UNIVERSITY OF SALENTO

Department of Engineering for Innovation

Master degree in *Computer Engineering*

Master Thesis in
High Performance Computing

Assessing the effect of Virtual Reality elements in Upper Alpha Neurofeedback

Supervisor

Prof. Lucio Tommaso De Paolis

Co-Supervisors

Prof. Sergi Bermúdez i Badia

Dr. Athanasios Vourvopoulos

Student

Floriana Accoto

ACADEMIC YEAR 2017/2018

Abstract

In recent years, researchers provided evidence of the positive impact of Virtual Reality (VR) on Neurofeedback (NF) training. The immersive properties of VR training scenarios have been shown to facilitate NF learning compared to traditional training methods. However, in the design of an immersive virtual environment, there are several factors that can be manipulated to influence the level of immersion, and how specific factors contribute to the improvement of NF performance has not yet be clarified.

Therefore, the aim of this thesis was to investigate the effects of vividness (i.e., the visual realism of the virtual environment), one of the immersion's dimension, on NF training outcome. To this end, we carried out an experiment in which participants received NF training to enhance Upper Alpha (UA) amplitude. Participants were divided into three experimental groups, each receiving feedback in a different NF training scenario with increasing level of vividness (i.e., low, medium, high). Furthermore, as a secondary objective, we examined the effect of the UA enhancement protocol on working memory performance.

Results of this research showed variable NF learning performance, with better performance in higher vividness compared to lower vividness groups. Moreover, highly vivid training scenarios had a positive effect on some user experience variables: they increased motivation and concentration of NF users and reduced boredom. This suggests that NF training scenarios can be improved by the design of virtual environments highly vivid and realistic.

Finally, results confirmed the efficacy of the UA enhancement protocol in improving working memory performance.

Acknowledgements

The work described in this thesis was conducted at NeuroRehabLab of Madeira Interactive Technologies Institute. Being part of NeuroRehabLab research group has been an enriching experience and gave me the opportunity to work in an interesting and innovative area.

I thank my supervisor Professor Lucio De Paolis for his advice and for encouraging me in this challenging experience.

I would like to express my gratitude to Professor Sergi Bermúdez i Badia and Doctor Athanasios Vourvopoulos for their supervision and guidance throughout this project.

I am very thankful to Afonso Gonçalves, whose collaboration and support has been invaluable for completing this work.

I received a precious help from Teresa Bucho. I would like to thank her for the stimulating discussion relevant to my project and the important materials she provided.

Also, I would like to say thanks to all the members of NeuroRehabLab and the participants involved in my experiment, who have shared their precious time with me.

Finally, I thank my family. The encouragement, comfort, and relief received are immeasurable.

Contents

List of figures

List of Abbreviations

ADHD	Attentional Deficit Hyperactivity Disorder
BCI	Brain-computer Interface
CAVE	Cave Automatic Virtual Environment
DS	Digit Span
EEG	Electroencephalography
FFT	Fast Fourier Transform
HMD	Head-Mounted Display
NF	Neurofeedback
NFT	Neurofeedback Training
OCD	Obsessive-Compulsive Disorder
PTSD	Post-Traumatic Stress Disorder
UA	Upper Alpha
SUD	Substance Use Disorder
VE	Virtual Environment
VR	Virtual Reality

1 Introduction

Brain-Computer Interfaces (BCI) are communication systems which enable a direct connection between the brain and a computer [1]. BCI systems acquire brain signals, usually through Electroencephalography, analyse them and translate them into output commands that operate a computer display or other device.

Interest on BCI technology has rapidly grown over the past two decades [2] and progress in BCI research is paving the way for new solutions and applications in many fields, such as rehabilitation, neuroscience and cognitive neuroscience [3, 4, 5].

Even though recently BCIs are beginning to show their potentiality in fields like gaming, entertainment, safety and security, in existing research, applications of BCI are mainly focused on two major areas [2, 5, 6, 7].

The first regards assistive and restorative technology. BCI makes it possible for people with motor disabilities to regain the interaction with external environment (through the control of prosthetic devices) in order to improve the quality of their lives [8] or promote neuroplastic changes in order to achieve the reorganization of motor networks to attain functional motor recovery [9]. The second area is called Neurofeedback (NF) and it is what we will focus on in this thesis.

Neurofeedback is one of the most promising field of Biofeedback. Also known as EEG (Electroencephalogram) Biofeedback, it entails learning to self-regulate brain activity, with the aim of improving mental states or processes. During Neurofeedback Training the user receives real-time feedback of one’s own electrical brain activity assessed with the electroencephalogram. Specific components of the EEG are extracted online and fed back to the user, for instance via visual or auditory feedback. This enables the user to consciously perceive their own electrical brain activity, which is otherwise impossible since there are no somatic receptors to register the electrical brain activity as measured by the EEG. Consequently, the user learns associations between specific mental states and desired brain activation patterns [10].

It has been proved that voluntary modulation of specific EEG parameters generally leads to improvements in behaviour and cognition [11]. This makes Neurofeedback an important tool, whether or not in clinical conditions: it represents a method for cognitive enhancement in healthy subjects and also a therapeutic tool for neurological patients.

Although Neurofeedback has demonstrated benefits in many aspects, a critical issue in NF studies is that not all subjects showed satisfactory learning ability to regulate electrical brain activity [12]: about 15-30% of NF users cannot attain control over their brain signals [10].

In the BCI community, this inability to use BCI applications is called “BCI-illiteracy phenomenon”. There are different attempts to explain this phenomenon [13], but the definite reason why some people cannot control their own brain signals remains largely elusive.

Nevertheless, there are some prior studies providing evidence for psychological aspects influencing BCI and NF performance. For instance, motivation of the user turned out to play an important role [14]. It should also be considered that, to obtain cognitive or behavioural improvements, a large number of repeated NF training sessions are mandatory, and this can make NF users bored and tired. Furthermore, NF practice requires users to stay focused and concentrated on the NF task over a long training period [15].

In this context, the feedback design might play a crucial role. Traditional feedback modalities use auditory (e.g., a tone changes its volume or pitch in dependence on the brain activity level) and/or two-dimensional (2D) visual (e.g., simple bars or circles increase/decrease in size in dependence on the brain activity level) stimuli. Such relatively monotonous feedback methods might not attract users to focus on them [5], leading to decreased motivation, interest, concentration, and finally to a lower NF performance and success rate [14]. Hence, an increasing number of recent NF and BCI studies use Virtual Reality (VR) based feedback designs [10].

Nevertheless, still little is known about the effectiveness of VR based NF modalities. To date, studies on this topic mainly focused on the effects of dimensionality (comparing traditional 2D vs. 3D VR based feedback) and results suggest that neurofeedback training is more effective with immersive virtual environments (VE) when compared with traditional 2D feedback modalities.

Aim of this thesis

With this thesis we want to contribute to research on this topic, trying to examine in more detail the effects of immersive VR on neurofeedback. Even if literature highlights the importance of immersion, the significance of the feedback design is mainly unexplored, and it is still not clear which aspects of immersion are more significant and contribute to better NF performance.

According to a well-defined framework for immersive VE [16], we can identify different dimensions that contribute to make a VE immersive. One of these is Vividness, that is associated with the visual fidelity and resolution with which the real world is rendered in the virtual environment.

The first objective of this thesis is to investigate the effect of vividness on NF performance and subjective user experience. Hence, in the neurofeedback study performed, we designed three virtual environments, differing from one another in the level of vividness. Participants were divided into three groups and underwent five neurofeedback training sessions. Each group was exposed to feedback in a different virtual environment during the Neurofeedback procedure.

An Upper Alpha (UA) neurofeedback protocol was used, in which participants should learn to voluntarily increase their brain activity in the UA frequency band. Alpha training is one of the most commonly used protocols since Alpha is widely shown to be correlated with cognitive performance. As stated by Zoefel et al. [17], it is reasonable to choose a frequency band that is associated with certain cognitive functions to increase the probability of reliable behavioural effects as well as applicability. Hence, to promote interpretability of neurofeedback study results, we chose this protocol. In particular, according to literature, Upper Alpha is correlated with working memory. Thus, the second objective of this study is to assess the effect of UA neurofeedback training on working memory performance.

Thesis structure

In the next chapter, we will provide some background information about neurofeedback. We will introduce what brain waves are and how neurofeedback is used to train them. We will describe the main applications of neurofeedback and review the literature regarding, in particular, the use of Upper Alpha neurofeedback for memory enhancement. Finally, we will describe how immersive VR is defined in literature and we will provide a review of the NF studies investigating the effect of immersive VR in neurofeedback practice.

In chapter 3, we define our research questions and briefly summarize how the literature review brought us to the definition of this study.

In chapter 4, we provide all the details about the study and how we conducted it. We explain which our independent and dependent variables are and describe the experimental procedure.

In chapter 5, we present and discuss the results of the study.

In chapter 6, we draw our conclusion and discuss directions for future research.

2 Background

In order to better understand the science and implications of NF training, it is important to first provide some background regarding neurofeedback (*section 2.2*), along with aspects of brain behaviour that are implicated in neurofeedback (*section 2.1*). In *section 2.3*, a literature review focused on immersive Virtual Reality and its use for Neurofeedback training.

2.1 Electroencephalography (EEG) and Brain Waves

The roots of neurofeedback and the related field of electroencephalography can be traced back to Berger, a German psychiatric who recorded the first human EEG in 1924 [18]. EEG is a non-invasive recording method to measure electrical activity of the brain.

The human brain constitutes of billions of neurons, that generate electrical impulses to communicate with one another (neural firing). By placing electrodes on the scalp, this electrical activity can be detected and recorded, and the resulting output is known as the electroencephalogram (EEG). More specifically, the EEG results from the synchronous firing of a specific type of neurons in the cortex, known as pyramidal [18]. This synchronous electrical activity is referred to as brain oscillations or brain waves.

In general, a raw EEG recording is comprised of a collection of neural oscillations in several frequencies. After raw brainwave signal is recorded in a digital format, it can be transformed into brainwave data, by extracting information about the extent of unique frequency bands that is contributing to the overall power of a waveform. To do that Fast Fourier Transform (FFT) is a common processing method applied to EEG recordings that breaks down the raw EEG into the discrete patterns of electrical activity oscillating within it [19].

These patterns of electrical activity are distinguished into different brain waves based on their frequency, that represents how fast the waves oscillate, as measured by number of waves per second or Hertz (Hz). Each brain wave has an amplitude in microvolt (μV), which determines the power of the wave.

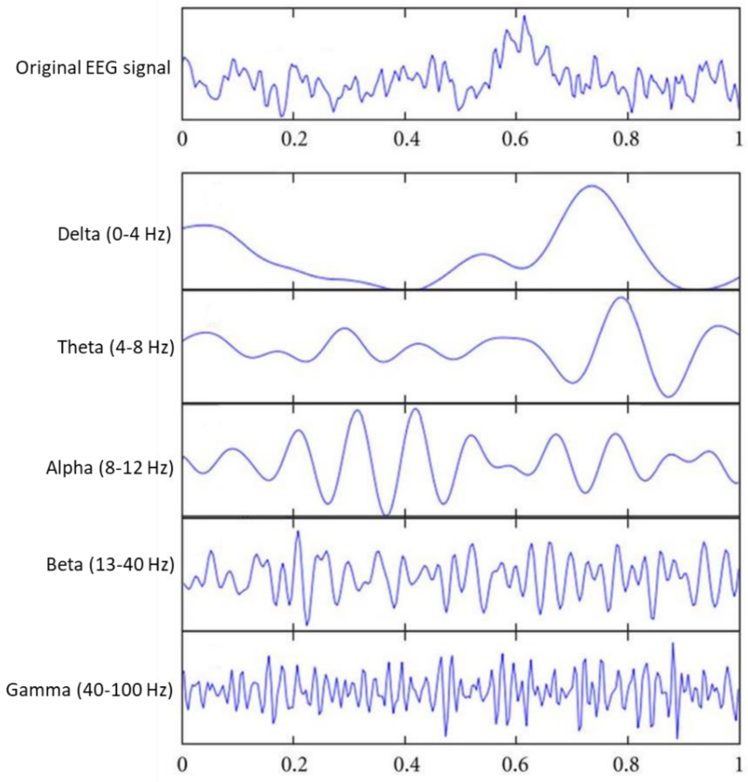


Figure 2.1 Example of a raw EEG signal filtered into its component frequencies [20]

With EEG, researchers had the opportunity of identifying the relationship between brain oscillations and different mental or behavioural states. Berger himself was the first to describe a predominant emerging rhythm of our brain. This rhythm oscillated between 7.8 and 13 Hz when subjects had their eyes closed, and it was replaced by a faster one when subjects opened their eyes. He also verified how this phenomenon was reproduced in response to other sensory stimuli, which made him conclude that those waves should be the fundamental activity of the cortex [18]. Today, we call to these brain waves "alpha waves", also known as "Berger's waves". Since then, the scientific community has found a wider variety of different brain waves associated with different subjective phenomena.

Brain waves are traditionally classified into delta (<4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (13-40 Hz), and gamma (>40 Hz) [21]. The designation of the range of Hz covered by these frequency bands is somewhat arbitrary and not always consistent in the literature. Moreover, these frequency components have subsets. For example, sensorimotor rhythm (SMR) frequency band (13-15 Hz) is related to the sensorimotor rhythm and entitled as low beta [22]. Alpha rhythm is usually divided in two subsets: lower alpha in the range of 8-10 Hz and upper alpha in the range of 10-12 Hz [22].

It is important to note that all of the traditional frequency bands are present at all times across the scalp, but it depends on the task being undertaken by the individual and the scalp location in question as to which is the most prevalent. In general, the faster the oscillation of the most prevalent frequency band the more alert the individual is thought to be. So, delta waves tend to dominate the EEG when the individual is asleep, theta when the individual is drowsy, alpha when the individual is relaxed but alert, beta when the individual is alert and concentrating, and gamma when the individual is trying to solve problems [22]. However, this association between EEG rhythms and activation state is a convenient simplification, because each frequency band may reflect many diverse functional states of neural communication and may be generated through different processes [23].

In the next section a brief summary of the main types of brain waves and their associated functions.

2.1.1 Brain waves classification

Delta

Delta brain waves range from 0 to 4 Hz (lowest frequencies). They are the slowest band of brain waves but with the highest amplitude. Delta waves are associated with the unconscious mind and are created when in deep meditation and dreamless sleep. Delta waves are also linked with intuition, empathy and hormones such as the human growth hormone (HGH). They play an important role in intuition and understanding others on an emotional level. In general, people who have more delta brain wave activity often feel calmer, happier and can understand other people's feelings better.

Theta

Theta brain waves lie within the range of 4 to 8 Hz. They are slower frequencies but have greater amplitudes compared to alpha brain waves. They are identified with the subconscious mind. They occur most often when daydreaming, sleep and deep meditation. Theta waves are also associated with creativity and spirituality. They act as a gateway to learning and memory. By stimulating the brain to generate more theta waves (but not in excess during waking hours), this can improve intuition and creativity, and thus enhance learning ability. Theta brain waves are also involved in restorative sleep.

Alpha

Alpha brain waves range from 8 to 12 Hz. They are sinusoidal-shaped signal. Alpha range is in between beta and theta. They can be found in the occipital and parietal regions of the brain. Alpha brain waves are produced in a normal waking state of consciousness. More alpha waves are produced when people are in a relaxed condition compared to when they are concentrating. Hence, alpha waves are known as “relaxing waves”. An abundant amount of alpha brain waves is important to prevent an individual from encountering different mental and physical problems. Alpha is the resting mental state of the brain. With high alpha wave activity, one feels more relaxed and the heart rate slows down. Hence, body recovery by itself is faster than in a beta state. The alpha brain waves also act as a bridge between the subconscious (theta waves) and the conscious (beta waves) mental state. This enables information, feelings, creativity and memories, which are deep down in one’s mind, to become conscious. Therefore, the functions of alpha brain waves are to assist relaxation and to improve self-regulation, body and mind integration and learning.

Beta

Beta brain waves lie within the range of 13 to 40 Hz. They are high-frequency low amplitude brain waves. They can be found in the frontal and parietal regions of the brain. Beta brain waves are associated with normal consciousness and at an elevated state of alertness, judgemental and rational reasoning. The brain produces beta waves when an individual is reading, thinking and focusing to solve a problem or to complete a task. Hence, beta brain waves are also called the “thinking waves”. A healthy dose of beta brain waves enables us to concentrate and accomplish a task faster. Although beta brain waves are necessary for effective brain operation, having too much beta brain waves are harmful to us as it may cause stress, anxiety and restlessness. Too much beta brain waves will also inhibit the production of alpha brain waves which will lead to other health issues such as depression.

Gamma

Gamma brain waves range from 40 to 100 Hz and are the highest frequency brain waves. They are involved in higher processing tasks such as memory storage, learning and formation of ideas and data processing. It has been known that people with learning disabilities tend to have lower average gamma activity.

2.1.2 The 10-20 International Electrode Placement System

Depending on the location of the scalp where electrodes are placed, a different kind of brain waves can be found. The 10-20 International Electrode Placement System [24] was developed as a way of standardising the locations on the scalp and thus enabling comparability of data. The numbers ‘10’ and ‘20’ refer to the fact that the distances between adjacent electrodes are either 10% or 20% of the total front-back or right-left distance of the skull (*Figure 2.2*).

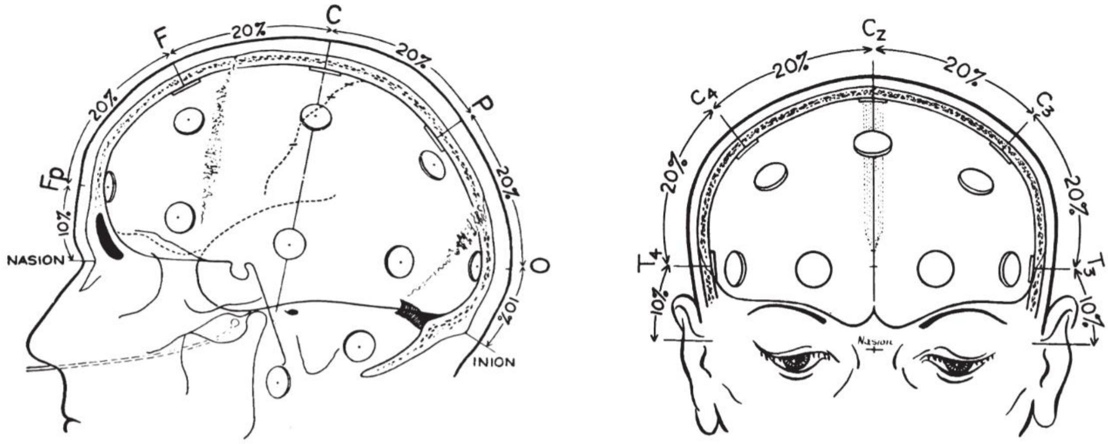


Figure 2.2 Front and side view of the skull, showing the methods of measurements for electrodes placement in the 10-20 system *[24]*

Each electrodes location is named using a letter and a number, with letters referring to the lobe of the brain they are positioned over and the numbers relating to the hemisphere and location of that part of the hemisphere (see *Figure 2.3* below). So, the letters F, P, T, O and C mean that the scalp locations are over the frontal, parietal, temporal, occipital and central regions of the brain respectively. Odd numbers (1, 3, 5 …) refer to scalp locations on the left side of the brain and even numbers (2, 4, 6 …) refer to scalp locations on the right side of the brain. When the letter z replaces a number (i.e., Fz, Cz, Pz, Oz) this indicates that the scalp location falls along the central line running between the nasion (bridge of the nose) and the inion (base of the occipital bone which protrudes from the back of the skull). Additional locations are indicated by the letters Fp that stands for frontal poles, and A that indicates earlobes, where usually ground and reference electrodes are placed.

