The Effect of Neurofeedback Training in CAVE-VR for Enhancing Working Memory

Floriana Accoto, Athanasios Vourvopoulos, Afonso Gonçalves, Teresa Bucho, Gina Caetano, Patrícia Figueiredo, Lucio De Paolis, and Sergi Bermudez i Badia

Abstract In recent years, increasing evidence of the positive impact of Virtual Reality (VR) on neurofeedback training has emerged. The immersive properties of VR training scenarios have been shown to facilitate neurofeedback learning while leading to cognitive enhancements such as increased working memory performance. However, in the design of an immersive VR environment, there are several covariates that can influence the level of immersion. To date, the specific factors which contribute to the improvement of neurofeedback performance have not yet been clarified. This research aims to investigate the effects of vividness in a Cave automatic virtual environment (CAVE-VR) on neurofeedback training outcome, and to assess the effect on working memory performance. To achieve this, we recruited 21 participants, exposed to neurofeedback training inside a CAVE-VR environment. Participants were divided

F. Accoto (✉) · L. De Paolis
University of Salento, Lecce, Italy
e-mail: floriana.accoto@studenti.unisalento.it

L. De Paolis
e-mail: lucio.depaolis@unisalento.it

A. Vourvopoulos (✉) · T. Bucho · P. Figueiredo
Instituto Superior Técnico, Lisbon, Portugal
e-mail: athanasios.vourvopoulos@tecnico.ulisboa.pt

T. Bucho
e-mail: teresa.bucho@gmail.com

P. Figueiredo
e-mail: patricia.figueiredo@tecnico.ulisboa.pt

A. Gonçalves · S. B. Badia
Universidade da Madeira, Funchal, Portugal
e-mail: afonso.goncalves@m-iti.org

S. B. Badia
e-mail: sergi.bermudez.badia@gmail.com

G. Caetano
Universidade de Lisboa, Lisbon, Portugal
e-mail: caetanogina@gmail.com

T. Dingler and E. Niforatos (eds.), *Technology-Augmented Perception and Cognition*, Human–Computer Interaction Series, https://doi.org/10.1007/978-3-030-30457-7_2

into three experimental groups, each of which received feedback in a different neurofeedback training scenario with increasing level of vividness (i.e., low, medium, high) while also assessing the effect of neurofeedback on working memory performance. Current findings show that highly vivid feedback in CAVE-VR results in increased neurofeedback performance. In addition, highly vivid training scenarios had a positive effect on user's motivation, concentration, and reduced boredom. Finally, current results corroborate the efficacy of the neurofeedback enhancement protocol in CAVE-VR for improving working memory performance.

