4		0,033	-0,098	0,118	0,469	0,447	0,220
V1	THETA	-0,105	-0,296	-0,189	0,069	-0,182	0,014
V2		0,002	-0,019	-0,026	-0,041	0,221	0,018
V3		-0,065	-0,283	-0,093	0,197	0,111	0,175
V5		-0,082	-0,331	-0,228	-0,138	-0,124	-0,035
V6		-0,082	-0,331	-0,228	0,096	-0,059	-0,005
V7		0,109	0,145	2,211	-0,114	-0,528	-0,084
V8		-0,138	-0,386	-0,284	0,480	0,536	0,235
V4		0,031	0,073	0,095	-0,138	-0,124	-0,035
V1	LA	0,002	-0,009	0,026	-0,045	-0,084	-0,019
V2		-0,010	-0,199	-0,044	-0,037	0,354	0,087
V3		-0,034	-0,101	-0,106	0,189	-0,016	0,091
V5		0,022	0,015	0,133	0,213	0,352	0,106
V6		0,022	0,015	0,133	0,561	0,610	0,900
V7		-0,255	-0,186	-0,635	-0,270	-0,415	-0,135
V8		-0,037	0,019	0,368	-0,173	-0,018	-0,145
V4		0,096	0,085	0,168	0,213	0,352	0,106
V1	BETA	0,024	0,068	0,116	-0,011	0,142	0,004
V2		0,005	0,043	0,126	0,108	-0,110	0,007
V3		0,023	0,093	0,153	-0,100	-0,027	-0,031
V5		0,011	0,077	0,063	-0,071	-0,098	-0,029
V6		0,011	0,077	0,063	-0,231	-0,285	-0,049
V7		-0,009	-0,049	-0,335	0,085	0,396	0,036
V8		0,032	0,056	0,235	0,121	0,334	0,018
V4		0,008	0,041	0,127	-0,071	-0,098	-0,029

Table A.2: All band-specific learning indexes for all the participants of the Immersive-VR modality group.

Subject	Band	IntraS	IntraA1	IntraA2	InterA1	InterA2	InterS
A1	UA	0,085	0,185	0,099	-0,143	-0,048	-0,277
A2		0,130	0,379	0,356	0,007	-0,444	-0,072
A3		0,049	0,176	0,165	0,154	-0,169	0,128
A4		-0,034	-0,060	-0,038	-0,256	-0,147	-0,079
A1	THETA	-0,072	-0,160	-0,175	0,198	-0,141	0,085
A2		-0,078	-0,193	-0,145	-0,055	-0,032	0,013
A3		-0,078	-0,210	-0,174	0,103	0,180	0,143
A4		-0,034	-0,060	-0,038	-0,020	-0,095	0,027
A1	LA	-0,031	0,004	-0,041	0,171	0,043	0,101
A2		0,080	0,180	0,218	0,008	0,079	-0,006
A3		0,048	0,154	0,140	0,163	0,613	0,140
A4		0,022	-0,052	0,041	-0,166	-0,093	-0,078
A1	BETA	-0,005	-0,013	-0,063	0,322	0,452	0,073
A2		-0,007	-0,021	-0,051	0,038	0,003	0,006
A3		0,013	0,032	0,096	-0,095	-0,160	-0,052
A4		0,059	0,070	0,080	0,057	0,071	0,012

Appendix B

Data processing and analysis code

```
file =  #INSERT FILE NAME HERE
raw = readrawgdf(file, preload=True)
chnames = raw.chnames
montage = readcustommontage( #LOCATION FILE )
montage.plot(kind='topomap')
chloc = montage.chnames
c = dict(zip(chnames, chloc)) #Combines names with electroded locations
raw.renamechannels(c)
raw.setmontage(montage)
print('Number of channels marked as bad:', len(raw.info['bads']))
```

Listing B.1: Code: data uploading and montage file reading.

```
# Apply band-pass filter (1-45 Hz) and re-reference to average
lofreq = 1.
hifreq = 45.
filtered = raw.filter(lofreq, hifreq, firdesign='firwin', skipbyannotation='
    edge')
filtered.seteegreference(refchannels='average', projection = True).applyproj()
```

Listing B.2: Code: data filtering and re-referencing to average.

```
#EPOCHING
trialsid=-'33282':6
eventstrials,   = eventsfromannotations(raw, eventid=trialsid)
tmin2=0
tmax2=60
croppedtrials = raw.copy()

pickstrials = picktypes(croppedtrials.info, meg=False, eeg=True, stim=False,
                        eog=False,
                  exclude='bads')

epochstrials = Epochs(croppedtrials, eventstrials, trialsid, tmin2,
                      tmax2, proj=True,picks=pickstrials, preload=True,
                      baseline=None,eventrepeated='merge')
```