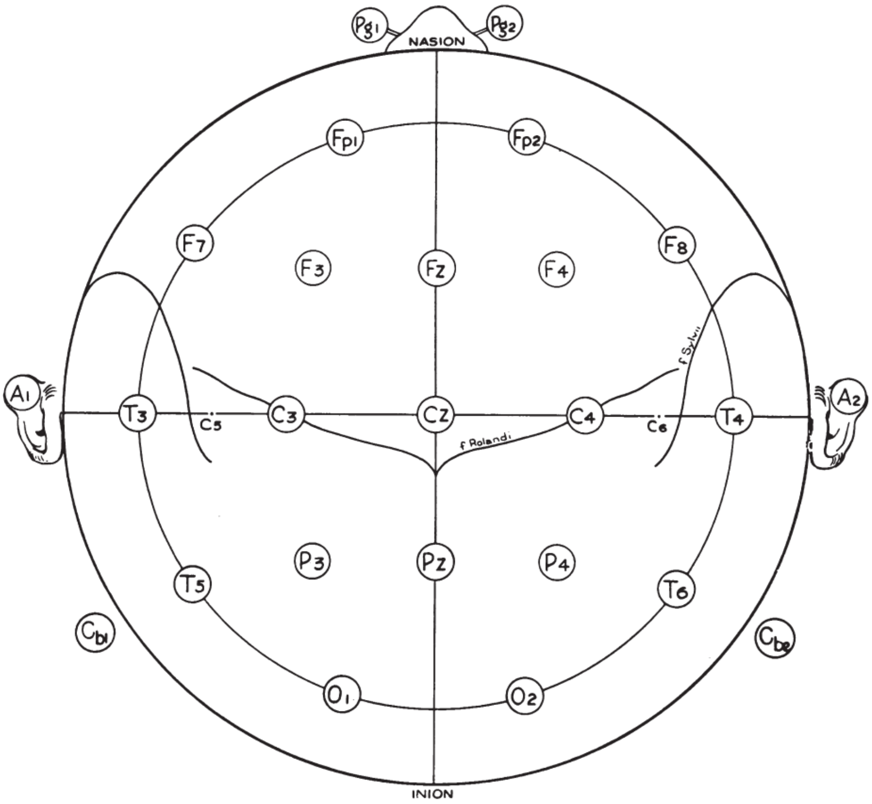


Figure 2.3 Single plane projection of the head, showing the standard electrodes positions *[24]*

2.2 Neurofeedback

Neurofeedback is part of a wider group of biofeedback applications, all of which have the goal of facilitating the self-regulation of physiological functions with the goal of normalising them in clinical populations or optimising them in healthy subjects. Biofeedback is an operant conditioning procedure in which participants learn to gain self-control over physiological functions (e.g., muscle activity, respiration, heart rate) that usually are not consciously perceived or controlled [25]. Operant conditioning is a method of learning that occurs through rewards and punishments for behaviour. Through operant conditioning, an individual makes an association between a given behaviour and a consequence: positive consequences increase the likelihood of the behaviour, whereas negative consequences decrease it [6].

In the 1960's Joseph Kamiya, today considered the father of neurofeedback, was the first to verify whether operant conditioning methods could be used to induce direct changes in the EEG [26]. He conducted experiments in order to investigate if subjects had the ability to distinguish, in a subjective way, which kind of waves were being generated by their brain. In these first studies subjects were asked to keep their eyes closed and periodically prompted to report whether they were producing dominant alpha waves or not. Participants were also told whether they were responding correctly, and they exhibited an increasing ability to associate subjective experience with the presence of alpha EEG oscillations. They also demonstrated their ability to produce alpha oscillations on demand, effectively bringing EEG parameters under operant control.

Kamiya was the first researcher to demonstrate human's ability of controlling one's own alpha waves. Since then, many studies have been conducted that confirm the effectiveness of neurofeedback in self-control of the brain activity. Researchers developed several protocols, which entail the upregulation or suppression of the amplitude of specific brain waves.

This ability of consciously controlling brain activity through neurofeedback is of great importance and can be used in at least two ways [6]:

- as a therapeutic tool to normalize neurological patients' deviating brain activity in order to influence symptoms;
- as so-called peak-performance training to enhance cognitive performance in healthy participants.

2.2.1 Clinical use of NF

As previously described in *section 2.1*, in the EEG signal we can identify several frequency components. Each brain wave is predominant in certain mental states and, in neurologically healthy population, all of them occur simultaneously in a certain balanced proportion [21]. However, by comparing measures from healthy reference populations, neuroscientist have found different dysregulated EEG patterns or abnormalities associated with mental disorders.

For instance, slower waves (e.g., theta 4–8 Hz) are reported to be globally elevated in Attentional Deficit Hyperactivity Disorder (ADHD). Similarly, obsessive-compulsive disorder (OCD) patients demonstrate low-frequency power excess (2–6 Hz) in the resting state [21]. Another example is post-traumatic stress disorder (PTSD), which is observed to have decreased power of the alpha rhythm, potentially reflecting cortical. In contrast, disorders such as substance use disorders (SUD) and schizophrenia are characterised by synchronization deficits of the faster beta and gamma rhythms [21].

In this context, neurofeedback is used as a therapeutic tool to normalize patients' deviating brain activity in order to influence symptoms. Research on clinical application of neurofeedback recognized it as a valid alternative to drugs or behavioural therapy [22]. Based on the specific application, different protocols are used trying to compensate for the EEG dysregulation.

For instance, in ADHD disorder treatment the goal is to decrease the brain activity in the theta band and to increase its activity in the beta band (i.e., to decrease theta/beta ratio) [22]. Several studies proved the efficacy of such a protocol on ADHD patients, with improvements in attention, focus and memory. PTSD patients showed improvement in symptoms when treated with neurofeedback protocols designed to regulate alpha or both alpha (upregulation) and theta (suppression) waves [27]. And also in the case of schizophrenia, positive outcomes were reached by applying NF protocols establishing the increase of beta and decrease delta and theta activity [28].

The literature on this topic is vast. The abovementioned cases are only a few examples of the applications of NF in clinical context, but we could cite numerous others, like depression, anxiety, epilepsy, autistic spectrum disorder, drugs abuse, learning disabilities [21].

NF clinical applications are not our main focus; with these examples, we just wanted to give the reader an idea of NF potentiality in the treatment of brain disorders.

2.2.2 Peak-performance NF

When it comes to the use of neurofeedback in healthy populations, the goal is to optimise cognitive, and sometimes also artistic, performance. The rationale for undertaking NFT to improve performance is based on research showing that brain waves are correlated with cognitive functions. Most researchers had focused on some main neurofeedback protocols, i.e., theta, beta, gamma and alpha training.

Several studies have demonstrated a close relationship between theta band synchronization (i.e., power increase) and good cognitive performance. For example, intense theta activity has been observed while performing a continuous attention task or a task with high cognitive demands. Theta band activity has been then associated with focused attention, concentration and creativity as well. Moreover, theta band power increase has been associated with facilitating episodic memory and encoding new information, thus learning processes [29].

Regarding beta training, Egner & Gruzelier [30] have reported that low beta conditioning resulted in better results and improved perceptual sensitivity in a continuous performance task. Moreover, uptraining of beta band components is shown to affect different aspects of attentional processing, resulting in faster reaction time and reduced omission errors [31].

About gamma training, power increase in the gamma band has been observed when subjects are performing attention-related tests [32]. It has also been demonstrated that gamma enhancing [33] neurofeedback could lead to increased flexibility in handling episodic bindings (i.e., bindings between two features of visual objects, such as their shape and location), which suggests a role of gamma band activity in control of memory retrieval. Moreover, there is evidence that gamma activity is important for controlling and organizing memory traces in both short-term binding and long-term memory [33].

As regards alpha band, it has been linked to different cognitive abilities. Oscillations in the alpha band increase with memory load [34], suggesting that alpha waves are correlated with memory performance. Hanslmayr et al. [35] showed that participants who were capable of learning to increase their alpha power performed better on a mental rotation task. Alpha NFT was also used to obtain musical performance enhancement [36]. Moreover, it has been suggested that the alpha frequency range should be separated into lower alpha (8-10 Hz) and upper alpha (10–12 Hz), based on the findings showing that the lower alpha band is predominantly associated with attentional processes, whereas upper alpha is primarily associated with memory processes [37].

For the scope of this thesis we focused in particular on this relationship between upper alpha band and memory performance. Hence, in the following section we give a brief literature review of the most recent studies addressing this topic.

2.2.2.1 Upper Alpha NF and memory performance

Different studies proved the hypothesis that neurofeedback training in the upper alpha sub band (10-12 Hz) can lead to memory performance enhancement.

Hanslmayr and colleagues [35] conducted an experiment in which neurofeedback training was applied in order to increase upper alpha power. A mental rotation task was performed before and after upper alpha NFT. They observed that only those subjects who were able to increase their upper alpha power performed better on mental rotations after NFT, showing that training success (extent of NFT-induced increase in upper alpha power) was positively correlated with the improvement in cognitive performance.

Similarly, the impact of upper alpha NF training on cognitive abilities was assessed by Zoefel et al. [17]. Participants in the study were divided into two groups. One group underwent five NFT sessions with the aim of increasing the amplitude of upper alpha brain waves, while the other (control group) did not receive NF treatment. A mental rotation test was given before and after the NFT period to test cognitive abilities. The expectation of an enhancement of cognitive performance was confirmed: the NFT group obtained an increase in the upper alpha activity and the increase in scores of mental rotation was significantly larger for the NFT group than for the control group.

Since mental rotation is an ability that involves working memory (WM) [38], these results suggested that upper alpha NFT has a positive effect on working memory.

Working memory refers to the ability of the brain to provide temporary storage and manipulation of information, necessary for complex cognitive task as language comprehension, learning and reasoning. The definition of WM evolved from the concept of short-term memory and it is often confused with it. The difference lies in the fact that working memory requires the simultaneous maintenance and manipulation of information, while short-term memory refers to the temporary storage of information only, without the attention component of working memory. Although they are conceptually different, the use of the terms short-term memory and working memory in literature is not always strict and there is evidence for a large or even complete overlap between the two constructs [39].

The specific effect of upper alpha NFT on working memory was further investigated by Escolano et al. (2011) [40].

Their experiment consisted of five NF sessions, during which participants learnt to increase their upper alpha amplitude. Participants in the study were divided into a NFT group and a control group. Only the NF group was exposed to NF training. Working memory was tested for every subject before and after the NFT period using a conceptual span test. Results show that participants in the NF group obtained an increase in the upper alpha activity as well as a significant enhancement in memory performance compared to the control group. Hence, they confirmed the hypothesis that an increase of upper alpha activity is associated with an enhancement of the working memory performance.

In 2012, Nan and colleagues [41] proposed the use of alpha neurofeedback to improve short-term memory performance. In this case, the NF protocol established the training of brain activity in the whole alpha (not only upper alpha) band. Short-term memory was evaluated by a digit span test. The experimental results showed that the participants were able to learn to increase the amplitude in the alpha band during 20 sessions of NFT and short-term memory performance was significantly enhanced by NF training. More importantly, further analysis revealed that the improvement of short-term memory was positively correlated with the increase of the amplitude only in the upper alpha sub band.

Hsueh and colleagues (2016) [42] explored the effects of alpha NFT on memory performance as well. Participants in their study received 12 sessions of alpha neurofeedback. Working memory was assessed by both a backward digit span task and an operation span task, and episodic memory was assessed using a word pair task. Subjects showed a progressive significant increase in the alpha amplitude. Accuracies of both working and episodic memories were significantly improved in a large proportion of participants, particularly for those with remarkable alpha amplitude increases. In this case the NF training was not limited to the upper alpha sub band, but on the whole alpha band. However, results are consistent with those of the other studies.

In all the studies mentioned above, the NF treatment was applied to healthy subjects. Another study, by Kober et al. (2015) [43], verified if the same results could be obtained with stroke patients.

They investigated the ability of stroke patients to control their own brain activity via NF and evaluated specific effects of different NF protocols, including upper alpha training, on cognition, in particular recovery of memory. To evaluate the NF training outcome, a test battery assessing different cognitive functions (attention, executive functions, long-term memory, working memory) was performed before and after NF training. According to the results, Upper Alpha patients specifically improved their working memory performance.

All these studies provide promising results and encourage the use of upper alpha neurofeedback training for working memory enhancement.

2.2.3 Neurofeedback loop

The neurofeedback training procedure consists in the following phases [6].

(1) Data acquisition

The first step in a neurofeedback training procedure is the data acquisition of brain signals by means of EEG. Electrodes are placed on the scalp following the 10-20 system. Their location is established based on the specific application, considering the association between brain lobes and their functions and the scalp distribution of the target frequency band. Usually, only one or two active electrodes are used.

(2) Online data processing

Then, EEG data is analysed in real time to select and extract features that are used during neurofeedback. These features represent that pattern of brain activity that one wants to modulate [6], and the feedback given to the participant will be based on the data elaborated in this phase. The feedback parameter is usually the amplitude of a specific brain wave (e.g., alpha amplitude) or the ratio between the amplitudes of brain activity in two frequency bands (e.g., alpha/theta ratio). Hence, this step may simply translate to the selection of specific frequency bands of the EEG. Today, the most popular method of translating EEG into the frequency domain is by applying a Fast Fourier Transform (FFT). The FFT is superior to other frequency analysis methods because it can transform data as EEG is being recorded, which is necessary for real-time applications such as neurofeedback [44].

(3) Feedback signal generation

The chosen feedback parameter is then translated into a sensory (visual and/or auditory) stimulus that is presented and processed by the learner. The feedback is performed in real time, so it reflects the brain activity with a minimum constant time delay, a delay which is usually kept under one second. The type of feedback varies among studies. Auditory feedback might be a sound (e.g., the sound of a waterfall, birds singing), a tone or a melody changing its pitch [45] or volume [46] in dependence on the amplitude of the trained EEG frequency. Visual feedback often uses two-dimensional (2D) moving objects such as bars or circles changing their size or colour in dependence on the brain activity level [43, 15].

(4) Learning

Whatever the type of feedback is, it serves to give the subject information on their own brain activity. The participant is instructed on the use of the feedback and they know which aspect is associated with appropriate changes in the brain activity. For instance, they are told to increase the size of a bar in order to increase the brain wave amplitude. Being continuously exposed to the feedback, the learner will identify by trial and error the correct behaviour or strategy to alter the brain activity in the desired direction.

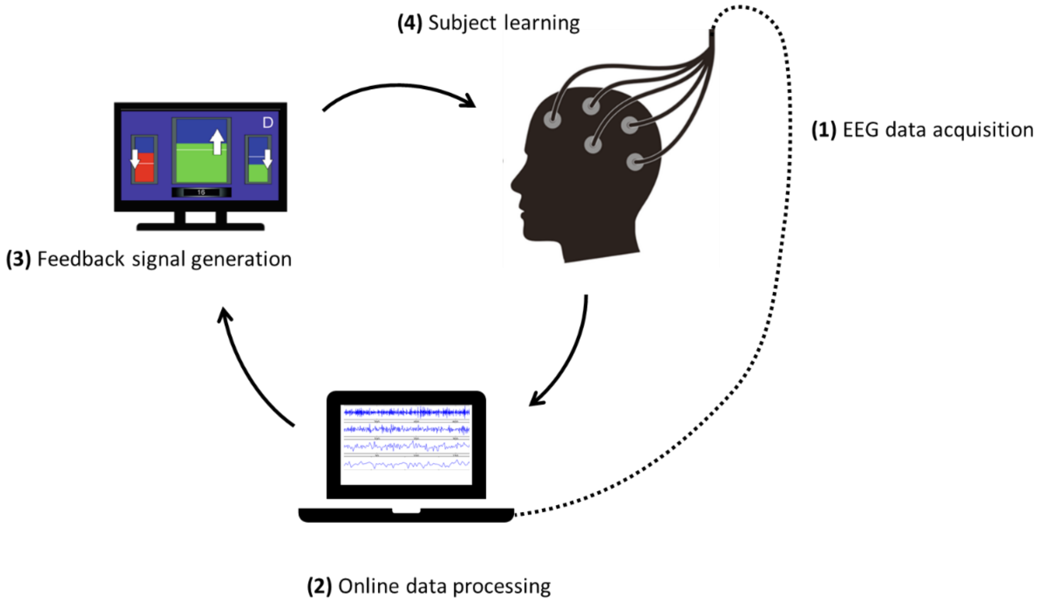


Figure 2.4 Feedback loop

2.2.4 Neurofeedback training efficacy

Undoubtedly, controlling brain activity is an ability that can be learnt. There is ample literature from the last fifty years providing evidence of the effectiveness of neurofeedback. However, individuals differ in their ability to learn how to regulate the brain activity by neurofeedback. Little is known of how these individual differences arise and what enables one person to learn better or faster than the other. These differences may exist in internal and external factors.

Learner internal characteristics that determine the success of neurofeedback training have become the focus of attention recently [6]. Learner specific aspects such as positive mood states [47], motivation [14, 48], locus of control [49], all turned out as being relevant for the prediction of individual learning success. Evidence also suggests that the morphology of brain areas generating EEG features used for neurofeedback training may be associated with training success [50].

Variability in external factors can be found by comparing the design of training protocols between studies. To date, there is no consensus on the parameters that should lead to an effective NF protocol [51]. The duration of sessions applied in different studies can vary within a range of 30 to 60 minutes. The number of sessions can differ from 5 [17, 40] to more than 40 [52]. Spacing of sessions over time also differs, but most studies involve two or three sessions a week. Even training frequency bands vary in width and range amongst studies. Sometimes several frequencies are trained simultaneously, as alpha enhancement paired with theta inhibition training, while other researchers argue that training a single frequency is more effective. Furthermore, researchers can employ a variety of forms of feedback, some using visual feedback such as dynamic shapes and others use auditory feedback or a combination of both.

All the above-mentioned aspects may affect the efficacy of the training. There is increasing awareness that the effects of changing such parameters should be explored further, in order to define an effective NF protocol. In particular, researchers recently started to focus on the effects that feedback design can have on NF training.

Traditional feedback modalities, often using two-dimensional objects, can be relatively monotonous and not encourage users to focus on them. Since mood, motivation and interest are relevant aspects for successful NF learning, it is important the feedback to be engaging and attractive. For this reason, an increasing number of recent NF studies use Virtual Reality based feedback designs [10, 53], showing that VR is more effective than traditional modalities. We will describe the results of these studies in the next section.

2.3 Immersive Virtual Reality

Virtual Reality is defined as “a medium composed of interactive computer simulations that sense the participant's position and actions and replace or augment the feedback to one or more senses, giving the feeling of being mentally immersed or present in the simulation (a virtual world)” [54].

A concept frequently mentioned in VR is “immersion”, that is the perception of being physically present in a non-physical world. This perception is created by means of images, sounds or other stimuli that surround the user, providing a very absorbing environment. A VR system is immersive when the simulated world is perceptually convincing, it looks authentic and real, and the user has the feeling of “being there" [55].

Even if immersion seems to be a crucial element, VR can also be non-immersive when it “places the user in a 3D environment that can be directly manipulated, but it does so with a conventional graphics workstation using a monitor, a keyboard, and a mouse” [56]. This is referred to as desktop VR.

In the next sections we give a detailed definition of immersion and present a literature review on the use of immersive VR in Neurofeedback practice.

2.3.1 Immersion

Immersion is the measurable feature of Virtual Reality technology that could make a user feel present in a virtual environment. Slater & Wilbur [16] have laid out a series of definitions for immersion that will be used in this work.

Immersion is what a technology delivers from an objective measure and describes the extent to which users can feel part of the environment. The more a system conveys displays that preserve fidelity in relation to their corresponding real-world sensory modalities, the more it is immersive. Immersion refers to what is, in principle, a quantifiable description of a technology and it can be objectively assessed based on capabilities of the hardware and software being used. However, there are no widely accepted methods for objectively quantifying immersion.

Several factors influence the level of immersion. Specifically, Slater & Wilbur identified five primary factors of immersion. According to them, it “includes the extent to which the computer displays are extensive, surrounding, inclusive, vivid and matching” [16].

> The displays are more extensive the more sensory systems that they accommodate.
>
> They are surrounding to the extent that information can arrive at the person's sense organs from any (virtual) direction, and the participant can turn towards that direction receiving the appropriate directional sensory signals. The notion of surrounding also includes the greater the reproduction of the natural modes of sensory presentation (visual and auditory stereopsis for example).
>
> They are inclusive to the extent that all external sensory data (from physical reality) is shut out. In the ideal situation, for example, an inclusive system would have an HMD completely weightless, so that this aspect of external reality is not perceived by the participant.
>
> Their vividness is a function of the variety and richness of the sensory information they can generate. Vividness is concerned with the richness, information content, resolution and quality of the displays.
>
> Finally, immersion requires that there is match between the participant's proprioceptive feedback about body movements, and the information generated on the displays. A turn of the head should result in a corresponding change to the visual display, and, for example, to the auditory displays so that perceived sound direction is invariant to the orientation of the head. Matching requires body tracking, at least head tracking, but generally the greater the degree of body mapping, the greater the extent to which the movements of the body can be accurately reproduced.

Vividness

Vividness is related to the resolution, photo-realism and visual fidelity of the virtual scenario [16].

We are particularly interested in studying vividness because of its heavy reliance on visual stimuli. Since virtual environments are graphical interfaces, humans heavily rely on their visual sensory system to perceive their surroundings. Hence, modifications to the scene vividness should result in significant effects.

Additionally, from an experimental design perspective, researchers can quickly manipulate and study different levels of vividness.

For example, Slater & Wilbur [16] use shadows as an example of vividness manipulation. It has been shown that scenes where shadows and reflections are present are perceived as more realistic [57].