1 Introduction

Technology has undoubtedly impacted our cognition and the way our mental actions like attention, problem-solving, and working memory is formed (Dingler et al. 2016). One way of improving behavior and cognition is by controlling certain brain signals in a closed feedback loop called neurofeedback (Enriquez-Geppert et al. 2013; Huster and Herrmann 2017; Zoefel et al. 2011). Neurofeedback is a form of biofeedback that self-regulates brain activity, with the aim of improving mental states or processes (Gruzelier 2014). During neurofeedback training, the user receives real-time feedback of one's own electrical brain activity acquired through electroencephalographic (EEG) signals. Specific components of the EEG signal (or EEG bands) are extracted in real-time from the user and presented via visual or auditory feedback. This enables the user to consciously perceive their own 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 forms associations between specific mental states and desired brain activation patterns (Kober et al. 2017). It has been shown that voluntary modulation of specific EEG bands leads to improvements in behavior and cognition (Gruzelier 2014). Moreover, studies of working memory training have shown that specifically designed mental exercises (i.e., cognitive training paradigms) could be used to enhance cognitive performance (Morrison and Chein 2011). Working memory refers to the ability of the brain to provide temporary storage and manipulation of information, necessary for cognitive tasks such as language, learning, and reasoning (Baddeley 1992). Although neurofeedback has demonstrated benefits in many aspects, a critical issue in neurofeedback studies is that not all subjects showed satisfactory learning ability to regulate electrical brain activity (Wan et al. 2014): about 15–30% of neurofeedback users cannot attain control over their brain signals (Kober et al. 2017). There are different attempts to explain this phenomenon (Kober et al. 2013), but the specific 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 neurofeedback performance. For instance, motivation of the user turned out to play an important role (Kleih et al. 2010). It should also be considered that to obtain cognitive or behavioral improvements, a large number of repeated neurofeedback training sessions are mandatory, and this can induce fatigue to the user. Furthermore, neurofeedback practice requires users to stay focused and concentrated on the neurofeedback task over a long training period (Kober et al. 2016). 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 (Yan et al. 2008), leading to decreased motivation, interest, concentration, and finally to a lower neurofeedback performance and success rate (Kleih et al. 2010). Hence, an increasing number of recent neurofeedback studies have utilized Virtual Reality (VR) in their feedback design (Kober et al. 2017). In spite of that, still little is known about the effectiveness of VR-based neurofeedback training and the effect it might have on working memory performance. 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 when compared with traditional 2D feedback modalities (Kober et al. 2016). Moreover, concerning the effect of vividness on neurofeedback training performance, the literature suggests that the immersive properties of virtual environments are effective in cognitive training (Cho et al. 2002). To address current limitations. The objective of this study is twofold. First, to investigate the effect of vividness in VR in terms of neurofeedback performance and subjective user experience, and second, to assess the effect of upper-Alpha neurofeedback training on working memory performance. To achieve this, we designed three virtual environments, with different level of vividness in a CAVE-VR environment. 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 neurofeedback protocol was used, in which participants learned to increase their brain activity in the upper-Alpha frequency band voluntarily. Alpha band training is one of the most commonly used protocols since upper-Alpha has consistently been shown to be correlated with cognitive performance (Zoefel et al. 2011). Hence, to promote interpretability of neurofeedback study results, a similar protocol was selected.

2 Background

Here we give a brief background on brain electrical activity referred to as brain oscillations, using electroencephalography (EEG) and the way it is utilized in neurofeedback. Moreover, we present the importance of immersion and vividness in neurofeedback through the use of Virtual Reality (VR). Finally, we present the impact of neurofeedback training in working memory.

2.1 EEG and Brain Waves

The root of neurofeedback and the related field of electroencephalography can be traced back to Hans Berger, a German psychiatrist who recorded the first human electroencephalogram (EEG) in 1924 (Berger 1933). EEG is a noninvasive recording method to measure the electrical activity of the brain. The human brain contains 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. More specifically, the EEG results from the synchronous firing of a specific type of neurons in the cortex, known as pyramidal neurons (Teplan 2002). 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 the raw brainwave signal is recorded in digital format, it can be transformed into brainwave data, by extracting information about the extent of specific frequency bands that are contributing to the overall power of a waveform. These patterns of electrical activity are split into different brain waves based on their frequencies, that represent how fast the waves oscillate, as measured by the number of waves per second or Hertz (Hz). With EEG, researchers had the opportunity of identifying the relationship between brain oscillations and different mental or behavioral states. Berger himself was the first to describe a predominant emerging rhythm of the human brain. This rhythm increased in power between 7.8 and 13 Hz when subjects had their eyes closed and decreased 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 represent fundamental activity at the cortical level (Teplan 2002). In present days, these brain waves are referred to as "Alpha waves", also known as "Berger 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) (Ros et al. 2014). 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, the sensorimotor rhythm (SMR) frequency band (13–15 Hz) is related to motor tasks (even during movement imagination) and entitled as low Beta (Marzbani et al. 2016). The Alpha rhythm is usually divided into two subsets: lower Alpha in the range of 8–10 Hz and upper-Alpha in the range of 10–12 Hz (Marzbani et al. 2016). It is important to note that all of the traditional frequency bands are always present across the scalp, but which is the most prevalent depends on the task being undertaken by the individual and the scalp location in question. In general, the more prevalent the higher frequency bands are, 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 (Marzbani et al. 2016). 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 (Gruzelier and Egner 2005).