Listing B.3: Code: Epoching.

```

```

```
# Applying ICA for noise removal

#RUN ICA on Raw cropped
icaraw = ICA(ncomponents=19,method='fastica').fit(croppedsets) #Applies ICA to
    the epoched data
icaraw.plotsources(croppedsets) # Shows the time series of the ICs (SHOW ONLY 1st
     SET)
components = icaraw.plotcomponents(cmap='jet') # plot components
components.save(curdir, 'JPEG')

#VISUALIZING ICA AND CHOOSE AT LEAST ONE COMPONENT
see = int(input("Choose ICA components to see:"))
icaraw.plotproperties(croppedsets,picks=see)
icaraw.plotoverlay(croppedsets, exclude=[see], picks='eeg')

# Removing components based on plots and comparison
#lista = []
#remove = [int(input("Choose ICA components to remove:"))]
icaraw.exclude = [0,1]
reconstepochs = croppedsets.copy()
icaraw.apply(reconstepochs)
reconstepochs.plot()

#Generating new Epochs (after ICA)
newepochssets = Epochs(reconstepochs, eventssets, setsid, tmin1, tmax1, proj=
    True,picks=pickssets, preload=True,baseline=None)
```

Listing B.4: Code: Applying ICA.

```

#textfile = open('') # HERE THE INPUT IS A .TXT FILE WITH EACH SUBJECT'S IAF
iaf = textfile.read().split(',')
iaf = np.array(iaf, dtype=float) #IAF CHANGES ACCORDING TO THE SUBJECT IN QUESTION
    1-7
subject =int(input("Subject number minus 1:"))

IAF = iaf[subject]
LTF = IAF - 2
HTF = IAF + 2



FREQBANDS = -"theta": [4, LTF],
              "alpha": [LTF,HTF],
              "upper alpha": [IAF,HTF],
              "lower alpha": [LTF,IAF],
              "beta": [HTF, 30],
              "normal":[4,30]


set1 = slice(0,6)
set2 = slice(6,12)
set3 = slice(12,18)
set4 = slice(18,24)
set5 = slice(24,30)

```

```
27
28 sets = [set1,set2,set3,set4,set5]
29
30 bpstotal=[]
31 bpsband=[]
32 bpsrelative=[]
33
34 def psdcz(epochs,setnum,band,db=False): #CURRENTLY WORKING FOR CZ CHANNEL
35     #CALCULATE TOTAL PSD FOR ALL EPOCHS BETWEEN 4-30 HZ
36     psds,freqs = psdmultitaper(epochs,fmin=4,fmax=30)
37
38     # DEFINE OUR BP ACCORDING TO THE SET + AVERAGING THE PSD + CALCULATE TOTAL BP
39     bp = psds[setnum,24,:].mean(0)
40     bp = (sum(bp))/(30-4)
41     bpstotal.append(bp)
42
43     # DEFINE HERE OUR BAND FREQ VECTOR AND GET PSD IN THAT RANGE
44     idxband = np.logicaland(freqs ¿= min(band), freqs ¡= max(band))
45     bpband = psds[setnum,24,idxband].mean(0) #24 BECAUSE Cz = ch.24
46     bpband = (sum(bpband))/(max(band)-min(band))
47     bpsband.append(bpband)
48
49     if db:
50         psds = 10. * np.log10(psds)
51         bp = psds[setnum,24,:].mean(0)
52         bp = (sum(bp))/(30-4)
53         idxband = np.logicaland(freqs ¿= min(band), freqs ¡= max(band))
54         bpband = psds[setnum,24,idxband].mean(0)
55         bpband = (sum(bpband))/(max(band)-min(band))
56         bpsband.append(bpband)
57
58     #GATHER RELATIVE VALUES AND PRINT THEM
59     bprelative= bpband/bp
60     bpsrelative.append(bprelative)
61     print('UA Band power: ', bpband)
62     print('Relative UA power: ', bprelative)
63
64
65 for i in sets:
66     psdcz(epochstrials,i,FREQBANDS["upper alpha"])
67     #psdcz(epochstrials,i,FREQBANDS["theta"])
68     #psdcz(epochstrials,i,FREQBANDS["lower alpha"])
69     #psdcz(epochstrials,i,FREQBANDS["beta"])
```

Listing B.5: Code: Calculating Band Power

References

Merriam-Webster.com dictionary. Merriam-Webster., 2022. Virtual Reality.

P. A. Abhang, B. W. Gawali, and S. C. Mehrotra. *Introduction to EEG-and speech-based emotion recognition*. Academic Press, 2016.