Wang & Doube [58] considered image roughness and shadow softness as perceivable characteristics of realism. It has been shown that images appear more realistic when the surfaces of their objects are perceived to be rough. Conversely, they appear less realistic when the surfaces of their objects appear smooth. Moreover, images in which objects project hard shadows under the illumination of strong, directional light are perceived as less 'real' than images in which soft shadows are projected under normal diffused illumination.

Toczek [59] used a texture resolution approach, populating high and low vividness conditions with objects of varying pixel resolution.

2.3.2 Neurofeedback and VR

Immersive virtual environments are supposed to be more effective with respect to the acquisition of several abilities and have a positive impact on human performance, compared to other digital approaches. This is because the brain recognizes the virtual world as real and this facilitates the transfer of the learned skills to the real world. For this reason, NF researchers started to investigate the hypothesis that virtual reality feedback causes an improvement in NF learning performance, compared to traditional feedback modalities.

Results from Berger and colleagues (2017) [60] seem to confirm this hypothesis. They used neurofeedback to train subjects to increase their level of alpha amplitude, providing half of the participants with feedback in a 3-dimensional (3D) virtual reality environment, while the other half received feedback in a 2D environment. Both groups visualized the feedback through an HMD. Participants in the 2D group were watching a simple cinema screen with a square floating when the alpha amplitude was above a threshold, while those in the 3D group were placed in the middle of a virtual room. They could look around in the virtual room, and the feedback consisted of a vase floating when the alpha amplitude exceeded a threshold. After five neurofeedback training sessions, they found out that learning slopes were higher in participants who receive feedback in the 3D virtual environment, while the training of the 2D group was unsuccessful. Researchers hypothesized that the 2D environment itself hindered learning: compared to 3D, it was monotonous and dark, and this could increase boredom. Hence, the study corroborates the 3D advantages, highlighting the importance of immersion and engagement.

In a similar study, Kober et al. [15] compared two different types of feedback. In the 2D condition, they used vertically moving bars changing size visualized on a conventional computer screen. In the 3D condition, the feedback was a virtual 3D stereoscopic rendering of a human body, in which the organs changed in appearance depending on NFT results. To provide the 3D feedback they used a stereoscopic screen and stereoscopic glasses. Different neurofeedback protocols were applied (increase of SMR or upper alpha activity, decrease of theta/beta ratio). In this case, NFT performance was comparable in 2D and 3D conditions: participants in both groups were able to modulate the trained EEG frequencies in the desired direction and 3D VR feedback did not improve the NF training performance compared to a traditional feedback modality. However, researchers also investigated the effects on user experience, assessing that interest, motivation and challenge were higher in the 3D group. This suggests that VR applications have the potential to attract and motivate users more than classical 2D feedback screens.

In another study by Gruzelier et al. [61], participants learned through neurofeedback to control their SMR activity. For the feedback, researchers recreated a theatre auditorium in a virtual environment. The lighting level and the audience noise in the VE changed according to the EEG activity. Two levels of immersion were examined. In one the auditorium was rendered on a conventional computer screen. This was compared with a CAVE-like system, a more immersive medium, where the seated participant was surrounded by the same theatre auditorium projected seamlessly on the surrounding walls. EEG analysis revealed that the presence enhancing properties of the more immersive CAVE-like system context had benefits: NF learning was facilitated (participants learned faster) in the CAVE rendition of the theatrical space vs. the computer screen, even though the same auditorium was depicted.

Cho and colleagues [62] obtained consistent results as well. In their study participants were trained to reinforce their beta activity. As feedback, they visualized a virtual classroom and they earned scores as positive reinforcement when the EEG signal was greater than a threshold. A group of participants visualized the VE on a computer monitor, while the other used an HMD and was able to look around in the virtual classroom. Both groups learnt to reinforce beta activity, but there was a tendency for the HMD group to obtain better learning results.

The reviewed studies make comparisons between different types of feedback on different plans, sometimes comparing the same virtual reality content in different settings (e.g., screen vs. CAVE, or screen vs. HMD), sometimes comparing VR contents with traditional non-VR feedback. Even if the VR modalities used in these studies are heterogeneous, we can notice that in every comparison the feedback resulting more effective was the more immersive one: being immersed in a virtual room was better than looking at 2D objects on a screen; visualizing virtual contents with an HMD or in a CAVE was better than through a computer screen.

The overall conclusion we can deduce from these results is that the immersive properties of virtual reality bring advantages in neurofeedback training, either in facilitating NF learning or increasing motivation and interest. However, as previously described, immersion is defined by many factors and we cannot find in literature studies that examined immersion at the level of the core principles which define it.

3 Problem statement

A well-known problem in the neurofeedback community is that there exists a large variability in people's ability to acquire a certain degree of control over their brain activity. This variability can reflect on the speed of learning. Moreover, not all NF users have shown satisfactory learning ability to regulate their brain oscillations.

Given the wide range of NF applications (section *2.2*), trying to understand what influence and how to enhance NF learning performance is of primary importance.

The efficacy of NF training can be affected by numerous factors we cited in section *2.2.4*. One of these is the feedback modality. To date, there is not exhaustive literature investigating the impact of the feedback design on NF learning performance. Few studies (see literature review in section *2.3.2*) provide evidence that when using immersive virtual reality, NF training is more effective. The immersive properties of virtual reality seem to have a positive effect on NF practice: users learn better or faster to control brain activity, and they feel more interested and motivated. However, the precise mechanism by which an immersive virtual environment facilitates NF learning is yet unclear.

We think that, in order to understand what makes an immersive virtual environment a good feedback modality and how to choose the best design for virtual environments, there is the need to systematically examine immersion-related variables and their effects on training outcomes. However, there are no studies in literature exploring this issue.

In this study we want to contribute to research on this topic, focusing on one of the immersion’s dimensions: vividness. We want to assess the effects that a more vivid training scenario has on neurofeedback learning performance as well as on subjective user experience. We focused on psychological factors such as such as motivation, concentration, stress, boredom, and feeling of control that can influence, in turn, NF learning performance. Our final aim is to understand whether a more vivid feedback brings advantages to neurofeedback practice.

The primary research questions addressed are:

- *Does a more vivid feedback lead to better NF learning performance?*
- *How does vividness affect subjective variables, such as motivation, concentration, stress, boredom, and feeling of control?*

Our hypothesis is that a more vivid (thus immersive) training scenario, being more engaging and attractive, makes NF users feel more interested and focused on the NF training. Conversely, we would expect users in a lower vividness training scenario to feel more bored and frustrated. Consequently, we hypothesize that NF users in a highly vivid scenario attain higher NF learning performance.

In the choice of the NF protocol, we followed the principle of interpretability [17]. This means that we chose a frequency band that is associated with specific cognitive functions to increase the probability of reliable behavioural effects as well as applicability. We decided to adopt an upper alpha enhancement protocol. According to literature (see literature review section *2.2.2.1*), upper alpha is correlated with memory performance. Thus, the second objective of this study is to assess the effect of upper alpha neurofeedback training on working memory performance.

The secondary research question is:

- *Is the upper alpha activity increase, obtained through NF, correlated with working memory improvement?*

We would expect that NF users, after the NF training, show improved working memory performance and that this improvement is correlated with the UA enhancement.

4 Methodology

In order to investigate our research questions, we designed a NF-VR task where participants underwent five neurofeedback training sessions. We adopted an upper alpha (10-12 Hz) enhancement neurofeedback protocol, in which subjects learned to increase their relative upper alpha amplitude. The feedback was given to participants through an immersive Cave automatic virtual environment (CAVE). In *section 4.3*, we describe all the details related to the neurofeedback procedure.

We wanted to test the effects of vividness on NF learning and user experience, so we implemented the neurofeedback software and created three virtual environments differing in the level of vividness (*section 4.2*). Participants were divided into three experimental groups, each receiving feedback in a different VE. After every session, participants were asked to report through questionnaires some user experience variables (*section 4.4*).

In order to assess the effect of upper alpha enhancement on working memory, before the start and after the end of the neurofeedback training period, we measured working memory performance through three tests (*section 4.4*).

Moreover, after the five NF training sessions, participants underwent an additional NF Transfer session. In this session, all participants received feedback with the same modality, with no distinction between groups. This served to assess if the ability to control upper alpha, acquired during the NF training, could generalize to other visual stimuli.

4.1 Participants

Twenty-one participants (6 female and 15 male), ranging in age from 20 to 42 years old (M = 28, SD = 5.2), took part in the experiment. Participants were recruited based on their motivation to participate among students and staff at the Madeira Interactive Technologies Institute (M-ITI), Funchal, Portugal. Inclusion criteria for participation in the study included the following: (i) be over 18 years old; (ii) can understand English; (iii) have no past of brain injuries and no neurological disorders. Finally, all participants were informed and signed an informed consent to participate and to publish their data.

Participants were quasi-randomly (by order of enrolment in the study) assigned to the three experimental groups:

(i) group A – Low vividness;
(ii) group B – Medium vividness;
(iii) group C – High vividness.

Each group consisted of 7 subjects (2 female and 5 male) and there was no significant difference in age between groups.

4.2 Independent variable

Vividness

Our prime objective was to assess the effects (in terms of NF learning performance and user experience) of using different types of Virtual Reality feedback in a NF training practice. More precisely, we wanted to test virtual environments changing in the level of vividness.

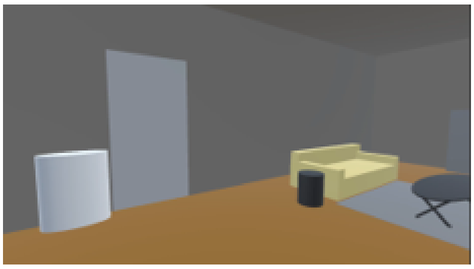
This experiment used three levels of vividness, those being low, medium and high.

Vividness is associated with the resolution and fidelity simulated within a particular modality. High vividness scenarios were designed to be the "most realistic" while the low vividness scenarios were designed to be the "least realistic". These differences were made evident by changing the geometric complexity of the elements in the environment and by the use of textures, shadows, and reflections (*Table 4.1*). As introduced in *section 2.4*, the presence of shadows and reflections in a virtual scene makes it seem more realistic [57]. Moreover, images appear less realistic when the surfaces of their objects are perceived to be smooth [58] while applying textures to the objects can increase the photo-realism.

Table 4.1 Vividness levels classification

	Low vividness	**Medium vividness**	**High vividness**
Geometric complexity	Low geometric complexity	Higher geometric complexity	High geometric complexity
Textures	Smooth surfaces	Limited textures	High-resolution textures
Shadows/reflections	No shadows/reflections	No shadows/reflections	Soft shadows/reflections

The virtual environments were developed using Unity (Unity Technologies, San Francisco, CA), a cross-platform game engine. We reproduced three versions of the same living room at different levels of vividness. In the low vividness level, we used simple geometric shapes (i.e., cube, cylinder, sphere) to reproduce objects. Each additional vividness level was created incrementally from the previous one, by implementing new details, modifying textures and shadows and using more elaborate 3D models[1].



(a) Scenario A – Low vividness:
Low geometric complexity
Smooth surfaces
No shadows/reflections

(b) Scenario B – Medium vividness:
Higher geometric complexity
Limited object textures
No shadows/reflections

(c) Scenario C – High vividness:
High geometric complexity
High resolution textures
Soft shadows/reflections

Figure 4.1 A view of the virtual environments differing in level of vividness

---

[1] We used 3D models available in Unity Asset Store
[https://assetstore.unity.com/packages/3d/environments/archvizpro-interior-vol-1-41721]

4.3 Experimental procedure

Participants received five neurofeedback training session on consecutive days (except weekend days), from Day 1 to Day 5.

On Day 1, before the start of the NF training, participants signed an informed consent form (*Appendix A*) and provided some basic demographic information (i.e., age, gender, *Appendix B*). Then they did three working memory tests (Pre-tests): a digit span test and N-back tests (in the 2-back and 3-back versions).

After that, they started the neurofeedback session (as described in section *4.3.5*). The same NF procedure was repeated from Day 2 to 5, and after every session, the participant filled out a set of questionnaires to assess some subjective user variables.

On Day 5, after the end of the NF session, they did an additional NF session (Transfer session) and they repeated the same working memory tests performed on Day 1 (Post-tests). The transfer session consisted in the same NF training of the previous sessions, but with a different type of feedback.

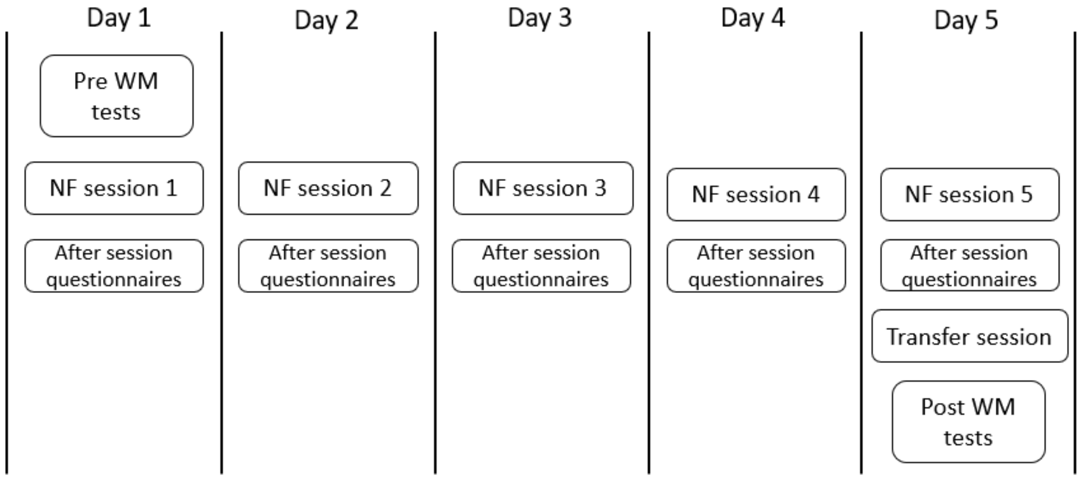


Figure 4.2 Overview of the experimental procedure

4.3.1 Neurofeedback protocol

We adopted an upper alpha enhancement protocol, with the objective of increasing the amplitude of the brain activity in the upper alpha frequency band (10-12 Hz).

Absolute EEG amplitude has large individual difference owing to influences of many factors (such as anatomical and neurophysiological properties of the brain, cranial bone structure, and electrode impedances) [63]. Furthermore, additional confounding factors across sessions could result from changes in time of day [64], mood or spontaneous cognitive activity [65].

Simple ratios between EEG band amplitudes are commonly used in neurofeedback protocols as relative measures are less sensitive to differences in these uncontrolled factors that modulate absolute EEG amplitudes [63].

Hence, in order to ensure comparability across participants and sessions, we used the upper alpha relative amplitude as feedback parameter.

The upper alpha relative amplitude was defined to the analysed frequency band (upper alpha: 10-12 Hz) amplitude relative to the EEG band amplitude from 4 to 30 Hz [41, 12]. For brevity, we will refer to the UA relative amplitude as UA ratio.

$$UA\ relative\ amplitude = \frac{UA\ (10-12\ Hz)\ amplitude}{EEG\ (4-30\ Hz)\ amplitude}$$

4.3.2 EEG data acquisition

For EEG acquisition, the Enobio 8 (Neuroelectrics, Barcelona, Spain) system was used. Enobio is a wearable, wireless EEG sensor with 8 EEG channels, provided with a battery-operated device that connects through Bluetooth to the Neuroelectrics Instrument Controller (NIC) software running on a computer. It records 24-bit EEG data at 500 Hz.

Eight dry electrodes were placed following the 10-20 system in the locations *F3, F4, C3, Cz, C4, Pz, O1, O2* (as shown in *Figure 4.3*) and the reference channels were placed at the left mastoid. A neoprene cap held the sensors in place (*Figure 4.4*).

Although signals were recorded from all the eight channels, only data from the *Cz* channel was used for the feedback and analysis.

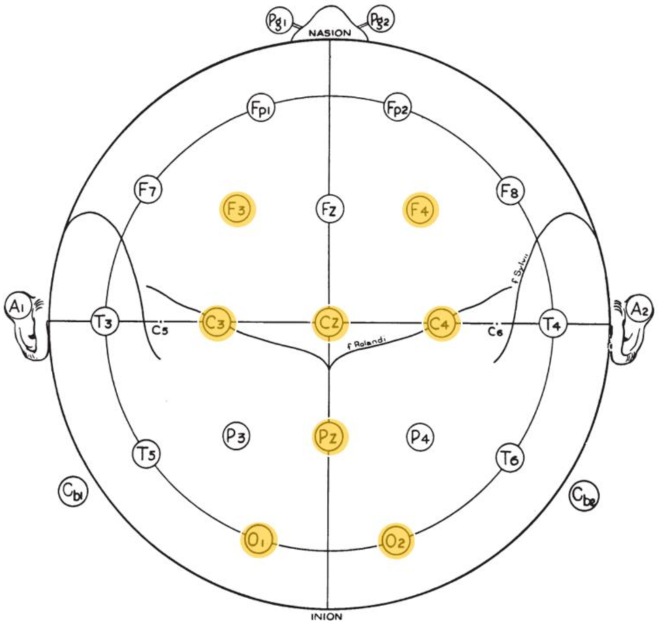


Figure 4.3 Electrodes configuration used for the experiment, based on 10-20 system

Figure 4.4 On the left, the Enobio8 system. It has 8 EEG channels and 2 reference channels. On the right, a participant wearing the neoprene cup with the Enobio device attached. The headphones were used for sound isolation during the NF training session.

4.3.3 EEG data processing

Enobio was connected via Bluetooth to a desktop computer for the EEG data processing. Data was recorded and processed through OpenVibe, an open source software platform for designing BCI experiments.

The OpenVibe acquisition server was used to connect to and get data from the EEG cap. The acquisition server sent data to the OpenVibe designer, where we design the EEG signal processing scenario.

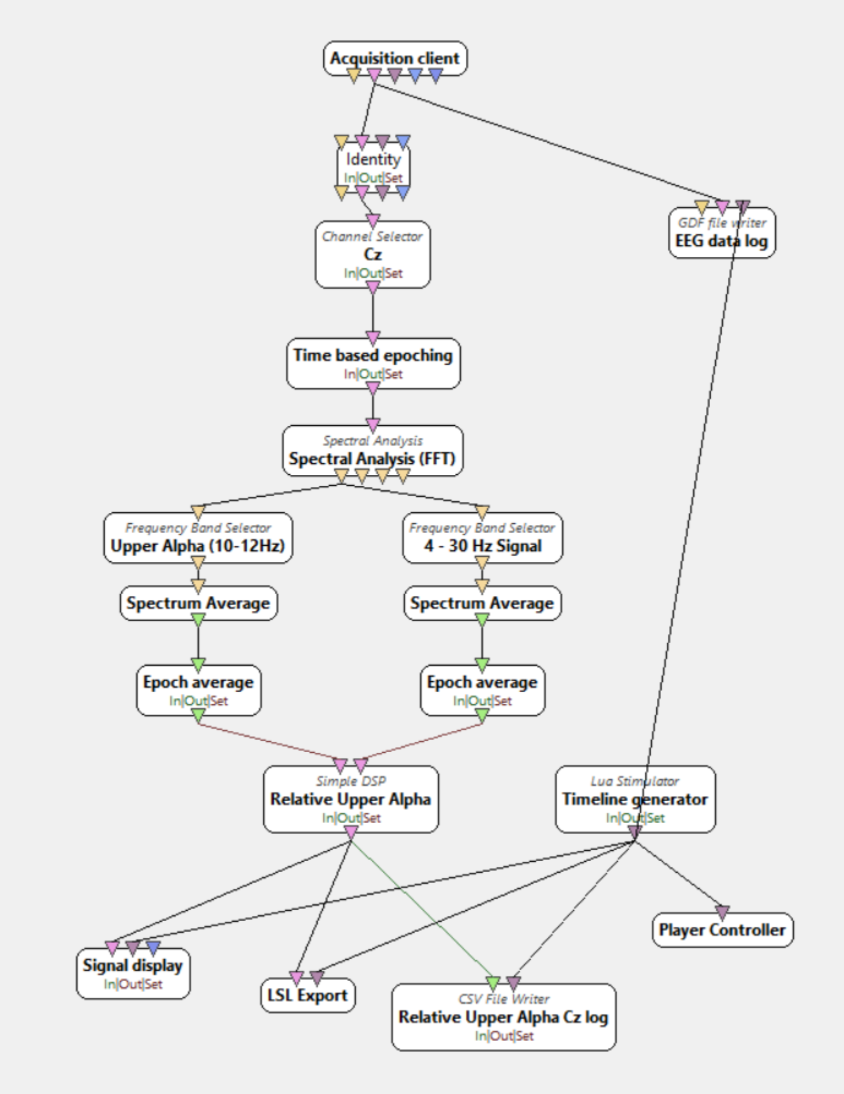


Figure 4.5 OpenVibe scenario for EEG data processing.