2.2 Neurofeedback Training

Neurofeedback is part of a broader group of biofeedback applications, all of which have the goal of facilitating the self-regulation of physiological functions with the purpose of normalizing them in clinical populations or optimizing 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 (Heinrich et al. 2007). Operant conditioning is a method of learning that occurs through rewards and punishments for behavior. Through operant conditioning, an individual makes an association between a given behavior and a consequence: positive consequences increase the likelihood of the behavior, whereas negative consequences decrease it (Huster and Herrmann 2017). 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 (Peper and Shaffer 2010). 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 increased ability to associate the 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. Joseph Kamiya was the first researcher to demonstrate the human's ability to control 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 (Hanslmayr et al. 2005; Heinrich et al. 2007; Zoefel et al. 2011). 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: (1) as a therapeutic tool to normalize neurological patients' deviating brain activity, in order to influence symptoms; (2) as so-called peak-performance training to enhance cognitive performance in healthy participants (Huster and Herrmann 2017).

2.3 Neurofeedback Training Efficacy

Undoubtedly, controlling brain activity is an ability that can be learned. 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 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 (Huster and Herrmann 2017). Learner specific aspects such as positive mood states (Subramaniam and Vinogradov 2013), motivation (Kleih and Kübler 2013; Kleih et al. 2010), focus of control (Witte et al. 2013), 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 (Halder et al. 2013). 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 neurofeedback protocol (Enriquez-Geppert et al. 2013). The duration of sessions applied in different studies can vary within a range of 30–60 min. The number of neurofeedback sessions can differ from 5 (Escolano et al. 2011; Zoefel et al. 2011) to more than 40 (Lofthouse et al. 2012). 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 among studies. Sometimes several frequencies are trained simultaneously, such 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. Late research on the impact of the type of feedback showed that auditory feedback may be as effective as the more commonly used visual feedback (Bucho et al. 2019).

All the aspects mentioned above 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 neurofeedback protocol. In particular, researchers recently started to focus on the effects that feedback design can have on neurofeedback 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 neurofeedback learning, it is crucial that the feedback is engaging and attractive. For this reason, an increasing number of recent neurofeedback studies use Virtual Reality based feedback designs (Kober et al. 2017; Vourvopoulos and Bermúdez i Badia 2016), showing that VR is more effective than traditional modalities. We will describe the results of these studies in the next section.

2.4 Neurofeedback Training in Working Memory

As mentioned previously, the term working memory refers to the temporary storage and manipulation of the information necessary for complex cognitive tasks (Baddeley 1992). The definition of working memory 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 (Aben et al. 2012).

Different studies proved the hypothesis that neurofeedback training in the upper-Alpha sub-band (10–12 Hz) can lead to memory performance enhancement. Hanslmayr et al. (2005) showed that only those subjects who were able to increase their upper-Alpha power performed better on mental rotations after neurofeedback training, showing that training success (extent of neurofeedback training-induced increase in upper-Alpha power) was positively correlated with the improvement in cognitive performance (Hanslmayr et al. 2005). Similarly, the impact of upper-Alpha neurofeedback training on cognitive abilities was assessed by Zoefel et al. (2011). The expectation of an enhancement of cognitive performance was confirmed when their study participants in the neurofeedback training group obtained an increase in the upper-Alpha activity and the increase in performance of mental rotations (rotation of mental representations of objects) was significantly larger for the neurofeedback training group than for the control group. Since mental rotation is an ability that involves working memory (Prime and Jolicoeur 2009), these results suggest that upper-Alpha neurofeedback has a positive effect on working memory. The specific effect of upper-Alpha neurofeedback training on working memory was further investigated by Escolano et al. (2011). Their experiment consisted of five neurofeedback sessions, during which participants learned to increase their upper-Alpha amplitude as described in previous sections. Results show that participants in the neurofeedback group obtained an increase in the upper-Alpha activity, as well as a significant enhancement in memory performance compared to the control group. In 2012, Nan et al. (2015), proposed the use of Alpha neurofeedback to improve short-term memory performance. In this case, the neurofeedback 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 neurofeedback training and short-term memory performance was significantly enhanced by neurofeedback 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 et al. (2016) showed that subjects had a progressive significant increase in the Alpha amplitude following neurofeedback training, where the 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 neurofeedback training was not limited to the upper-Alpha sub-band, but on the whole Alpha band. Nonetheless, results are consistent with prior studies.