F. Accoto, A. Vourvopoulos, A. Gonçalves, T. Bucho, G. Caetano, P. Figueiredo, L. De Paolis, and S. B. i Badia. The effect of neurofeedback training in cave-vr for enhancing working memory. In *Technology-Augmented Perception and Cognition*, pages 11–45. Springer, 2021.

C. Amo, L. De Santiago, R. Barea, A. López-Dorado, and L. Boquete. Analysis of gamma-band activity from human eeg using empirical mode decomposition. *Sensors*, 17(5):989, 2017.

C. Berhanu et al. Connectivity-based eeg-neurofeedback in vr. Master's thesis, Instituto Superior Técnico, 2019.

Bitbrain. All about eeg artifacts and filtering tools, 2020. [Online]. Available at: `https://www.bitbrain.com/blog/eeg-artifacts`.

T. Bucho, G. Caetano, A. Vourvopoulos, F. Accoto, I. Esteves, S. B. i Badia, A. Rosa, and P. Figueiredo. Comparison of visual and auditory modalities for upper-alpha eeg-neurofeedback. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5960–5966. IEEE, 2019.

Y. Cao, C. Han, X. Peng, Z. Su, G. Liu, Y. Xie, Y. Zhang, J. Liu, P. Zhang, W. Dong, et al. Correlation between resting theta power and cognitive performance in patients with schizophrenia. *Frontiers in Human Neuroscience*, 16, 2022.

M. Chaumon, D. V. Bishop, and N. A. Busch. A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of neuroscience methods*, 250: 47–63, 2015.

W. Commons. Electrode locations of international 10-20 system for eeg recording., 2010. [Online]. Available at: `https://en.wikipedia.org/wiki/File:21electrodesofInternational10-20systemforEEG.svg`.

E. J. Davelaar. Mechanisms of neurofeedback: a computation-theoretic approach. *Neuroscience*, 378: 175–188, 2018.

T. Egner and M. B. Sterman. Neurofeedback treatment of epilepsy: from basic rationale to practical application. *Expert Review of Neurotherapeutics*, 6(2):247–257, 2006.

S. Enriquez-Geppert, R. J. Huster, and C. S. Herrmann. Eeg-neurofeedback as a tool to modulate cognition and behavior: a review tutorial. *Frontiers in human neuroscience*, 11:51, 2017.

S. Enriquez-Geppert, D. Smit, M. G. Pimenta, and M. Arns. Neurofeedback as a treatment intervention in adhd: Current evidence and practice. *Current psychiatry reports*, 21(6):1–7, 2019.

B. Fasanya, O. Omotoso, and O. Fasanya. Effect of age on inter and intra-subject variability in acceptable noise level (anl) in listeners with normal hearing. *Management science letters*, 3(2):385–394, 2013.

T. Fernández, J. Bosch-Bayard, T. Harmony, M. I. Caballero, L. Díaz-Comas, L. Galán, J. Ricardo-Garcell, E. Aubert, and G. Otero-Ojeda. Neurofeedback in learning disabled children: visual versus auditory reinforcement. *Applied psychophysiology and biofeedback*, 41(1):27–37, 2016.

J. A. Frederick. Psychophysics of eeg alpha state discrimination. *Consciousness and cognition*, 21(3): 1345–1354, 2012.

J. A. G. Gnecchi, J. C. H. Garcia, and J. d. D. O. Alvarado. Auxiliary neurofeedback system for diagnostic of attention deficit hyperactivity disorder. In *Electronics, Robotics and Automotive Mechanics Conference (CERMA 2007)*, pages 135–138. IEEE, 2007.

A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, et al. MEG and EEG data analysis with MNE-python. *Frontiers in neuroscience*, page 267, 2013.

J. H. Gruzelier. Eeg-neurofeedback for optimising performance. i: A review of cognitive and affective outcome in healthy participants. *Neuroscience & Biobehavioral Reviews*, 44:124–141, 2014.

S. Hanslmayr, P. Sauseng, M. Doppelmayr, M. Schabus, and W. Klimesch. Increasing individual upper alpha power by neurofeedback improves cognitive performance in human subjects. *Applied psychophysiology and biofeedback*, 30(1):1–10, 2005.

L. Hu and Z. Zhang. *EEG signal processing and feature extraction*. Springer, 2019.

R. J. Huster, Z. N. Mokom, S. Enriquez-Geppert, and C. S. Herrmann. Brain–computer interfaces for eeg neurofeedback: Peculiarities and solutions. *International journal of psychophysiology*, 91(1):36–45, 2014.