Data is received from the EEG device. Only data from Cz channel is selected. Amplitude of the upper alpha band and 4-30 Hz range is computed though FFT and used to compute the relative upper alpha amplitude (ratio). The relative UA amplitude is sent via LSL for feedback generation. Log files are saved for offline analysis.

With the OpenVibe scenario, we computed the upper alpha ratio, as described in the NF protocol. The amplitude was calculated by Fast Fourier Transforms, using a sliding window of 2 s, shifting every 0.125 s. The upper alpha signal was then sent in the network through the Lab Streaming Layer (LSL) protocol [66]. LSL is a system used for the real-time streaming of time series in research experiments. In this case, it was used to send processed data from OpenVibe to the Unity application used for the visualization of the feedback.

4.3.4 Providing feedback

Since we wanted to test different levels of vividness in virtual environments and vividness is an immersion-related variable, we needed an immersive virtual reality setting to deliver the feedback.

When using virtual reality for neurofeedback, it is very common the use of HMD for the feedback visualization [60, 62]. We opted for using the NeuroRehabLab CAVE at M-ITI (Madeira Interactive Technologies Institute), where the study was carried out. A CAVE is an immersive VR environment consisting of a cube-shaped VR room in which images are displayed by projection on the walls. Depending on the configuration, between three and six of the walls (including floor and ceiling) can be projection screens.

The NeuroRehabLab CAVE is a system developed by Gonçalves et al. [67]. It has a configuration of three orthogonal walls and a floor. It uses a Kinect sensor for tracking, thus enabling motion parallax effects and body interaction. The KAVE plugin [67] for Unity has been developed for the integration of Unity applications with the CAVE.

Figure 4.6 NeuroRehabLab CAVE

The neurofeedback application for the feedback visualization was then implemented using Unity and the KAVE plugin. The virtual environments created were described in section *4.2*. A series of C# scripts were developed and utilized in Unity to receive the EEG data from OpenVibe and control the visualization of the feedback.

More precisely, the feedback consisted in an object changing colour. We chose this type of feedback because it is often used in literature for upper alpha neurofeedback training [40, 35, 17]. In the low vividness environment, the object was a cylinder while in the other two environments it was the light from a lamp.

The colour scheme ranged from a highly saturated red to a highly saturated blue. The colour changed according to the upper alpha ratio: red and blue values symbolized an UA ratio above and below the THRESHOLD value, respectively; the full saturated red corresponded to an UA ratio greater than or equal to the MAX value; the full saturated blue corresponded to an UA ratio less than or equal to the MIN value; closer the UA ratio was to the THRESHOLD, whiter the colour became.

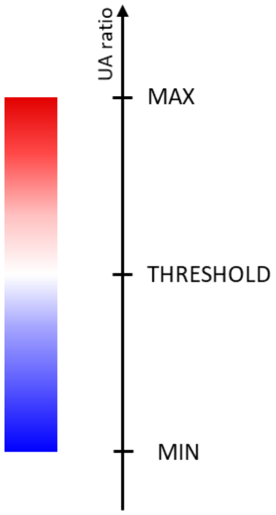


Figure 4.7 Colour scheme.

The colour changed from blue to red according to the UA ratio. A highly saturated red corresponded to a high UA relative amplitude. Participants' task was to make the colour as red as possible, in order to increase their UA relative amplitude.

Taking into account the differences that could occur in UA relative amplitude between sessions and between subjects, and in order to provide better feedback to participants, THRESHOLD, MAX and MIN values were not static. They were computed in every session, based on the UA relative amplitude recorded with eyes open at rest prior to each training session: THRESHOLD value was set to the median value of the UA relative amplitude of the baseline recording; MAX value was set to the 95th percentile of the UA relative amplitude of the baseline recording; MIN value was set to the 5th percentile of the UA relative amplitude of the baseline recording.

4.3.5 Neurofeedback training session

During the NF session, participants were placed in the CAVE, seated on a chair. The CAVE was in a dark and quiet room. The experimenter helped the participants to wear the EEG device and headphones for sound isolation. The preparation of the recording equipment took from five to ten minutes, during which the quality of the recorded signals and the contacts between skin and electrodes were checked. Participant were instructed not to move their head during the NF session to avoid interference with signal acquisition.

Participants were immersed in the virtual environment, facing the feedback object changing colour.

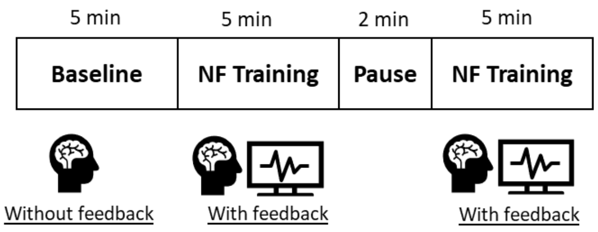


Figure 4.8 NF session structure

Each session was composed of three blocks: a resting Baseline block and two NF Training blocks. The Baseline block consisted in a 5-minutes recording in a resting state where subjects were instructed to stay relaxed and look at the object in front of them. During the Baseline recording they did not receive feedback about their brain activity (i.e., the colour of the object was fixed to white, didn't change).

The two following Training blocks lasted 5 minutes each, with a 2 minutes break in between. During the Training blocks, participants tried to modulate their brain activity in the desired direction. They were instructed to make the colour as red as possible. No other instruction or suggestion about strategies was given since the effective mental strategies vary among individuals [41]. Moreover, they were not allowed to close their eyes, because alpha activity naturally increases with eyes closed.

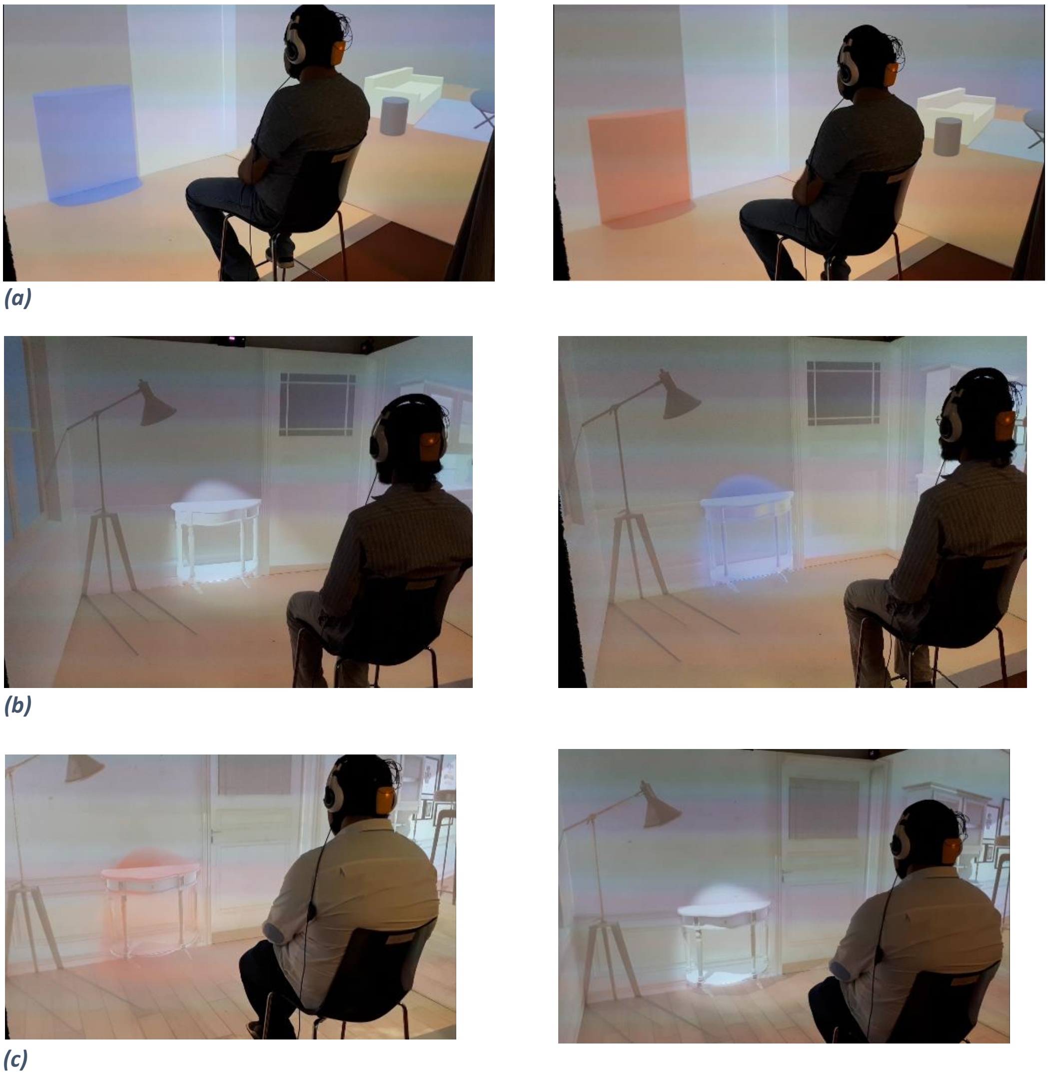


Figure 4.9 Participants during a NF training session.

In *(a)* a participant in *group A – Low vividness*. He is visualizing the virtual living room, seated in the CAVE. The cylinder in front of him is changing colour from blue to red, according to his UA relative amplitude.
In (b) a participant in *group B – Medium vividness*. In front of him in the virtual living room, there is a little table illuminated by a lamp. The light of the lamp is changing colour according to his UA relative amplitude.
In *(c)* a participant in *group C – High vividness*. The light of the lamp illuminating the table is changing colour according to his UA relative amplitude.

4.3.6 Neurofeedback transfer session

The Transfer session served to assess if the ability to modulate UA relative amplitude, acquired with NF in a specific modality (Low, Medium or High Vividness) could translate to other generic types of feedback.

One aim of NF is that, after multiple training sessions, the user should be able to reproduce the mental state, which is associated with the desired brain activity, also outside the NF training experience. Hence, the user should be able to modulate the brain activity with a different type of feedback or even without any feedback.

Transfer abilities are often tested asking participants to modulate their brain activity without providing feedback, but this can be a hard task and require long practice. We instead decided to test if learned ability could generalize to other types of feedback.

The Transfer session had the same described structure, but we used a standard 2-dimensional feedback used in literature: a square changing colour. The colour changed in the same way described for the previous modalities and was projected on the front wall of the CAVE. Thus, the settings and the NF task stayed the same as the previous training sessions.

4.4 Dependent variables

4.4.1 NF learning

A key factor in neurofeedback studies is what measures are used to identify whether an individual has been successful in altering the amplitude of a particular component of their EEG in the desired direction. Unfortunately, there is little consistency in the literature.

According to the review made by Dempster [68], the three most common measures used to identify changes in alpha activity are:

- amplitude, i.e., the changes in the mean level of amplitude during NF training;
- percent time, i.e., the percentage of time participants spend above or below the target threshold when attempting to enhance or inhibit their alpha;
- integrated alpha, a measure that combines both the amount of time spent over/under a pre-set threshold and the amplitude (e.g., integrated alpha = percent time x amplitude/100).

Moreover, the above measures can be used to identify possible changes in alpha in four possible methods. These include changes identified within the NFT session, changes across the NFT sessions, changes within sessions compared to a resting baseline and changes across sessions compared to a baseline.

From the study and considerations in [68], we can deduce some general rules:

I. A change in the EEG may be evident when looking at either amplitude or percent time, but changes in behaviour may only be evident with one. For instance, it could be the case that enhancing amplitude elicits a distinct effect on behaviour or cognition compared to enhancing percent time. Given that both amplitude and percent time measure different aspects of the EEG, it would seem prudent to include both measures but to look at them individually rather than combining them into a less sensitive measure. A combined measure does not provide a clear picture of where any changes occur, in amplitude, percent time, or both.
II. It is essential that baseline measures are included when attempting to identify evidence of learning via NFT. First, incorporating a baseline measure controls for natural fluctuations that can occur across sessions, due to uncontrolled factors (such as the amount of sleep on the preceding night, spontaneous cognitive activity and time between eating and EEG

recording). Second, it enables the researcher to see whether any changes seen during NFT exceed the amount of alpha participants naturally produce, or if it merely reflects a return to participants' natural levels after an initially suppressing effect, resulting from the completion of an unfamiliar task (as it has been observed in some studies).

III. Focusing on changes within NFT sessions may be more fruitful as possible changes across sessions may be confounded by shifting baselines.

For these reasons, we look at both amplitude and percentage of time to assess NF learning, and we always incorporate the baseline measures in our learning indices.

Within-session

We defined the following within-session learning indices. The indices L_1 and L_2 described the average learning ability in short term [12].

Average UA relative amplitude compared to baseline

To quantify the changes in the UA ratio within a session, we subtracted the average UA ratio during the resting baseline from the average UA ratio during the training session. This means that any resulting means which are positive in value represent enhancement above baseline and any negative values represent falling below baseline.

For every participant, we computed the average within-session change:

$$L_1 = \frac{\sum_{i=1}^{Nsess} mean(UA\ ratio\ training_i) - mean(UA\ ratio\ baseline_i)}{Nsess}$$

where $Nsess$ was the total number of NF session, i.e., 5 in our case.

Percentage of time above threshold

For every session, we considered the percentage of time during which the UA ratio was above the threshold, where the threshold was the median value of the UA ratio during the corresponding pre-training resting baseline.

For every participant, we computed the average percentage of time above threshold:

$$L_2 = \frac{\sum_{i=1}^{Nsess}\ \%\ time\ above\ threshold_i}{Nsess}$$

Across sessions

In order to check how the two measures (UA relative amplitude and percentage of time) changed across sessions, we defined the following across-sessions learning indices L_3 and L_4, which presented the learning ability across the whole training process and indicated accumulative training effects.

Average UA relative amplitude compared to baseline

For every training session, we considered the UA ratio increase from baseline. This means that we subtracted the average UA ratio during the resting baseline from the average UA ratio during the training session, like we did when computing L_1.

$$UA\ ratio\ increase_i = mean(UA\ ratio\ training_i) - mean(UA\ ratio\ baseline_i)$$

Then, for every participant we computed L_3 as the linear regression slope of that value over the 5 sessions.

Percentage of time above threshold

For every session, we considered the percentage of time during which the UA ratio was above the threshold, where the threshold was the median value of the UA ratio during the corresponding pre-training resting baseline.

Then, for every participant we computed L_4 as the linear regression slope of that value over the 5 sessions.

HYPOTHESIS:

We would hypothesize that a more vivid (thus more immersive) feedback would facilitate NF learning. The learning indices defined should be higher for participants in higher vividness groups.

4.4.2 NF transfer

The transfer session served to assess if the ability to control upper alpha, acquired during the NF training in a particular modality, could generalize to other types of feedback. In the ideal situation, a good NF training should translate into good performance during the Transfer session.

Performance during the Transfer session was measured using the same metrics described before:

- UA relative amplitude increase compared to baseline:

$$UA\ ratio\ increase_i = mean(UA\ ratio\ training_i) - mean(UA\ ratio\ baseline_i)$$

- Percentage of time above threshold: percentage of time during which the UA ratio was above the threshold, where the threshold was the median value of the UA ratio during the corresponding pre-training resting baseline.

HYPOTHESIS:

> *Based on the hypothesis that a more vivid feedback facilitates learning, we would expect participants from higher vividness groups to be able to transfer better their ability to control UA relative amplitude.*

4.4.3 Subjective Presence

Besides assessing the effect of vividness on NF learning, we measured the effect it could have on a subjective measure of presence. Presence is a state of consciousness concomitant with immersion and is related to a sense of being in a place [16].

We used the Slater-Usoh-Steed (SUS) questionnaire, that aims at measuring presence in immersive VEs [69]. A series of studies have been conducted showing correlation between the questionnaire results and objective measures of immersion [70, 71].

SUS questionnaire (*Appendix C*) was composed of 5 questions, each on a 1 to 7 scale where the higher score indicates greater presence. The overall score was computed as the mean value from responses to the five questions.

HYPOTHESIS:

> *More vivid virtual environments are more immersive. This should translate to the fact that subjects perceive greater presence in them.*

4.4.4 Motivation, Concentration, Stress, Sleepiness

As we repeated in the previous chapters, subjective user experience and psychological variables can be significative and influence NF performance.

Motivation and concentration can have an important role in NF training because NF practice requires participants to stay focused and concentrated on a NF task over a long training period. Moreover, the difficulty of the task and the length of the training period can make users feel tired and frustrated, this undermining NF performance.

We then selected the four variables Motivation, Concentration, Stress, and Sleepiness, in order to assess how different types of feedback affect them.

Through the questionnaire in *Appendix D*, participants were asked to rate how often they felt such sensations on a scale from 1 (Never) to 5 (Constantly).

HYPOTHESIS:

> *We would expect that a more vivid (thus more immersive) feedback is more engaging and attractive. Hence, participants in higher vividness groups should feel more motivated and concentrated and, conversely, less stressed and bored.*

4.4.5 Perceived competence

Perceived competence, i.e., the sense of mastery in executing the task, is also a variable that affects intrinsic motivation and, in turn, NF performance. We selected one of the subscales of the Intrinsic Motivation Inventory (IMI) to measure Perceived Competence.

Participants answered 6 questions, rating on a scale from 1 to 7 how much they felt competent during the task (*Appendix E*). The overall score was the mean of the rating of each question.

HYPOTHESIS:

> *Based on the hypothesis that a vivid feedback facilitates learning, we would expect that participants in higher vividness groups perceive higher competence in NF task.*

4.4.6 Perceived workload

We assessed the perceived workload for every session with the NASA Task Load Index (TLX) [72]. NASA-TLX gives a subjective estimate of workload considering the six factors of Mental Demand, Temporal Demand, Physical Demand, Performance, Effort, and Frustration. Each factor is rated in a scale with 20 points (1 = very low, 20 = very high, *Appendix F*).

The original version of the NASA-TLX requires a weighting process of the six subscales in order to obtain the overall score of the questionnaire. We used one of the most common modification of the NASA-TLX, the Raw TLX [72], in which the overall task load index is obtained by averaging the rating of each subscale.

4.4.7 Working memory

Working memory involves the ability to keep information active in mind for a short time (2-3 seconds) to be able to use it for further processing. Two commonly used tests for working memory assessment are the Digit Span test and the N-back test [73, 74].

We used Presentation[2] (Neurobehavioral Systems Inc.), a software application for psychological and neurobehavioral experiments, to run a Digit Span test and two N-back tests, respectively in the 2-back and 3-back versions.

Digit Span

The Digit Span (DS) is a test consisting of two tasks: a forward and a backward task.

In the forward task, participants listen to a sequence of numbers and are required to recall back the sequence correctly. The length of the sequences increases every two trials (i.e., there are two trial of length 3, then two trials of length 4, and so on). The forward digit span is defined as the length of the longest sequence the participant can repeat back in correct order on at least one of the two trials. The test ends when the person fails to recall correctly both the sequences of a given length.

The same holds for the backward task, except for the fact that the participants listen to the sequence of number and must recall it back in the reverse order. Thus, the backward digit span is the length of the longest sequence the participant can remember correctly in backward order.

---

[2] http://www.neurobs.com/

We considered both the measure forward DS and backward DS, although the backward DS is regarded to be more related to working memory, while the forward DS is to attention [75].

N-back

In the N-back task, subjects are presented with a stream of stimuli one-by-one. In our case, participants visualized a sequence of letters. The task is to decide for each item whether it matches the one presented N items before. An item that matches the one presented N steps before is called Target, otherwise it is a Distractor. When a Target item was recognized, participants had to report it (by clicking the mouse button); while Distractor items should be ignored.

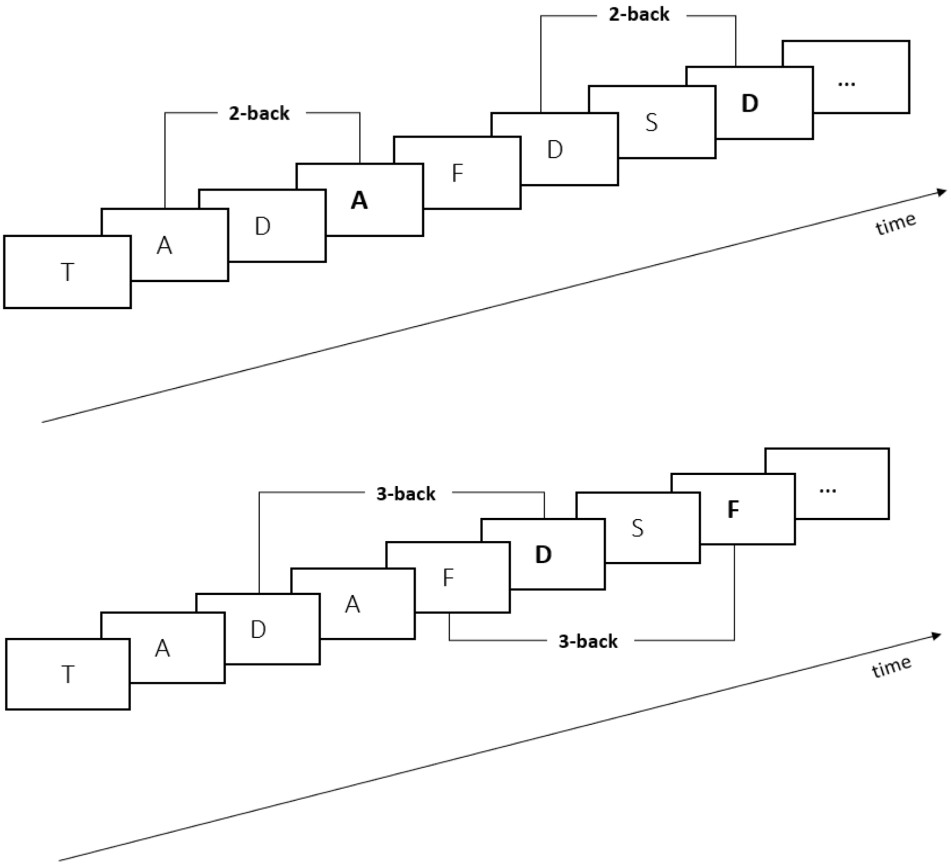


Figure 4.10 Examples of 2-back and 3-back tasks. The highlighted letters are Target items, the remaining are Distractors.

We measured performance in the N-back test considering both the accuracy of the subject in identifying Target items and the accuracy in identifying Distractor items (i.e., the percentage of correctly identified Targets/Distractors).

We decided to test two level of difficulty: 2-back and 3-back (in which subjects must find a match with the item presented 2 and 3 steps before, respectively).

Thus, we had four metrics of N-back performance:

- Target accuracy in 2-back
- Distractor accuracy in 2-back
- Target accuracy in 3-back
- Distractor accuracy in 3-back

In summary, our working memory performance measures are:

Table 4.2 Working memory performance metrics

WM Test	**Metric**	**Definition**
Digit span	Forward DS	Length of the longest sequence participants can repeat back in the correct order on at least 50% of trials
	Backward DS	Length of the longest sequence participants can remember correctly in backward order on at least 50% of trials
2-back	Target accuracy (2-back)	Percentage of correctly identified Targets in the 2-back task
	Distractor accuracy (2-back)	Percentage of correctly identified Distractors in the 2-back task
3-back	Target accuracy (3-back)	Percentage of correctly identified Targets in the 3-back task
	Distractor accuracy (3-back)	Percentage of correctly identified Distractors in the 3-back task

HYPOTHESIS:

We hypothesize to find an increase in these WM performance measure between pre-training and post-training tests. We would expect that such increase is positively correlated with the NF ability to modulate UA relative amplitude: a higher increase in UA ratio or percentage of time above threshold should result in a greater enhancement in WM performance.

5 Results

In this chapter we present the results of the analysis conducted on collected data.

As regards EEG data, the upper alpha relative amplitude recorded during the NF sessions was saved in *.csv* files for offline analysis. MATLAB (MathWorks Inc.) was used for deriving the NF learning indices (described in *section 4.4.1*) for every participant. Data from the questionnaires were analysed considering for each questionnaire the average score of all the training sessions. The Statistical Package for Social Science (SPSS) software package (IBM Corp.) was used to perform statistical analysis on data.

This study used an independent groups design, with one independent variable (level of vividness) and several dependent variables (NF learning performance, NF transfer performance, subjective presence, motivation, concentration, stress, sleepiness, perceived competence, perceived workload). The sample size was too small for a parametric analysis (i.e., not all the data were normally distributed) and some variables were measured at an ordinal level. Hence, to assess differences between groups, the non-parametric Kruskal-Wallis test by rank was used for each dependent variable. For each test, we reported the mean rank of each group and the test statistic H, with its degree of freedom and its significance. Significance was considered for p-values below 0.05.

For working memory data analysis, we computed the differences between the scores in the post-tests and the scores in the pre-tests. Then, we used correlation analysis between these variables and the indices of NF learning. The Spearman's rank coefficient was used as non-parametric measure of rank correlation.

Because of missing values in the data, the sample size can vary among different tests. We reported the sample size N for each test.

In the following section *5.1* we summarize the results of the statistical analysis with tables and box plots. In section 5.2 we discuss the obtained results.

5.1 Statistical analysis

5.1.1 NF learning

The table below shows the descriptive statistics for the learning indices defined in the previous chapter (*section 4.4.1*).

Descriptive Statistics

	Condition	N	Mean	SD	Median
L_1 - Average UA ratio increase from baseline	A - Low vividness	7	-0,0018	0,025	-0,0095
	B - Medium vividness	7	0,0027	0,037	0,0119
	C - High vividness	7	0,0443	0,050	0,0409
L_2 - Average % time above threshold	A - Low vividness	7	48,9	5,576	46,2
	B - Medium vividness	7	48,6	5,016	49
	C - High vividness	7	55,3	5,093	55,4
L_3 - Slope UA ratio increase from baseline	A - Low vividness	7	-0,0039	0,013	-0,0028
	B - Medium vividness	7	-0,0225	0,035	-0,0192
	C - High vividness	7	-0,0173	0,039	-0,0102
L_4 - Slope % time above threshold	A - Low vividness	7	-1,129	2,429	-1,3
	B - Medium vividness	7	-2,914	3,93	-2,9
	C - High vividness	7	-0,757	3,141	-0,1

Within session

L_1 - Average UA relative amplitude increase from baseline

The index L_1 represents how much a participant managed to increase the UA relative amplitude (compared to the baseline level) within a session.

In the Box plot below, we can see that there is a tendency for participants in higher vividness groups to have a higher L_1 value.

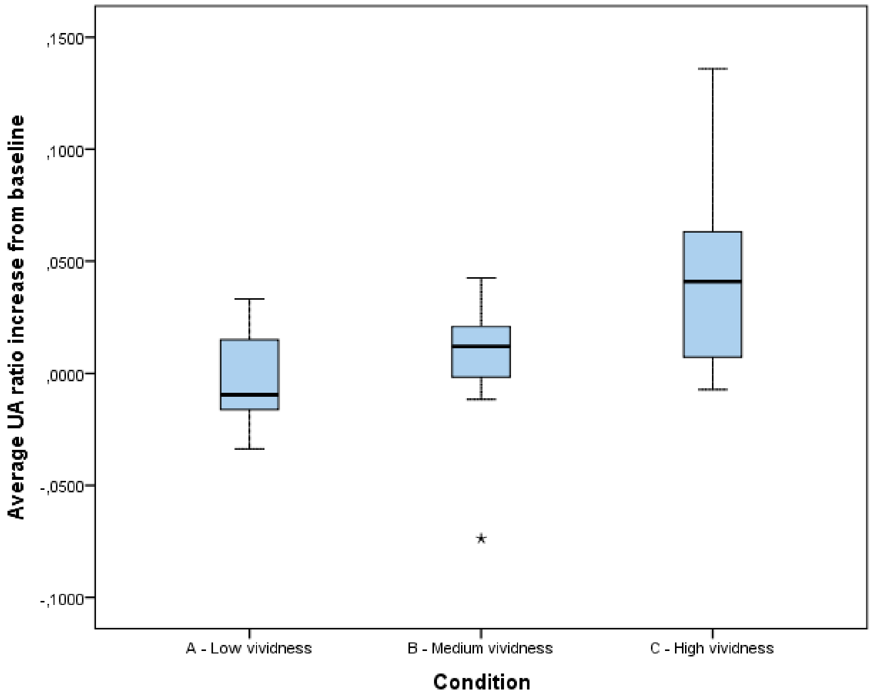


Figure 5.1 Box plot – Average UA ratio increase from baseline

Group A (Low vividness) has a median value below 0, this meaning that participants didn’t manage to successfully increase the UA relative amplitude. Group B (Medium vividness) and C (High vividness) have a positive median value, thus managed to modulate UA relative amplitude in the desired direction, with group C performing better than group B.

Participants in higher vividness groups tended to better modulate UA relative amplitude in the desired direction compared to lower vividness groups. However, a Kruskal-Wallis H test showed that there was not a statistically significant difference in the UA ratio increase between the different groups, *H(2) = 4.839, p > 0.05*, although close to significance (*p = 0.089*).

Ranks

	Condition	N	Mean Rank
Average UA ratio increase from baseline	A - Low vividness	7	7,86
	B - Medium vividness	7	10,14
	C - High vividness	7	15,00
	Total	21	

Kruskal-Wallis test Statistics

	Average UA ratio increase from baseline
H	4,839
df	2
Asymp. Sig.	0,089

L_2 - Percentage of time above threshold

The index L_2 represents the percentage of time a participant kept the UA relative amplitude above the threshold (baseline level) within a session.

As for L_1, we can notice from the plot a tendency for participants in higher vividness groups to have a higher L_2 value.

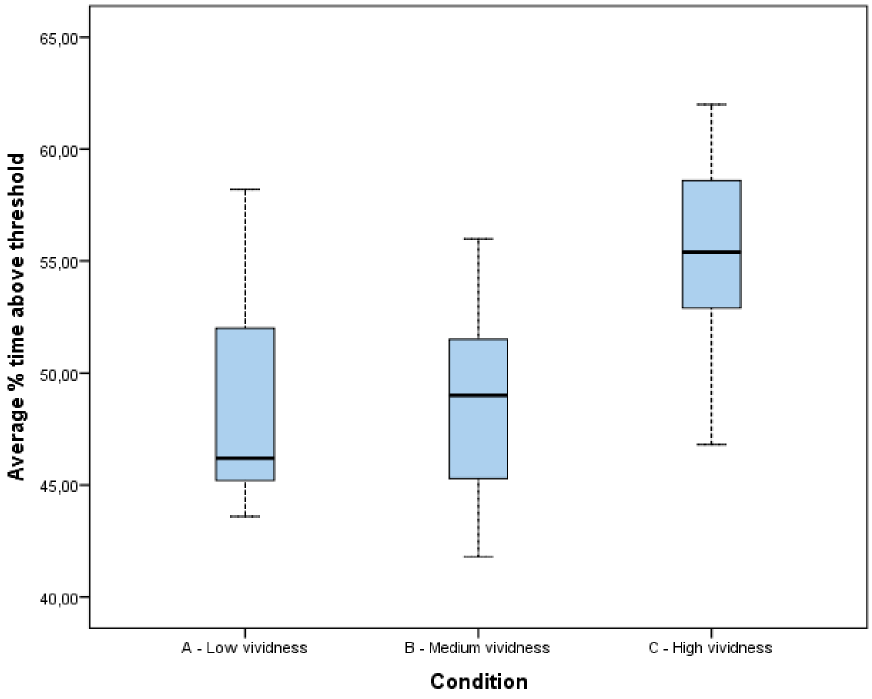


Figure 5.2 Box plot – Average % time above threshold

During a training session, participants in Group C managed to modulate their UA relative amplitude above the threshold level for a longer time than participants in group B. And the same holds for group B participants compared to group A participants.

A Kruskal-Wallis H test showed that there was not a statistically significant difference in the percentage of time above threshold between the different groups, $H(2) = 5.705, p > 0.05$. However, the p-value was close to the significance level ($p = 0.058$).

Ranks

	Condition	N	Mean Rank
Average % time above threshold	A - Low vividness	7	8,64
	B - Medium vividness	7	8,79
	C - High vividness	7	15,57
	Total	21	

Kruskal-Wallis test Statistics

	Average % time above threshold
H	5,705
df	2
Asymp. Sig.	0,058

Across sessions

L_3 – Slope of UA relative amplitude increase baseline

The index L_3 corresponds to the linear regression slope of the UA ratio increase over the five training sessions. It represents the evolution of the UA ratio increase from the baseline level across the whole training period. A positive value of L_3 indicates that the participant could enhance the UA ratio across sessions.

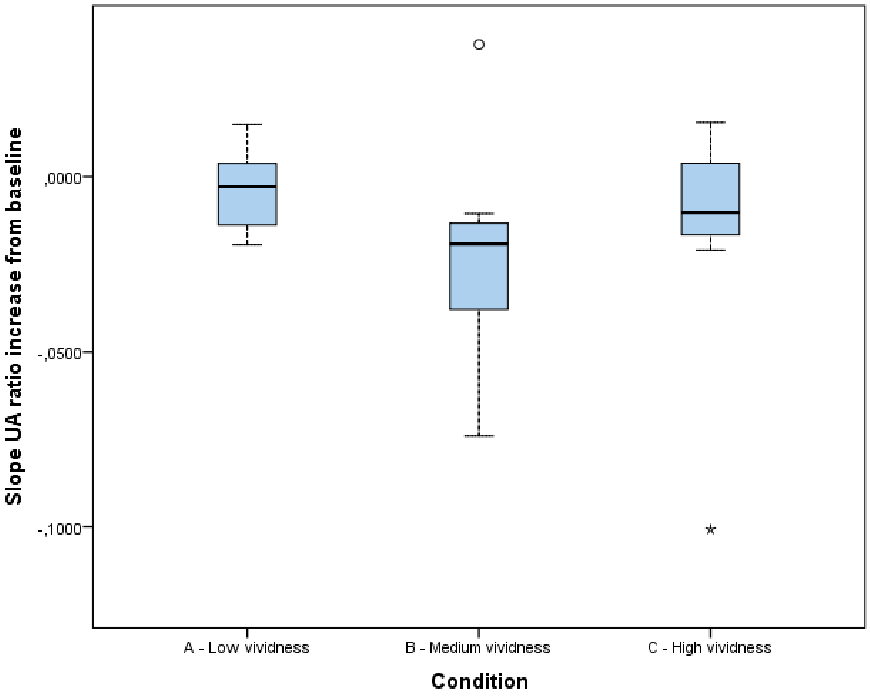


Figure 5.3 Box plot – Linear regression slope of the UA ratio increase from baseline

The box plot shows that the median value of L_3 is negative for all the groups, suggesting there was not an overall increase of UA ratio across sessions.

The Kruskal-Wallis test found no significant difference in the regression slope between groups, *H(2) = 2.494, p > 0.05*.

Ranks

	Condition	N	Mean Rank
Slope UA ratio increase from baseline	A - Low vividness	7	13,29
	B - Medium vividness	7	8,14
	C - High vividness	7	11,57
	Total	21	

Kruskal-Wallis test Statistics

	Slope UA ratio increase from baseline
H	2,494
df	2
Asymp. Sig.	0,287

L_4 – Slope of percentage of time above threshold

The index L_4 corresponds to the linear regression slope of the percent time above threshold over the five training sessions. It represents how the percent time changed across sessions, with a positive value of L_4 suggesting that the participant could increase the percentage of time spent above threshold across sessions.

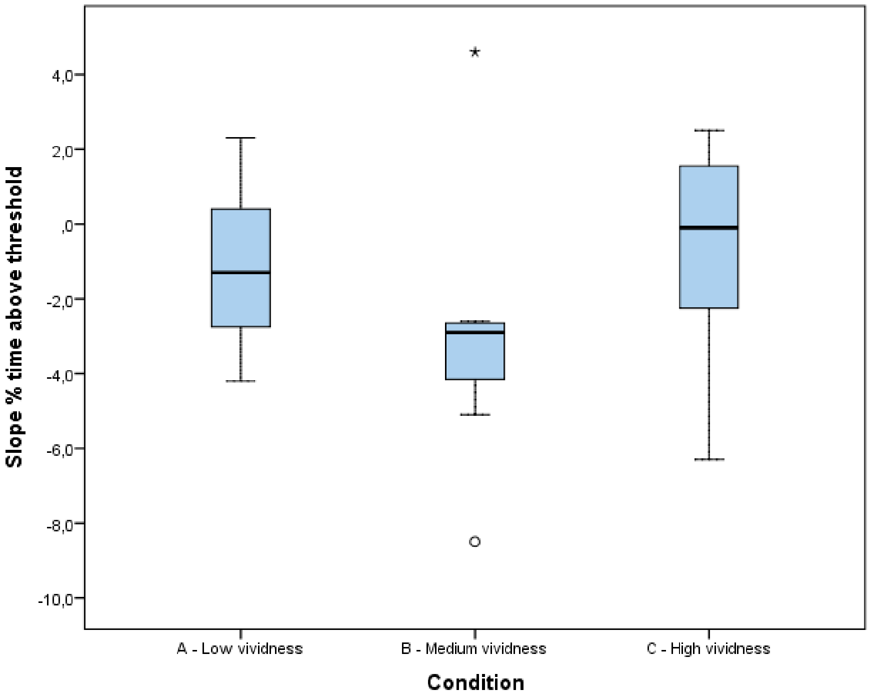


Figure 5.4 Box plot – Linear regression slope of the % time above threshold

The median value of L_4 is negative for all the groups, suggesting there was not an increase in the percentage of time above threshold across sessions.

A Kruskal-Wallis test showed that there was not a statistically significant difference in the L_4 value between the different groups, *H(2) = 1.955, p > 0.05*

Ranks

	Condition	N	Mean Rank
Slope % time above threshold	A - Low vividness	7	11,64
	B - Medium vividness	7	8,43
	C - High vividness	7	12,93
	Total	21	

Kruskal-Wallis test Statistics

	Slope % time above threshold
H	1,955
df	2
Asymp. Sig.	0,376

Correlation between learning indices

For exploratory purposes, a correlation analysis between the four learning indices was performed, using the Spearman's rank correlation coefficient.

We found a statistically significant positive relationship between L_1 and L_2 (*r = 0.93, p < 0.01, N = 21*) and between L_3 and L_4 (*r = 0.94, p < 0.01, N = 21*). L_1 and L_2 measure the change of UA relative amplitude and percentage of time above threshold respectively, within a session. While L_3 and L_4 measure the change of UA amplitude and percentage of time across sessions. Hence, the results indicate a strong correlation between the two metrics UA relative amplitude and percentage of time above threshold: an increase of UA amplitude within session corresponds to an increase of percentage of time above threshold within session; similarly, an increase of UA amplitude across sessions corresponds to an increase of percentage of time across sessions.

Moreover, a negative correlation between L_1 and L_3 was found, even though not statistically significant (*r = -0.38, p = 0.09, N = 21*). This relationship suggests that participants who performed better within session, achieving higher increase of UA ratio with respect to the baseline level, tended to show lower increase of UA ratio across sessions. Conversely, participants who showed low UA ratio increase within session, attained a high UA ratio increase across sessions.

5.1.2 NF transfer

As described in the previous chapter (*section 4.4.2*), we analysed the NF performance during the transfer session in terms of UA relative amplitude increase from baseline level and percentage of time above threshold.

The table below shows the descriptive statistics for these measures.

Descriptive Statistics

	Condition	N	Mean	SD	Median
UA ratio increase from baseline	A - Low vividness	7	0,0048	0,0723	-0,0028
	B - Medium vividness	7	-0,0026	0,0333	0,0116
	C - High vividness	7	-0,0117	0,0902	-0,0274
% time above threshold	A - Low vividness	7	47,8	14,2	49
	B - Medium vividness	7	49,6	3,82	51
	C - High vividness	7	48,5	12,3	43

UA relative amplitude increase from baseline

The box plot below shows the increase in the UA relative amplitude during the transfer session per group. Only for group B the median value is above 0, indicating that participants during the transfer session successfully modulated the UA ratio above the baseline level.