X. Jiang, G.-B. Bian, and Z. Tian. Removal of artifacts from eeg signals: a review. *Sensors*, 19(5):987, 2019.

T.-P. Jung, S. Makeig, C. Humphries, T.-W. Lee, M. J. Mckeown, V. Iragui, and T. J. Sejnowski. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(2):163–178, 2000.

T. Kirschstein and R. Köhling. What is the source of the eeg? *Clinical EEG and neuroscience*, 40(3): 146–149, 2009.

Kiyoshi Ota. The playstation vr headset in tokyo has emerged as the world's top vr device. `https://www.ft.com/content/bff48624-daf5-11e7-a039-c64b1c09b482`, 2016. [Online; accessed 19-July-2022].

W. Klimesch, H. Schimke, G. Ladurner, and G. Pfurtscheller. Alpha frequency and memory performance. *Journal of Psychophysiology*, 1990.

M. Klug and K. Gramann. Identifying key factors for improving ica-based decomposition of eeg data in mobile and stationary experiments. *European Journal of Neuroscience*, 54(12):8406–8420, 2021.

S. E. Kober, D. Schweiger, J. L. Reichert, C. Neuper, and G. Wood. Upper alpha based neurofeedback training in chronic stroke: brain plasticity processes and cognitive effects. *Applied psychophysiology and biofeedback*, 42(1):69–83, 2017.

J. Kropotov et al. *Quantitative EEG, event-related potentials and neurotherapy*. Academic Press, 2010.

L. Leuchs et al. Choosing your reference–and why it matters. *Brain Products*, pages 03–05, 2019.

E. K. S. Louis, L. Frey, J. Britton, J. Hopp, P. Korb, M. Koubeissi, W. Lievens, and E. Pestana-Knight. Electroencephalography (eeg): An introductory text and atlas of normal and abnormal findings in adults. *Children, and Infants*, 2016.

J. F. Lubar and W. Bahler. Behavioral management of epileptic seizures following eeg biofeedback training of the sensorimotor rhythm. *Biofeedback and Self-regulation*, 1(1):77–104, 1976.

H. Marzbani, H. R. Marateb, and M. Mansourian. Neurofeedback: a comprehensive review on system design, methodology and clinical applications. *Basic and clinical neuroscience*, 7(2):143, 2016.

V. Menon et al. Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in cognitive sciences*, 15(10):483–506, 2011.

M. Mihara, I. Miyai, N. Hattori, M. Hatakenaka, H. Yagura, T. Kawano, M. Okibayashi, N. Danjo, A. Ishikawa, Y. Inoue, et al. Neurofeedback using real-time near-infrared spectroscopy enhances motor imagery related cortical activation. *PloS one*, 7(3):e32234, 2012.

S. Murakami and Y. Okada. Contributions of principal neocortical neurons to magnetoencephalography and electroencephalography signals. *The Journal of physiology*, 575(3):925–936, 2006.

W. Nan, F. Wan, M. I. Vai, and A. C. Da Rosa. Resting and initial beta amplitudes predict learning ability in beta/theta ratio neurofeedback training in healthy young adults. *Frontiers in human neuroscience*, 9:677, 2015.

C. S. Nayak and A. C. Anilkumar. Eeg normal waveforms. *StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing*, 2019.

M. Nazer, H. Mirzaei, and M. Mokhtaree. Effectiveness of neurofeedback training on verbal memory, visual memory and self-efficacy in students. *Electronic physician*, 10(9):7259, 2018.

M. Ninaus, S. Kober, M. Witte, K. Koschutnig, C. Neuper, and G. Wood. Brain volumetry and self-regulation of brain activity relevant for neurofeedback. *Biological psychology*, 110:126–133, 2015.

S. Niv et al. Clinical efficacy and potential mechanisms of neurofeedback. *Personality and Individual Differences*, 54(6):676–686, 2013.

P. L. Nunez, R. Srinivasan, et al. *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA, 2006.

Y. Okumura, Y. Kita, M. Omori, K. Suzuki, A. Yasumura, A. Fukuda, and M. Inagaki. Predictive factors of success in neurofeedback training for children with adhd. *Developmental Neurorehabilitation*, 22 (1):3–12, 2019.

A. Plerou, P. Vlamos, and A. Margetaki. EEG analysis of the neurofeedback training effect in algorithmic thinking. In *GeNeDis 2016*, pages 313–324. Springer, 2017.

D. Purves, G. J. Augustine, and D. Fitzpatrick. *Neuroscience*. Oxford University Press, 6 edition, 2018.

J. L. Reichert, S. E. Kober, C. Neuper, and G. Wood. Resting-state sensorimotor rhythm (smr) power predicts the ability to up-regulate smr in an eeg-instrumental conditioning paradigm. *Clinical Neurophysiology*, 126(11):2068–2077, 2015.