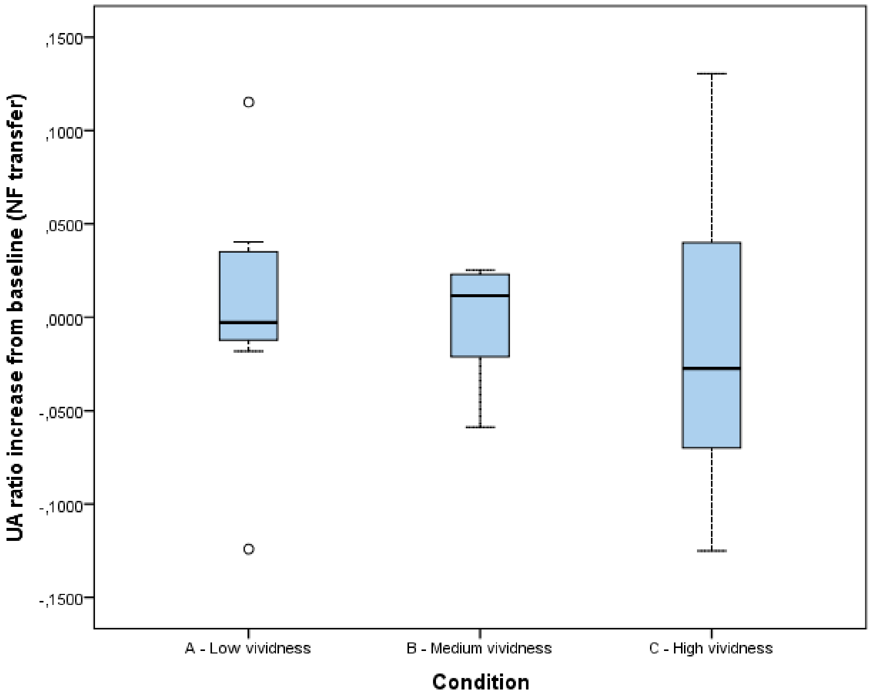


Figure 5.5 Box plot – UA ratio increase from baseline during the NF Transfer session

A Kruskal-Wallis test showed no statistically significant difference between groups, *H(2) = 0.475, p > 0.05*.

Ranks

	Condition	N	Mean Rank
UA ratio increase from baseline (NF transfer)	A - Low vividness	7	12,14
	B - Medium vividness	7	11,00
	C - High vividness	7	9,86
	Total	21	

Kruskal-Wallis test Statistics

	UA ratio increase from baseline (NF transfer)
H	0,475
df	2
Asymp. Sig.	0,789

Percentage of time above threshold

The box plot below shows the percentage of time the UA ratio was above the baseline level during the transfer session per group.

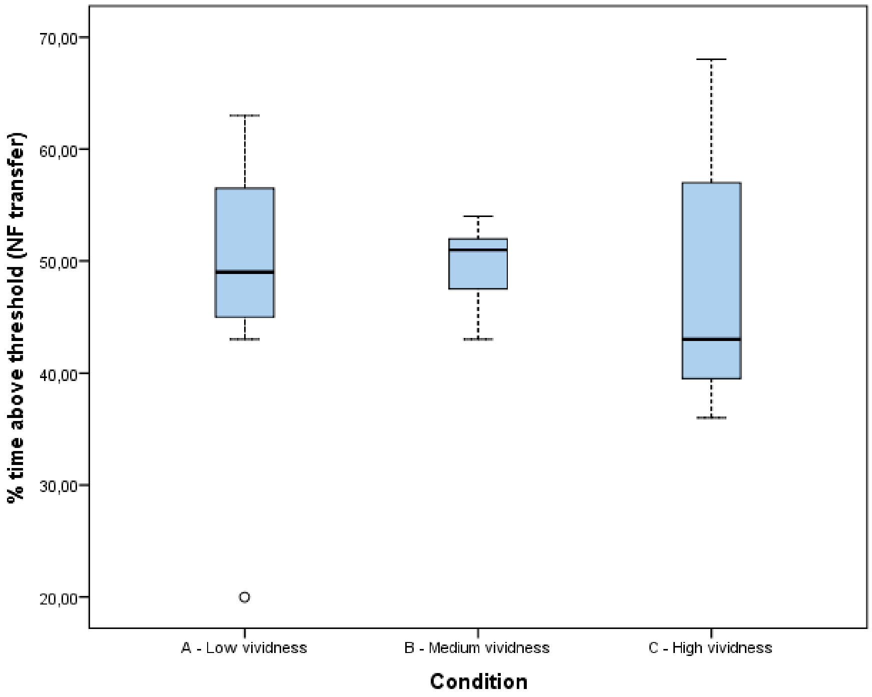


Figure 5.6 Box plot – Percentage of time above threshold during the NF transfer session

Results are comparable between groups and a Kruskal-Wallis test showed no statistically significant difference, *H(2) = 0.282, p > 0.05*.

Ranks

	Condition	N	Mean Rank
% time above threshold (NF transfer)	A - Low vividness	7	11,36
	B - Medium vividness	7	11,64
	C - High vividness	7	10,00
	Total	21	

Kruskal-Wallis test Statistics

	% time above threshold (NF transfer)
H	0,282
df	2
Asymp. Sig.	0,868

5.1.3 Subjective Presence

The next figures show results for the SUS questionnaire (*section 4.4.3*), that gives a subjective measures of presence in virtual environments in a scoring system ranging from 1 (low) to 7 (high).

Descriptive Statistics

	Condition	N	Mean	SD	Median
SUS score	A - Low vividness	7	2,43	1,21	2,4
	B - Medium vividness	7	4,14	1,22	3,8
	C - High vividness	7	3,29	1,01	2,8

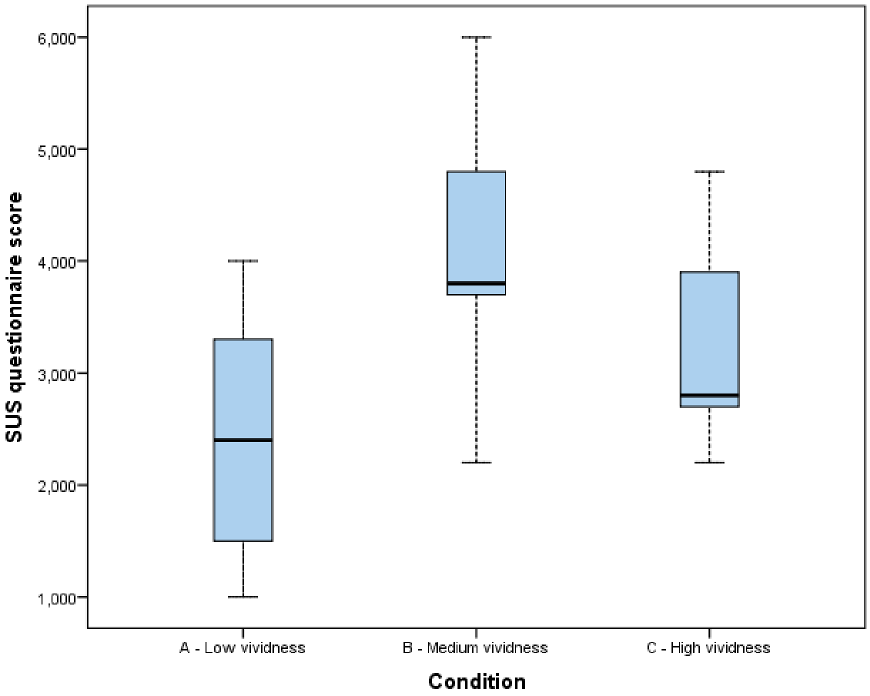


Figure 5.7 Box plot – SUS questionnaire score

Participants in group A (Low vividness) reported the lowest SUS score, but not statistically significant difference was found between groups, *H(2) = 4.954, p > 0.05*.

Ranks

	Condition	N	Mean Rank
SUS score	A - Low vividness	7	7,29
	B - Medium vividness	7	14,64
	C - High vividness	7	11,07
	Total	21	

Kruskal-Wallis test Statistics

	SUS score
H	4,954
df	2
Asymp. Sig.	0,084

5.1.4 Motivation, Concentration, Stress, Sleepiness

Participants rated their sensations of motivation, concentration, stress and sleepiness (*section 4.4.4*) during the NF training on a scale from 1 (Never felt) to 5 (Constantly felt).

Motivation

The next figures show results for the motivation variable. We can notice an increasing trend, with participants in higher vividness groups reporting to feel more motivated during NF training.

Descriptive Statistics

	Condition	N	Mean	SD	Median
Motivation score	A - Low vividness	5	2,64	0,95	2,8
	B - Medium vividness	6	3,13	0,60	3,3
	C - High vividness	5	3,64	0,67	3,6

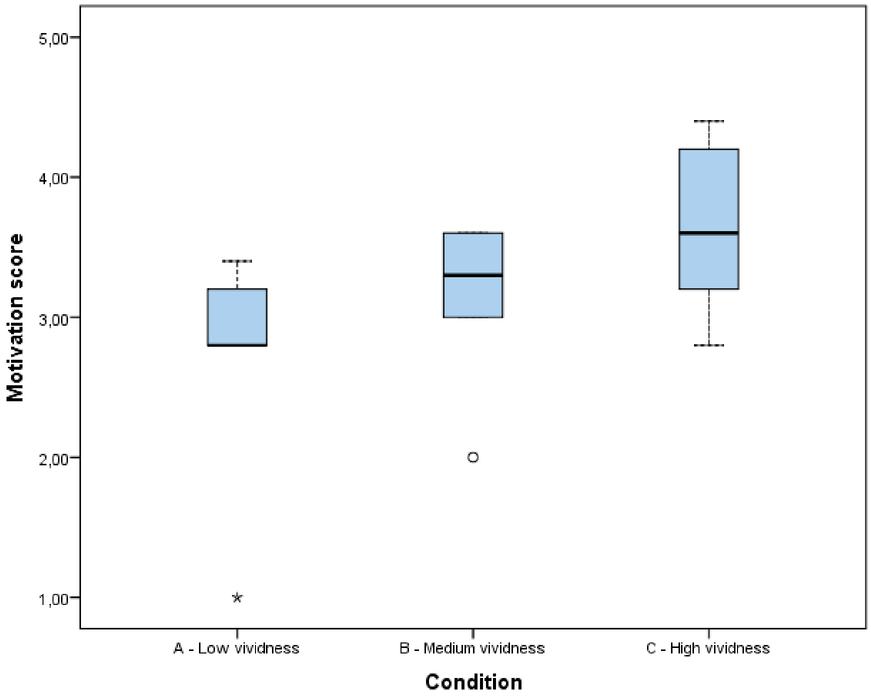


Figure 5.8 Box plot – Motivation score

However, a Kruskal-Wallis test showed no statistically significant difference between groups, *H(2) = 3.680, p > 0.05*.

Ranks

	Condition	N	Mean Rank
Motivation score	A - Low vividness	5	5,50
	B - Medium vividness	6	8,75
	C - High vividness	5	11,20
	Total	16	

Kruskal-Wallis test Statistics

	Motivation score
H	3,680
df	2
Asymp. Sig.	0,159

Concentration

The next figures show results for the concentration variable. As for motivation, we can notice that participants in higher vividness groups tended to feel more concentrated during NF training with respect to lower vividness groups.

Descriptive Statistics

	Condition	N	Mean	SD	Median
Concentration score	A - Low vividness	5	2,84	0,73	3,2
	B - Medium vividness	6	3,43	0,61	3,4
	C - High vividness	5	3,92	0,64	3,6

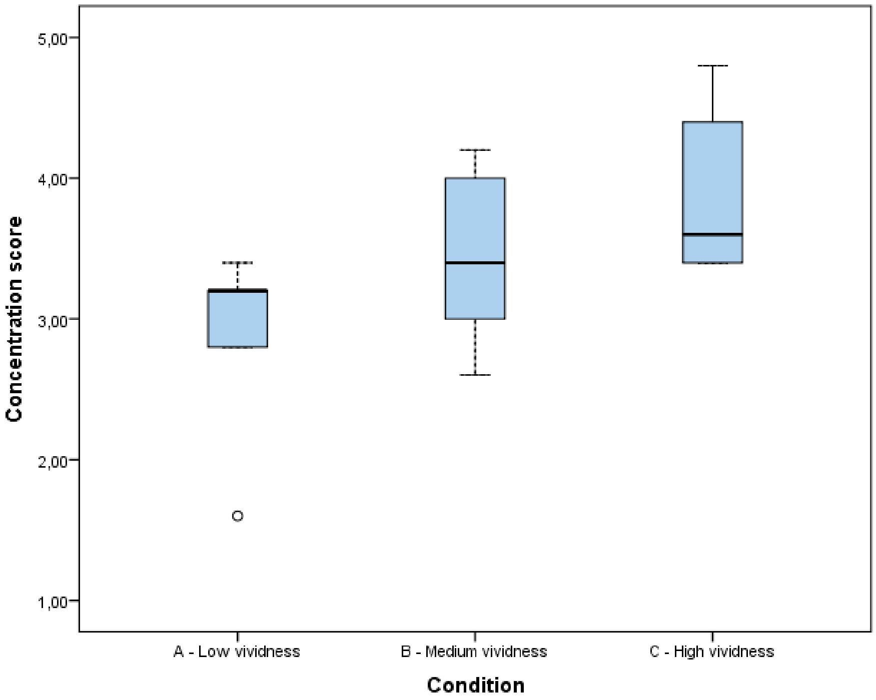


Figure 5.9 Box plot – Concentration score

A Kruskal-Wallis test showed no statistically significant difference between groups, *H(2) = 5.637, p > 0.05*.

Ranks

	Condition	N	Mean Rank
Concentration score	A - Low vividness	5	5,00
	B - Medium vividness	6	8,42
	C - High vividness	5	12,10
	Total	16	

Kruskal-Wallis test Statistics

	Concentration score
H	5,637
df	2
Asymp. Sig.	0,060

Stress

In the next figures we can see the results for the stress variable. All participants reported low stress scores.

Descriptive Statistics

	Condition	N	Mean	SD	Median
Stress score	A - Low vividness	5	2,00	0,73	2,0
	B - Medium vividness	6	1,60	0,71	1,3
	C - High vividness	5	1,72	0,64	1,8

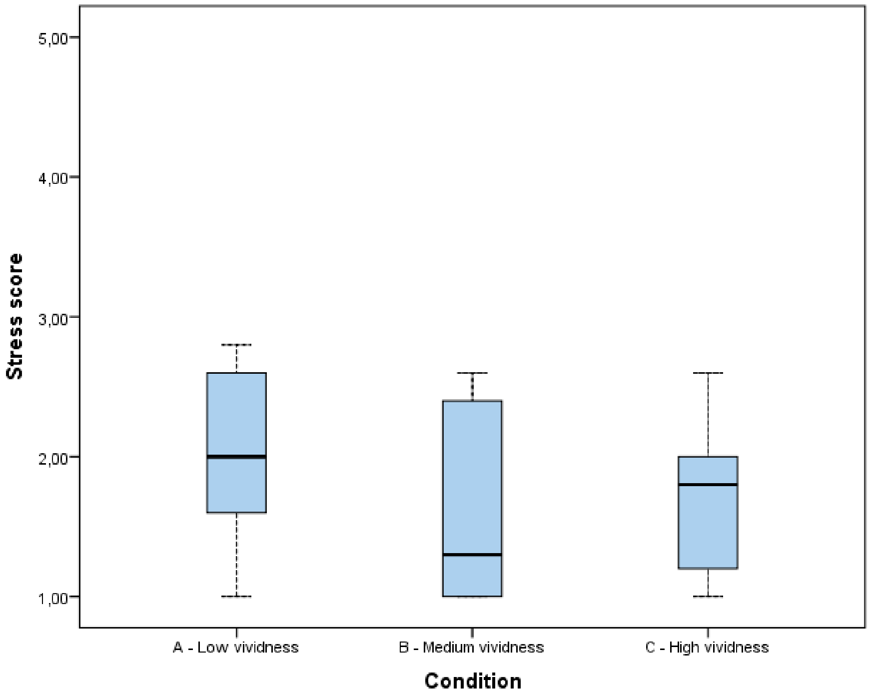


Figure 5.10 Box plot – Stress score

No statistically significant difference was found between groups, *H(2) = 1.085, p > 0.05*.

Ranks

	Condition	N	Mean Rank
Stress score	A - Low vividness	5	10,20
	B - Medium vividness	6	7,25
	C - High vividness	5	8,30
	Total	16	

Kruskal-Wallis test Statistics

	Stress score
H	1,085
df	2
Asymp. Sig.	0,581

Sleepiness

In the next figures we can see how participants rated their sensation of sleepiness. Participants in group A were the ones feeling drowsier during NF training sessions.

Descriptive Statistics

	Condition	N	Mean	SD	Median
Sleepiness score	A - Low vividness	5	3,28	0,98	3,4
	B - Medium vividness	6	2,53	0,74	2,4
	C - High vividness	5	2,44	0,89	2,2

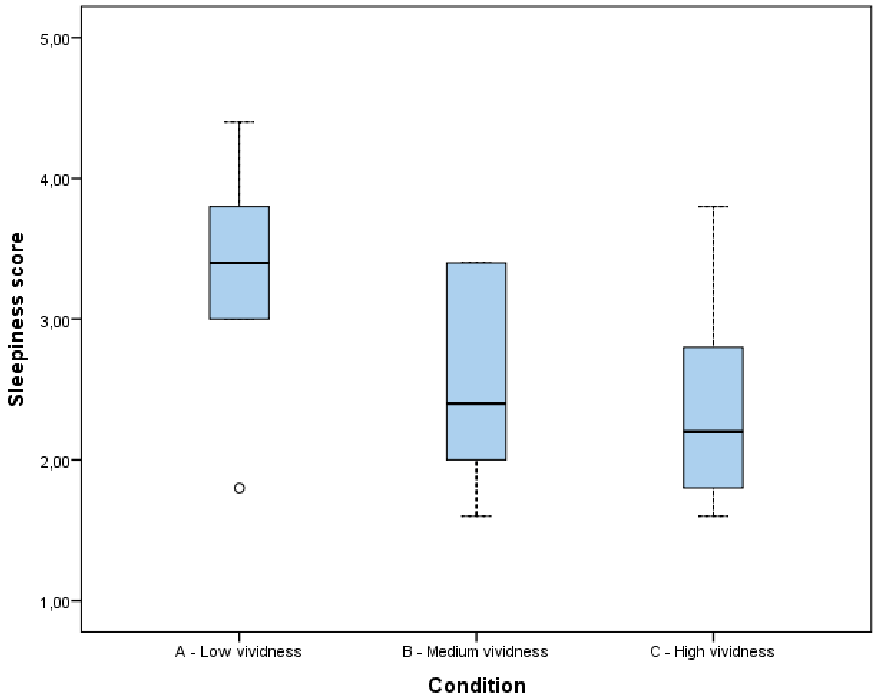


Figure 5.11 Box plot – Sleepiness score

A Kruskal-Wallis test showed no statistically significant difference between groups, *H(2) = 2.397, p > 0.05*

Ranks

	Condition	N	Mean Rank
Sleepiness score	A - Low vividness	5	11,20
	B - Medium vividness	6	7,50
	C - High vividness	5	7,00
	Total	16	

Kruskal-Wallis test Statistics

	Sleepiness score
H	2,397
df	2
Asymp. Sig.	0,302

5.1.5 Perceived competence

Participants rated on a scale from 1 (low) to 7 (high) their feeling of control (*section 4.4.5*) during the NF training sessions. We can notice from the figures below that the perceived competence level is comparable between groups.

Descriptive Statistics

	Condition	N	Mean	SD	Median
Perceived competence IMI score	A - Low vividness	7	3,35	1,52	3,5
	B - Medium vividness	7	3,86	0,88	3,6
	C - High vividness	7	3,29	0,84	3,4

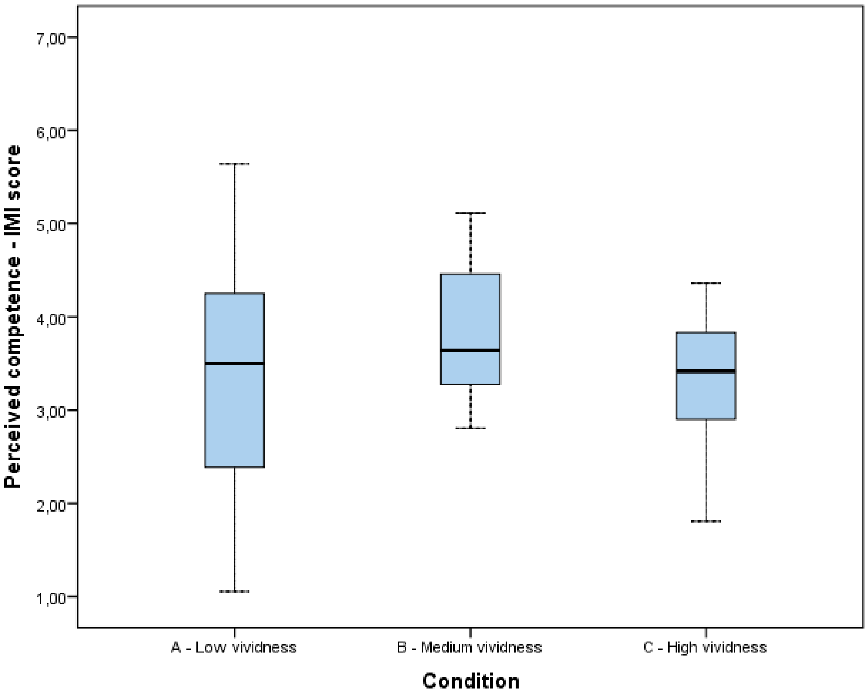


Figure 5.12 Box plot – IMI Perceived competence score

A Kruskal-Wallis test showed that there are not statistically significant difference between groups, *H(2) = 0.831, p > 0.05*

Ranks

	Condition	N	Mean Rank
Perceived competence IMI score	A - Low vividness	7	10,43
	B - Medium vividness	7	12,71
	C - High vividness	7	9,86
	Total	21	

Kruskal-Wallis test Statistics

	Perceived competence IMI score
H	0,831
df	2
Asymp. Sig.	0,660

5.1.6 Perceived workload

Through the NASA-TLX questionnaire (*section 4.4.6*), participants rated the perceived workload during the NF sessions on a scale from 1 (low) to 20 (high). The next figures report the results of TLX questionnaire.

Descriptive Statistics

	Condition	N	Mean	SD	Median
Perceived workload TLX score	A - Low vividness	7	8,13	2,10	7,83
	B - Medium vividness	7	9,19	2,04	9,87
	C - High vividness	7	10,23	2,26	10,43

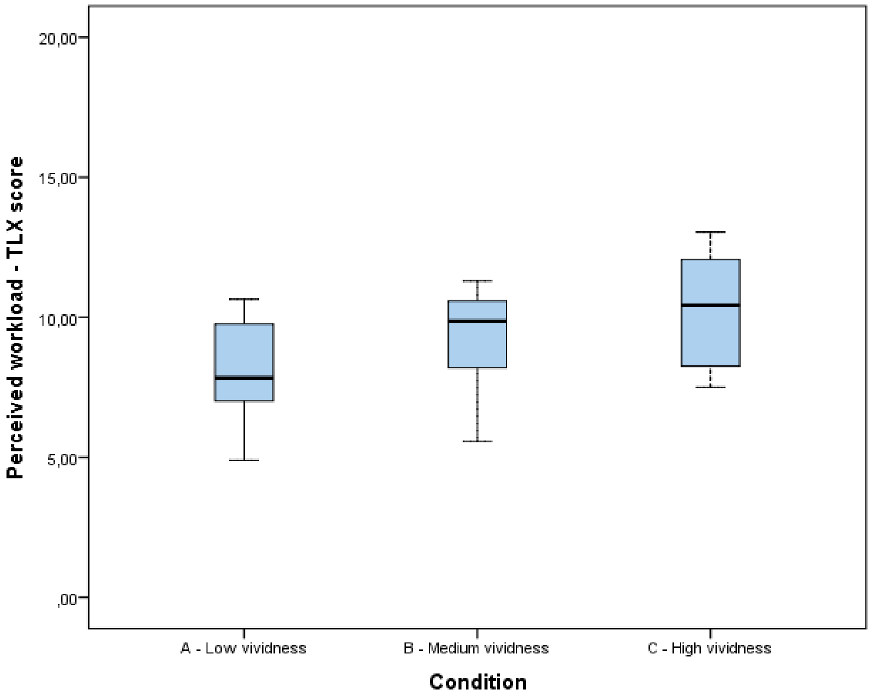


Figure 5.13 Box plot – TLX score

We can notice that the perceived workload tended to increase in higher vividness compared to lower vividness groups. However, no statistically significant difference was found in the TLX score between groups, *H(2) = 2.753, p > 0.05*.

Ranks

	Condition	N	Mean Rank
Perceived workload TLX score	A - Low vividness	7	8,21
	B - Medium vividness	7	11,07
	C - High vividness	7	13,71
	Total	21	

Kruskal-Wallis test Statistics

	Perceived workload TLX score
H	2,753
df	2
Asymp. Sig.	0,252

5.1.7 Working memory

For each working memory test (*section 4.4.7*), we computed the increase from the pre-test score to the post-test score.

Digit Span test

In the next figures we can see the results relative to the Digit Span test. The forward Digit span slightly increased in the post-test, while the backward Digit span stayed at the same level (median increase equal to 0).

Descriptive Statistics

	N	Mean	SD	Median
Forward Digit span increase	16	0,75	0,931	1,00
Backward Digit span increase	16	-0,31	1,302	0,00

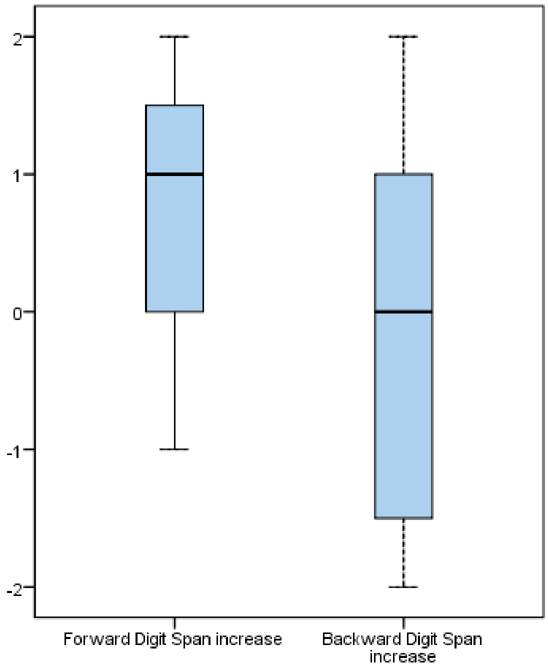


Figure 5.14 Box plot – Digit Span increase from pre to post-test

In the following table we can see the Spearman's rank correlation coefficients between forward and backward Digit Span increase and the indices of NF learning. For each pair of variables analysed, the significance value is above *0.05*. Thus, no statistically significant correlation was found between Digit Span test results and UA relative amplitude changes.

Correlations

		Forward Digit Span increase	**Backward Digit Span increase**
L_1 - Average UA ratio increase from baseline	Correlation Coefficient	,037	,139
	Sig. (1-tailed)	,446	,304
	N	16	16
L_2 - Average % time above threshold	Correlation Coefficient	-,064	,239
	Sig. (1-tailed)	,407	,187
	N	16	16
L_3 - Slope UA ratio increase from baseline	Correlation Coefficient	-,240	,121
	Sig. (1-tailed)	,186	,328
	N	16	16
L_4 - Slope % time above threshold	Correlation Coefficient	-,179	,143
	Sig. (1-tailed)	,253	,298
	N	16	16

N-back test

The next figures show the results relative to the N-back tests (in the 2-back and 3-back versions).

Descriptive Statistics

	N	Mean	SD	Median
Target accuracy increase – 2-back	16	0,112	0,163	0,050
Distractor accuracy increase – 2-back	16	-0,014	0,052	-0,025
Target accuracy increase – 3-back	16	0,025	0,211	0,000
Distractor accuracy increase – 3-back	16	-0,006	0,078	0,125

Only the Target accuracy in the 2-back and the Distractor accuracy in the 3-back slightly increased from the pre to the post-test (positive median value).

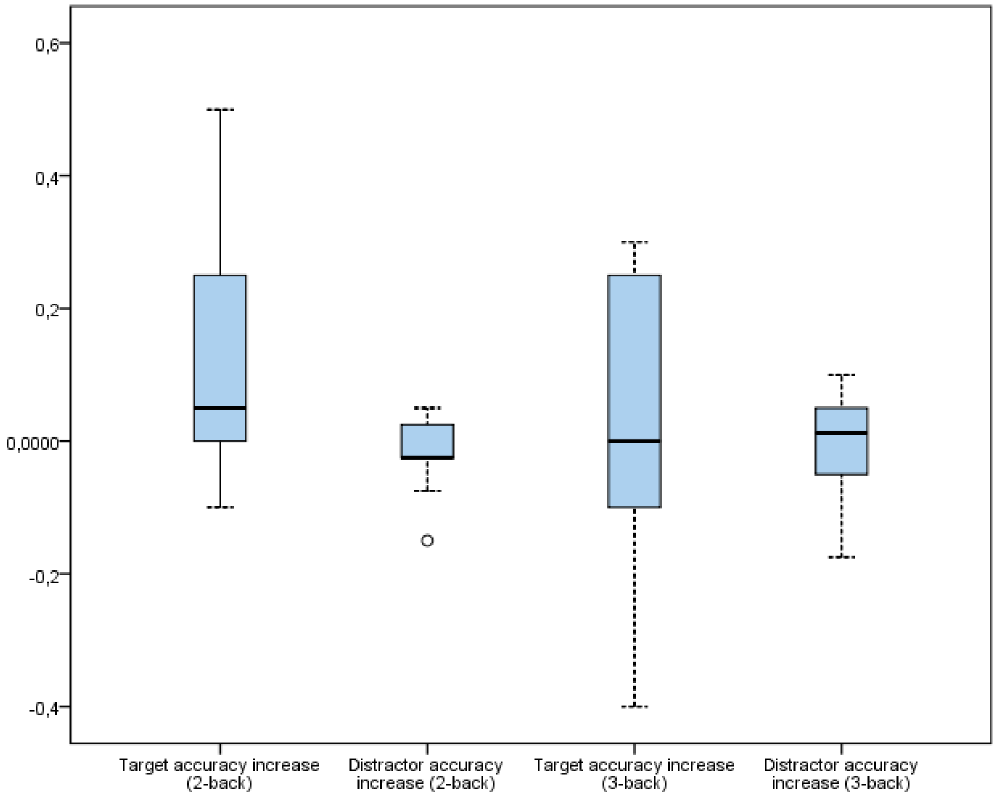


Figure 5.15 Box plot – 2-back and 3-back test results

The table below shows the Spearman's rank correlation coefficients calculated between the indices of NF learning and the performance measures in the N-back test.

A statistically significant relationship was found between the NF learning index L_3 and the Distractor accuracy increase in the 3-back test, with a correlation coefficient $r = 0.641, p < 0.01$. The Distractor accuracy increase in the 3-back test resulted to be significantly correlated with the NF learning index L_4 as well, $r = 0.639, p < 0.01$.

These positive correlation coefficients suggest that, when the UA ratio or the percentage of time above threshold increased across the NF sessions, this corresponded to an enhancement in the Distractor accuracy in the 3-back test.

Correlations

		Target accuracy increase (2-back)	**Distractor accuracy increase (2-back)**	**Target accuracy increase (3-back)**	**Distractor accuracy increase (3-back)**
L_1 - Average UA ratio increase from baseline	Correlation Coefficient	-,111	-,552	-,021	-,262
	Sig. (1-tailed)	,341	,013	,469	,164
	N	16	16	16	16
L_2 - Average % time above threshold	Correlation Coefficient	-,394	-,396	-,154	-,084
	Sig. (1-tailed)	,065	,064	,285	,379
	N	16	16	16	16
L_3 - Slope UA ratio increase from baseline	Correlation Coefficient	-,315	,244	,054	,641
	Sig. (1-tailed)	,117	,181	,421	,004
	N	16	16	16	16
L_4 - Slope % time above threshold	Correlation Coefficient	-,374	,111	,027	,639
	Sig. (1-tailed)	,077	,342	,461	,004
	N	16	16	16	16

5.2 Discussion

In the present study we manipulated the vividness of the virtual environment used for providing feedback in a NF procedure. Our aim was to assess the effect of different levels of vividness on user experience during NF training as well as on the NF training outcome. Furthermore, we evaluated the effect of UA neurofeedback training on working memory performance.

Effect of vividness on NF training performance

Since literature suggests that the immersive properties of virtual environments facilitate NF learning, our hypothesis was that a more vivid (thus more immersive) virtual environment would imply better NF performance. When speaking of NF performance, we measured it using two metrics: the increase of UA relative amplitude with respect to the baseline level and the percentage of time the participants spent above the baseline threshold. These two metrics were shown to be strongly positively correlated.

From data analysis, it emerges that participants in higher vividness groups tended to perform better within a NF session, in terms of both performance metrics, than participants in lower vividness groups: participants in group C attained better NF performance within a training session compared to participants in group B; the same holds, in turn, for participants in group B, who showed an improved NF training performance compared to group A. Statistical analysis showed that the difference between groups was only marginally significant, with a *p*-value slightly above 0.05. However, given the small sample size (N=7 for each group), this is a positive result and suggests a positive effect of vividness on NF training, in accordance with our hypothesis.

It is important to notice that participants in group A failed to attain control on their UA relative amplitude, showing no increase in the UA ratio with respect to the baseline level. We would have expected that every group was able to modulate the UA ratio in the desired direction, with an advantage for the higher vividness groups. An explanation could be that the Low Vividness virtual environment itself hampered participants in acquiring the UA self-regulation skill. The VE was monotonous and boring, thus not engaging compared to the higher vividness VEs. This could have made participants tired and reduced their dedication to the NF task [60]. This seems to confirm the importance of vividness and the advantage of a highly vivid virtual environment.

While evidence of NF learning within session was found, there was not improvement in NF performance across sessions (neither in UA amplitude increase, nor in time spent above threshold). This could be due to the length of the training period. Our training schedule consisted of five NF session, each composed of two 5-minutes training blocks, for a total of 50 minutes of NF training. Studies in which significant Upper Alpha NF learning across sessions was found used a NF procedure with longer sessions (25 minutes) [40, 17] and/or with a greater number of sessions (about 10) [42, 43], resulting in at least double the NF training time in our experiment. Furthermore, it has been shown in literature that when significant NF changes across sessions were found, it was coming from an increase between the first and the later sessions, while no significant differences were identified with any of the intermediate sessions [68]. This suggests that, in the early stage of the NF training, changes across sessions could not be detected. Thus, the fact that participants in this study underwent a relatively short (50 minutes overall) NF training can be the reason why an enhancement of NF performance across sessions was not found.

Furthermore, we noticed a negative relationship between the performance measure within session (ability to up-regulate the UA relative amplitude in a session) and across sessions (ability to enhance the UA relative amplitude across sessions). This indicates that participants who showed low increase of UA ratio within session, tended to enhance NF performance across sessions.

Finally, no significant difference in NF performance during the NF Transfer session was found between groups. The aim of the transfer session was assessing how the ability acquired during the NF training generalizes to another type of feedback. Since results were comparable between groups, we can argue that the vividness of the training scenario had no effect on NF transfer ability.

Effect of vividness on user experience

From the analysis of questionnaire data, it results that the vividness of the virtual environment had no statistically significant effect on subjective presence response, as measured with SUS questionnaire. Even though not significant, we could notice that perceived presence tended to be higher for groups B and C compared to group A. This is in line with findings in literature [70, 71] and seems to confirm that subjective presence response increases with greater levels of immersion. The fact that the greater difference was found between low and medium vividness levels, but not between medium and high, could be explained by the greater transitions of textures and geometric complexity. Specifically, in low-medium transition there was a jump from no textures to limited object textures and from simple geometric shapes to 3D models; while the high level was created by increasing the textures resolution and the complexity of 3D models. It appears that the transition from nothing to something (e.g., some textures vs. no texture) had a more profound effect on the way users perceive the environments and subjectively represent their sensation of presence.

No statistically significant effect of vividness was found on the variables motivation, concentration, stress and sleepiness. However, we could notice that both motivation and concentration tended to increase with greater level of vividness. Participants in higher vividness groups reported to feel more motivated and focused on the task during NF training compared to participants in lower groups. Even though not significant, we found these results encouraging, because they are consistent with NF training performance results and corroborate the assumption that a more vivid training scenario should increase interest and motivation [10, 15], consequently leading to an improved NF training performance [14, 48]. Furthermore, the results relative to sleepiness variable showed that participants in group A tended to feel drowsier during NF training compared to participants in group B and C: the low vivid training scenario made participants feel bored and lose interest in the NF training. As previously hypothesized, this could explain the fact that participants in group A did not achieve successful results in UA modulation.

The analysis of perceived competence and workload data showed no statistically significant difference between groups. The results of perceived competence were comparable between groups, suggesting that the vividness of the training scenario had no effect on the sense of mastery in executing the NF task. Although non-significant, there was an increasing trend in workload results, showing that the perceived workload increased with greater level of vividness.

Effect of UA neurofeedback on working memory

There is evidence in literature that UA enhancement training has the effect of improving working memory performance [17, 40, 41]. The hypothesis that an increase in UA activity is correlated with increasing WM performance seemed to be confirmed by the findings of this study. In fact, a statistically significant correlation was found between the improvement of performance in a 3-back test and the enhancement of NF performance across NF training sessions. Specifically, it has been shown that an increase in the UA ratio or in the percentage of time above threshold across NF sessions corresponded to an enhance in the Distractor accuracy in the 3-back test.

6 Conclusions

In recent years, researchers provided evidence of the positive impact that the immersive properties of VR training scenarios have on Neurofeedback training. This thesis is an initial attempt to investigate how distinct immersion-related factors contribute to the improvement of NF performance.

Specifically, the main objective of the present study was to analyse the effect of vividness on NF training outcome and user experience. To this end, we designed a between-subjects experiment in which participants received NF training to enhance Upper Alpha amplitude and we compared the results obtained through three different NF training scenarios with increasing level of vividness (i.e., low, medium, high). Furthermore, as a secondary objective, we assessed the efficacy of the UA enhancement protocol in improving working memory performance.

In summary, results concerning NF training performance revealed that participants' ability to modulate UA amplitude during a training session tended to increase with increased vividness. Although, statistical analysis revealed that this trend was only marginally significant. No effect of vividness was found on NF transfer ability. Moreover, NF performance did not improve across training sessions, suggesting that a longer training period is necessary to detect progress over time.

Regarding user experience, perceived competence and level of stress resulted not to be affected by vividness. In contrast, we found a consistent upward trend for motivation and concentration with increased vividness, indicating that a highly vivid training scenario made NF users feel more interested and focused on NF training compared to lower vivid scenarios. Conversely, a low vivid scenario was perceived as more boring, and this may have hindered subjects' abilities to successfully modulate brain activity. The influence of vividness on these variables was not effective enough to result in significant differences. However, given the small sample size, we considered these results encouraging because, even though not significant, we found the same consistent trends in the analysis of different variables.

Finally, a significant positive correlation was found between UA enhancement across sessions and the improvement of working memory performance.

In conclusion, the results of this study suggest that NF training scenarios can be improved by the design of virtual environments highly vivid and realistic, since vividness has been shown to be a factor that influences NF performance as well as variables related to subjective user experience, such as motivation, concentration and sleepiness. Furthermore, the results show that upper alpha NF training is an effective procedure to improve working memory performance, confirming the findings of prior studies.

Limitations and future improvements

There are some limitations that need to be accounted for a correct interpretation of these results. First, the sample size of the experimental groups was small, and this affects the statistical significance of the study. As mentioned in the description of the results, inspection of the visual representations of the data is required and has been used throughout this thesis to draw the presented conclusion. However, often the statistical analysis failed to detect significant differences between the experimental conditions. It is possible that the results did not reach significance because of the small sample size. Hence, further studies with a greater number of participants are needed to confirm our findings.

Another limit is represented by the fact that a short period of training (5 sessions) was conducted. There is evidence in literature that a longer NF practice is necessary to detect long-term effects. Therefore, the number of sessions in this study might not have been enough to show significant effects of vividness on NF transfer ability and on the improvement of NF performance across sessions. For future study, it is advisable to adopt a prolonged NF training schedule so that findings can be extended to long-term effects.

Furthermore, future research should consider investigating further immersive factors for effects on NF performance and subjective response measures. Besides vividness, other variables such as extensiveness, proprioceptive matching, and inclusiveness could be examined. It is possible that another factor could have significant effects on NF outcomes which vividness did not.

References

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller and T. M. Vaughan, “Brain–computer interfaces for communication and control,” *Clinical Neurophysiology,* vol. 113, no. 6, pp. 767-791, 2002.

[2] J. van Erp, F. Lotte and M. Tangermann, “Brain-computer interfaces: beyond medical applications,” *IEEE Computer,* vol. 45, no. 4, pp. 26-34, 2012.

[3] D. J. McFarland and J. R. Wolpaw, “Brain-computer interfaces for communication and control,” *Communications of the ACM,* vol. 54, no. 5, pp. 60-66, 2011.

[4] S. Sreedharan, R. Sitaram, J. S. Paul and C. Kesavadas, “Brain-Computer Interfaces for Neurorehabilitation,” *Critical Reviews in Biomedical Engineering,* vol. 41, no. 3, pp. 269-279, 2013.

[5] 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,* vol. 28, no. 3, pp. 167-172, 2008.

[6] 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,* vol. 11, no. 51, 2017.

[7] A. Vourvopoulos, A. Ferreira and S. Bermudez i Badia, “NeuRow: An Immersive VR Environment for Motor-Imagery Training with the Use of Brain-Computer Interfaces and Vibrotactile Feedback,” *Proceedings of the 3rd International Conference on Physiological Computing Systems - Volume 1: PhyCS,* pp. 43-53, 2016.

[8] G. Muller-Putz and G. Pfurtscheller, “Control of an Electrical Prosthesis With an SSVEP-Based BCI,” *IEEE Transactions on Biomedical Engineering,* vol. 55, no. 1, pp. 361-364, 2008.