M. Reiner, J. Gruzelier, P. D. Bamidis, and T. Auer. The science of neurofeedback: Learnability and effects. *Neuroscience*, 378:1–10, 2018.

Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, O. Bertrand, and A. Lécuyer. Openvibe: An open-source software platform to design, test, and use brain–computer interfaces in real and virtual environments. *Presence*, 19(1):35–53, 2010.

T. Ros, B. J Baars, R. A. Lanius, and P. Vuilleumier. Tuning pathological brain oscillations with neurofeedback: a systems neuroscience framework. *Frontiers in human neuroscience*, 8:1008, 2014.

D. L. Schomer and F. H. Lopes da Silva. *Niedermeyer's electroencephalography*. Lippincott Williams & Wilkins, 5 edition, 2010.

D. Seok, S. Lee, M. Kim, J. Cho, and C. Kim. Motion artifact removal techniques for wearable EEG and PPG sensor systems. *Frontiers in Electronics*, 2:685513, 2021.

W. R. Sherman and A. B. *Understanding virtual reality : interface, application, and design*. Cambridge, Usa Morgan Kufmann Publishers, 2 edition, 2019. ISBN 9780128183991.

ST Engineering Antycip. Virtual reality cave solutions. `https://steantycip.com/vr-cave/`, 2022. [Online; accessed 19-July-2022].

E. K. St. Louis, L. C. Frey, and J. W. Britton. *Electroencephalography (EEG)*. American Epilepsy Society, 2016.

M. Sterman, L. Macdonald, and R. Stone. Biofeedback training of sensorimotor electroencephalogram rhythm in man-effects on epilepsy. *Epilepsia*, 15(3):395–416, 1974.

T. Surmeli, A. Ertem, E. Eralp, and I. H. Kos. Schizophrenia and the efficacy of qeeg-guided neurofeedback treatment: a clinical case series. *Clinical EEG and Neuroscience*, 43(2):133–144, 2012.

W. O. Tatum, B. A. Dworetzky, and D. L. Schomer. Artifact and recording concepts in EEG. *Journal of clinical neurophysiology*, 28(3):252–263, 2011.

R. T. Thibault, M. Lifshitz, and A. Raz. The self-regulating brain and neurofeedback: experimental science and clinical promise. *cortex*, 74:247–261, 2016.

R. T. Thibault, A. MacPherson, M. Lifshitz, R. R. Roth, and A. Raz. Neurofeedback with fmri: A critical systematic review. *Neuroimage*, 172:786–807, 2018.

A. Vourvopoulos, O. M. Pardo, S. Lefebvre, M. Neureither, D. Saldana, E. Jahng, and S.-L. Liew. Effects of a brain-computer interface with virtual reality (VR) neurofeedback: A pilot study in chronic stroke patients. *Frontiers in human neuroscience*, 2019.

R. Vullings, C. Peters, R. Sluijter, M. Mischi, S. Oei, and J. Bergmans. Dynamic segmentation and linear prediction for maternal ecg removal in antenatal abdominal recordings. *Physiological measurement*, 30(3):291, 2009.

F. Wan, W. Nan, M. I. Vai, and A. Rosa. Resting alpha activity predicts learning ability in alpha neurofeedback. *Frontiers in human neuroscience*, 8:500, 2014.

E. Weber, A. Köberl, S. Frank, and M. Doppelmayr. Predicting successful learning of smr neurofeedback in healthy participants: methodological considerations. *Applied Psychophysiology and Biofeedback*, 36(1):37–45, 2011.

T. E. Wilens and T. J. Spencer. Understanding attention-deficit/hyperactivity disorder from childhood to adulthood. *Postgraduate medicine*, 122(5):97–109, 2010.

M. Witte, S. E. Kober, M. Ninaus, C. Neuper, and G. Wood. Control beliefs can predict the ability to up-regulate sensorimotor rhythm during neurofeedback training. *Frontiers in human neuroscience*, 7:478, 2013.

N. Yan, J. Wang, M. Liu, L. Zong, Y. Jiao, J. Yue, Y. Lv, Q. Yang, H. Lan, and Z. Liu. Designing a brain-computer interface device for neurofeedback using virtual environments. *Journal of Medical and Biological Engineering*, 28(3):167–172, 2008.

B. Zoefel, R. J. Huster, and C. S. Herrmann. Neurofeedback training of the upper alpha frequency band in eeg improves cognitive performance. *Neuroimage*, 54(2):1427–1431, 2011.