[9] B. Dobkin, “Brain-computer interface technology as a tool to augment plasticity and outcomes for neurological rehabilitation,” *The Journal of physiology,* vol. 579, no. 3, pp. 637-426, 2007.

[10] S. E. Kober, J. L. Reichert, D. Schweiger, C. Neuper and G. Wood, “Does Feedback Design Matter? A Neurofeedback Study Comparing Immersive Virtual Reality and Traditional Training Screens in Elderly,” *International Journal of Serious Games,* vol. 4, no. 3, 2017.

[11] J. H. Gruzelier, “EEG-neurofeedback for optimising performance. I: A review of cognitive and affective outcome in healthy participants,” *Neuroscience & Biobehavioral Reviews,* vol. 44, pp. 124-141, 2014.

[12] F. Wan, W. Nan, M. I. Vai and A. Rosa, “Resting alpha activity predicts learning ability in alpha neurofeedback,” *Frontiers in Human Neuroscience,* vol. 8, 2014.

[13] S. E. Kober, M. Witte, M. Ninaus, C. Neuper and G. Wood, “Learning to modulate one's own brain activity: the effect of spontaneous mental strategies,” *Frontiers in Human Neuroscience,* vol. 7, no. 695, 2013.

[14] S. Kleih, F. Nijboer, S. Halder and A. Kübler, “Motivation modulates the P300 amplitude during brain–computer interface use,” *Clinical Neurophysiology,* vol. 121, no. 7, pp. 1023-1031, 2010.

[15] S. E. Kober, J. L. Reichert, D. Schweiger, C. Neuper and G. Wood, “Effects of a 3D Virtual Reality Neurofeedback Scenario on User Experience and Performance in Stroke Patients,” in *GALA Conference 2016*, 2016.

[16] M. Slater and S. Wilbur, “A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments,” *Presence: Teleoperators and Virtual Environments,* vol. 6, no. 6, pp. 603-616, 1997.

[17] B. Zoefel, R. J. Huster and C. S. Herrmann, “Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance,” *NeuroImage,* vol. 54, no. 2, pp. 1427-1431, 2011.

[18] M. Teplan, “Fundamental of EEG Measurement,” *Measurement Science Review,* vol. 2(2), 2002.

[19] D. A. Kaiser, “Basic Principles of Quantitative EEG,” *Journal of Adult Development,* vol. 12, no. 2-3, pp. 99-104, 2005.

[20] A. Heraz and C. Frasson, “Towards a Brain-Sensitive Intelligent Tutoring System: Detecting Emotions from Brainwaves,” *Adv. Artificial Intellegence,* vol. 2011, pp. 1-13, 2011.

[21] 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,* vol. 8, no. 1008, 2014.

[22] H. Marzbani, H. Marateb and M. Mansourian, “Neurofeedback: A Comprehensive Review on System Design, Methodology and Clinical Applications,” *Basic and Clinical Neuroscience Journal,* vol. 7, no. 2, 2016.

[23] J. Gruzelier and T. Egner, “Critical validation studies of neurofeedback,” *Child and Adolescent Psychiatric Clinics of North America,* vol. 14, no. 1, pp. 83-104, 2005.

[24] G. H. Klem, H. O. Lüders, H. H. Jasper and C. Elger, “The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology,” *Electroencephalography and clinical neurophysiology. EEG Supplement,* vol. 52, pp. 3-6, 1999.

[25] H. Heinrich, H. Gevensleben and U. Strehl, “Annotation: Neurofeedback - Train your brain to train behaviour,” *Journal of child psychology and psychiatry, and allied disciplines,* vol. 48, no. 1, pp. 3-16, 2007.

[26] E. Peper and F. Shaffer, “Biofeedback History: An Alternative View,” *Biofeedback,* vol. 38, no. 4, pp. 142-147, 2010.

[27] K. Reiter, S. B. Andersen and J. Carlsson, “Neurofeedback Treatment and Posttraumatic Stress Disorder,” *The Journal of Nervous and Mental Disease,* vol. 204, no. 2, pp. 69-77, 2016.

[28] 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,* vol. 43, no. 2, pp. 199-144, 2012.

[29] J.-R. Wang and S. Hsieh, “Neurofeedback training improves attention and working memory performance,” *Clinical Neurophysiology,* vol. 124, no. 12, pp. 2406-2420, 2013.

[30] T. Egner and J. Gruzelier, “Learn self-regulation of EEG frequency components affects attention and event-related brain potentials in humans,” *Neuroreport,* vol. 12, no. 18, pp. 4155-4159, 2002.

[31] T. Egner and J. Gruzelier, “EEG Biofeedback of low beta band components: Frequency-specific effects on variables of attention and event-related brain potentials,” *Clinical neurophysiology,* vol. 115, no. 1, pp. 131-139, 2004.

[32] H. S. J. Tiitinen, K. Reinikainen, K. Alho, J. Lavikainen and R. Näätänen, “Selective attention enhances the auditory 40-Hz transient response in humans,” *Nature,* vol. 364, no. 6432, 1993.

[33] A. W. Keizer, M. Verschoor, R. S. Verment and B. Hommel, “The effect of gamma enhancing neurofeedback on the control of feature bindings and intelligence measures,” *International Journal of Psychophysiology,* vol. 75, no. 1, pp. 25-32, 2010.

[34] O. Jensen, J. Gelfand, J. Kounios and J. Lisman, “Oscillations in the Alpha Band (9-12 Hz) Increase with Memory Load during Retention in a Short-term Memory Task,” *Cerebral Cortex,* vol. 12, no. 8, pp. 877-882, 2002.

[35] 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,* vol. 30, no. 1, pp. 1-10, 2005.

[36] T. Egner and J. H. Gruzelier, “Ecological validity of neurofeedback: modulation of slow wave EEG enhances musical performance,” *Cognitive Neuroscience and Neuropsychology,* vol. 14, no. 9, pp. 1221-1224, 2002.

[37] D. J. Vernon, “Can Neurofeedback Training Enhance Performance? An Evaluation of the Evidence with Implications for Future Research,” *Applied Psychophysiology and Biofeedback,* vol. 30, no. 4, pp. 347-364, 2005.

[38] D. J. Prime and P. Jolicoeur, “Mental Rotation Requires Visual Short-term Memory: Evidence from Human Electric Cortical Activity,” *Journal of Cognitive Neuroscience,* vol. 22, no. 11, pp. 2437-2446, 2010.

[39] B. Aben, S. Stapert and A. Blokland, “About the Distinction between Working Memory and Short-Term Memory,” *Frontiers in Psychology,* vol. 3, no. 301, 2012.

[40] C. Escolano, M. Aguilar and J. Minguez, “EEG-based Upper Alpha Neurofeedback Training Improves Working Memory Performance,” *Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* pp. 2327-2330, 2011.

[41] W. Nan, J. Rodrigues, J. Ma, X. Qu, F. Wan, M. P. P. Mak, M. Vai and A. Rosa, “Individual alpha neurofeedback training effect on short term memory.,” *International Journal of Psychophysiology,* vol. 86, no. 1, pp. 83-87, 2012.

[42] J.-J. Hsueh, T.-S. Chen, J.-J. Chen and F.-Z. Shaw, “Neurofeedback training of EEG alpha rhythm enhances episodic and working memory,” *Human Brain Mapping,* vol. 37, no. 7, pp. 2662-2675, 2016.

[43] S. Kober, D. Schweiger, M. Witte, J. Reichert, P. Grieshofer, C. Neuper and G. Wood, “Specific effects of EEG based neurofeedback training on memory functions in post-stroke victims,” *Journal of NeuroEngineering and Rehabilitation,* vol. 12, no. 1, 2015.

[44] M. Akin, “Comparison of wavelet transform and FFT methods in the analysis of EEG signals,” *Journal of Medical Systems,* vol. 26, no. 3, pp. 241-247, 2002.

[45] T. Hinterberger, N. Neumann, M. Pham, A. Kübler, A. Grether, N. Hofmayer, B. Wilhelm, H. Flor and N. Birbaumer, “A multimodal brain-based feedback and communication system,” *Experimental Brain Research,* vol. 154, no. 4, pp. 521-526, 2004.

[46] 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,* vol. 41, no. 1, pp. 27-37, 2016.

[47] K. Subramaniam and S. Vinogradov, “Improving the neural mechanisms of cognition through the pursuit of happiness,” *Frontiers in human neuroscience,* vol. 7, 2013.

[48] S. C. Kleih and A. Kübler, “Empathy, motivation, and P300 BCI performance,” *Frontiers in Human Neuroscience,* vol. 7, 2013.

[49] 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,* vol. 7, 2013.

[50] S. Halder, B. Varkuti, M. Bogdan, A. Kübler, W. Rosenstiel, R. Sitaram and N. Birbaumer, "Prediction of brain-computer interface aptitude from individual brain structure," *Frontiers in Human Neuroscience,* vol. 7, 2013.

[51] S. Enriquez-Geppert, R. Huster and C. Herrmann, "Boosting brain functions: Improving executive functions with behavioural training, neurostimulation, and neurofeedback," *International Journal of Psychophysiology,* vol. 88, no. 1, pp. 1-15, 2013.

[52] N. Lofthouse, L. Arnold, S. Hersch, E. Hurt and R. DeBeus, "A review of neurofeedback treatment for pediatric ADHD," *Journal of Attention Disorders,* vol. 16, no. 5, pp. 351-372, 2012.

[53] A. Vourvopoulos and S. Bermúdez i Badia, "Motor priming in virtual reality can augment motor-imagery training efficacy in restorative brain-computer interaction: a within-subject analysis," *Journal of NeuroEngineering and Rehabilitation,* vol. 13, no. 1, 2016.

[54] W. R. Sherman and A. B. Craig, Understanding Virtual Reality: Interface, Application, and Design, Morgan Kaufmann Publishers Inc., 2002.

[55] L. Freina and A. Canessa, " Immersive vs Desktop Virtual Reality in Game Based Learning," *Conference: 9th European Conference on Games Based Learning (ECGBL),* 2015.

[56] G. Robertson, S. Card and J. Mackinlay, "Three views of virtual reality: nonimmersive virtual reality," *Computer,* vol. 26, no. 2, p. 81, 1993.

[57] M. Slater, P. Khanna, J. Mortensen and I. Yu, "Visual Realism Enhances Realistic Response in an Immersive Virtual Environment," *IEEE Computer Graphics and Applications,* vol. 29, no. 3, pp. 76-84, 2009.

[58] N. Wang and W. Doube, "How real is reality? A perceptually motivated system for quantifying visual realism in digital images," *Proceedings - 2011 International Conference on Multimedia and Signal Processing,* vol. 2, pp. 141-149, 2011.

[59] Y. Toczek, "The influence of visual realism on the sense of presence in virtual environments," Master dissertation, Eindhoven University of Technology, 2017.

[60] A. M. Berger and E. J. Davelaar, "Frontal alpha oscillations and attentional control: A virtual reality neurofeedback study," *Neuroscience,* vol. 378, pp. 189-197, 2017.

[61] J. Gruzelier, A. Inoue, R. Smart, A. Steed and T. Steffert, "Acting performance and flow state enhanced with sensory-motor rhythm neurofeedback comparing ecologically valid immersive VR and training screen scenarios," *Neuroscience Letters,* vol. 480, no. 2, pp. 112-116, 2010.

[62] B. H. Cho, S. Kim, D. Shin, J. Han Lee, S. Min Lee, I. Kim and S. Kim, "Neurofeedback Training with Virtual Reality for Inattention and Impulsiveness," *Cyberpsychology & behavior,* vol. 7, no. 5, pp. 519-526, 2004.

[63] 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,* vol. 9, no. 677, 2015.

[64] D. Aeschbach, T. T. Postolache, L. Sher, J. R. Matthews, M. A. Jackson and T. A. and Wehr, “Evidence from the waking electroencephalogram that short sleepers live under higher homeostatic sleep pressure than long sleepers,” *Neuroscience,* vol. 102, no. 3, p. 493–502, 2001.

[65] H. Laufs, K. Krakow, P. Sterzer, E. Eger, A. Beyerle, A. Salek-Haddadi and A. Kleinschmidt, “Electroencephalographic signatures of attentional and cognitive default modes in spontaneous brain activity at rest,” *Proceedings of the National Academy of Sciences of the United States of America,* vol. 100, no. 19, pp. 11053-11058, 2003.

[66] C. Kothe, “Lab streaming layer (LSL),” 2014. [Online]. Available: https://github.com/sccn/labstreaminglayer.

[67] A. Gonçalves and S. Bermúdez, “KAVE: Building Kinect Based CAVE Automatic Virtual Environments, Methods for Surround-Screen Projection Management, Motion Parallax and Full-Body Interaction Support,” *Proceedings of the ACM on Human-Computer Interaction,* vol. 2, no. 10, pp. 1-15, 2018.

[68] T. Dempster and D. Vernon, “Identifying Indices of Learning for Alpha Neurofeedback Training,” *Applied psychophysiology and biofeedback,* vol. 34, pp. 309-328, 2009.

[69] M. Slater, M. Usoh and A. Steed, “Depth of presence in virtual environments,” *Presence: Teleoperators and Virtual Environments,* vol. 3, no. 2, pp. 130-144, 1994.

[70] M. Usoh, E. Catena, S. Arman and M. Slater, “Using Presence Questionnaires in Reality,” *Presence: Teleoperators and Virtual Environments,* vol. 9, no. 5, pp. 497-503, 2000.

[71] M. Slater, M. Usoh and A. Steed, “Taking steps: the influence of a walking technique on presence in virtual reality,” *ACM Transactions on Computer-Human Interaction,* vol. 2, no. 3, pp. 201-219, 1995.

[72] S. G. Hart, “Nasa-Task Load Index (NASA-TLX); 20 Years Later,” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* vol. 50, no. 9, pp. 904-908, 2006.

[73] L. Ma, L. Chang, X. Chen and R. Zhou, “Working memory test battery for young adults: Computerized working memory assessment,” *PLOS ONE,* vol. 12, no. 3, 2017.

[74] U. León-Domínguez, J. F. Martín-Rodríguez and J. León-Carrión, “Executive n-back tasks for the neuropsychological assessment of working memory,” *Behavioural Brain Research,* vol. 292, pp. 167-173, 2015.

[75] H. Choi, D. Young Lee, E. Seo, M. Kyung Jo, B. Sohn, Y. M. Choe, M. S. Byun, J. W. Kim, S.-G. Kim, J. Choul Yoon, J. Hyeong Jhoo, K. W. Kim and J. Inn Woo, “A Normative Study of

the Digit Span in an Educationally Diverse Elderly Population," *Psychiatry investigation,* vol. 11, no. 1, pp. 39-43, 2014.

APPENDICES

APPENDIX A – Informed consent form

Consent Form for Participation in Research

Study Title: Neurofeedback training in VR

Investigator: Floriana Accoto, Università del Salento,
+39 291721006, floriana.accoto@gmail.com

Investigator: Eng. Afonso Gonçalves, Madeira Interactive Technologies Institute,
+351 291721006, afonso.goncalves@m-iti.org

Investigator: Dr. Athanasios Vourvopoulos, University of Southern California

Supervision: Dr. Sergi Bermudez i Badia, Madeira Interactive Technologies Institute,
+351 291721006, sergi.bermudez@m-iti.org

____

PURPOSE OF THE STUDY

The purpose of this study is to evaluate different types of visual feedback during neurofeedback training and its effect on working memory.

PROCEDURE

You have been invited to participate in a neuroscientific experiment. Before the experiment you will be asked to provide basic information.

For the experiment, 5 sessions are required on consecutive days (excluding weekends). In the first and last session you will be required to undergo 2 working memory tests. For the neurofeedback procedure, first you are going to use a Brain Computer Interface (BCI), a non-invasive device to measure electric activity patterns of your brain. After verifying the connections, to be sure that the position of the electrodes of the BCI system are in the correct position, you will be placed in a surround screen immersive VR (CAVE) and given a set of instructions to carry out. During this task we will record electroencephalographic (EEG) signals. You must try to execute the task as well as possible in the assigned period (20 min). Finally, a set of questionnaires will be supplied to be fill out after each session. The experimental data will be processed in such a way that your anonymity will be preserved.

INCLUSION CRITERIA

You are eligible for participation if you: are over 18 years old; can understand English; and have no past of brain injuries and no neurological disorders.

RISKS

The risk associated with participation in this study are no greater than those ordinarily encountered in daily life. The EEG electrodes are superficial and DO NOT have any risk for your health. The interaction requires executing mental tasks using a BCI on your head. You may experience fatigue and/or headache in some sessions.

BENEFITS

There are no personal benefits for your participation in the study. The results will contribute to the better understanding of neurofeedback learning mechanisms and its effect on working memory.

CONFIDENTIALITY

By participating in the study, you understand and agree that the researchers may be required to disclose your consent form, data and other personally identifiable information as required by law, regulation, subpoena or court order. Otherwise, your confidentiality will be maintained in the following manner. Data and information gathered during this study may be used by the researchers and published and/or disclosed by them to others for research purposes. However, your personal information will never be revealed in any publication or dissemination of the research data and/or results.

INFORMED CONSENT

I understand that all information derived from the study “Neurofeedback training in VR” is owned by the responsible research team. I give my consent for anonymous collection of data about me, which will be stored and processed for scientific evaluation. I understand the significance of this information, and any questions I had were answered satisfactorily. I had enough time to decide on my participation in this study. I hereby consent my participation and the collection of information.

___________________________________	_______________
Signature of the participant	Date

___________________________________	_______________
Signature of the researcher	Date

APPENDIX B

CHARACTERIZATION

Participant ID Code: ________________ Date: ___/____/201__

Name: ____

Date of Birth: ___/____/______ Gender: Male ☐, Female ☐

Video game experience:

None					Some				A lot

APPENDIX C – SUS (Slater-Usoh-Steed) Presence questionnaire

1. Please rate your sense of being there in the computer generated world...
In the computer generated world I had a sense of "being there"...

Not at all						*Very much*

2. To what extent were there times during the experience when the computer generated world became the "reality" for you, and you almost forgot about the "real world" outside?
There were times during the experience when the computer generated world became more real or present for me compared to the "real world"...

At no time						*Almost all the time*

3. When you think back about your experience, do you think of the computer generated world more as something that you saw, or more as somewhere that you visited?
The computer generated world seems to me to be more like...

Images that I saw						*Somewhere that I visited*

4. During the time of the experience, which was strongest on the whole: your sense of being in the computer generated world, or of being in the real world of the laboratory?
I had a stronger sense of being in . . .

The real world of the laboratory.						*The virtual reality computer generated world.*

5. During the time of the experience, did you often think to yourself that you were actually just standing in an office wearing a helmet or did the computer generated world overwhelm you?
During the experience I often thought that I was really standing in the lab wearing a helmet . . .

Most of the time I realized I was in the lab.						*Never because the virtual computer generated world overwhelmed me.*

APPENDIX D – After session survey

Please select the column which best describes how often you felt each of the following states/sensations during the session.

1. **Sleepiness**

1	2	3	4	5
Never				*Constantly*

2. **Motivation**

1	2	3	4	5
Never				*Constantly*

3. **Concentration**

1	2	3	4	5
Never				*Constantly*

4. **Stress**

1	2	3	4	5
Never				*Constantly*

APPENDIX E – IMI (Intrinsic Motivation Inventory) Perceived competence scale

Participant ID Code: __________ Session: __________ Date: ______________

For each of the following statements, please indicate how true it is for you, using the following scale:

1	2	3	4	5	6	7
Not at all			Somewhat true			Very true

1. **I think I am pretty good at this activity.**

1	2	3	4	5	6	7
Not at all			*Somewhat true*			*Very true*

2. **I think I did pretty well at this activity, compared to other students.**

1	2	3	4	5	6	7
Not at all			*Somewhat true*			*Very true*

3. **After working at this activity for awhile, I felt pretty competent.**

1	2	3	4	5	6	7
Not at all			*Somewhat true*			*Very true*

4. **I am satisfied with my performance at this task.**

1	2	3	4	5	6	7
Not at all			*Somewhat true*			*Very true*

5. **I was pretty skilled at this activity.**

1	2	3	4	5	6	7
Not at all			*Somewhat true*			*Very true*

6. **This was an activity that I couldn’t do very well.**

1	2	3	4	5	6	7
Not at all			*Somewhat true*			*Very true*

APPENDIX F – NASA TLX Workload questionnaire

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand How mentally demanding was the task?

Very Low Very High

Physical Demand How physically demanding was the task?

Very Low Very High

Temporal Demand How hurried or rushed was the pace of the task?

Very Low Very High

Performance How successful were you in accomplishing what you were asked to do?

Perfect Failure

Effort How hard did you have to work to accomplish your level of performance?

Very Low Very High

Frustration How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low Very High