2.5 Immersion and Vividness in Virtual Reality

Virtual Reality (VR) 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)" (Sherman and Craig 2002). A concept frequently mentioned in VR is "immersion", which is defined as the perception of being physically present in a nonphysical 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" (Freina et al. 2015). For example, immersive VR has been utilized for therapeutic purposes, such as stroke rehabilitation (Vourvopoulos 2019b), investigating ethical decision-making by enacting moral dilemmas (Niforatos et al. 2020), remote learning and virtual tourism by "placing" one in a virtual classroom (Bailenson et al. 2008) or at a virtual location (Marchiori et al. 2017), respectively. 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" (Robertson et al. 1993). Immersion is the measurable feature of VR technology that could make a user feel present in a virtual environment. Slater and Wilbur (1997) have laid out a series of definitions for immersion that will be used in this work. Immersion is what technology delivers from an objective measure and describes the extent to which users can feel part of the environment. The more a system conveys view that preserve fidelity in relation to their corresponding real-world sensory modalities, the more immersive it is. Finally, immersion requires that there is a match between the participant's proprioceptive feedback about body movements—the sense of the relative position of one's own body and movement—and the information vividly generated on the displays with the richness, information content, resolution, and quality of the displays. Vividness is related to the resolution, photo-realism, and visual fidelity of the virtual scenario (Slater and Wilbur 1997). We are particularly interested in studying vividness because of its heavy reliance on visual stimuli. Since virtual environments are mainly 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. For example, Slater and Wilbur (1997, use shadows as a way of vividness manipulation. It has been shown that the scenes where shadows and reflections are present are perceived as more realistic (Slater et al. 2009). Wang and Doube (2011) 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. Further, Toczek (2016), used a texture resolution approach, populating high and Low vividness conditions with objects of varying pixel resolution. Finally, VR settings with increased vividness could help in forming better VR memories with increased performance in recall tasks (Marchiori et al. 2018).

2.6 Neurofeedback in Virtual Reality

Immersive VR is considered to be more effective concerning the acquisition of several abilities and has a positive impact on human performance, compared to non-VR approaches (Vourvopoulos and Bermúdez i Badia 2016; Zimmons and Panter 2003). VR can simulate aspects of everyday life, helping to transfer the learned skills to the real world. For this reason, neurofeedback researchers started to investigate the hypothesis that virtual reality feedback causes an improvement in neurofeedback learning performance in many applications, such as motor recovery and movement re-learning (Hubbard et al. 2017; Vourvopoulos and Bermúdez i Badia 2016). Berger et al. (2018) used neurofeedback to train subjects to increase their level of Alpha amplitude. After five neurofeedback training sessions, they found out that learning slopes were higher in participants who received feedback in the 3D virtual environment, while the training of the 2D group was unsuccessful. On the other hand, Kober et al. (Kober et al. 2016), compared 2D versus 3D feedback with no significant differences. However, regarding user experience, they found that motivation and challenges were higher in the 3D group. In another study by Gruzelier et al. (Gruzelier et al. 2010), the lighting level and the audience noise in the virtual environment 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 theater 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: neurofeedback learning was facilitated (participants learned faster) in the CAVE rendition of the theatrical space versus the computer screen, even though the same auditorium was depicted. Prior studies make comparisons between different types of feedback on different plans, sometimes comparing the same VR content in different settings (e.g., screen vs. CAVE, or screen vs. head-mounted display), sometimes comparing VR contents with traditional non-VR feedback. Even if the VR modalities used in these studies are heterogeneous, in every comparison, the most immersive feedback resulted in more effective training. Specifically, being immersed in a virtual room was better than looking at 2D objects on a screen; visualizing virtual contents with a head-mounted display or in a CAVE was better than a computer screen. Overall, the immersive properties of VR bring advantages in neurofeedback training, either in facilitating neurofeedback learning or increasing motivation and interest.

3 Study

In this study, our target is to investigate the effects of vividness in VR on neurofeedback training outcome and to assess the effect on working memory performance. In order to achieve this, we designed a study by incorporating EEG data acquisition for real-time neurofeedback in a virtual environment delivered through a Cave Automatic Virtual Environment (CAVE).

3.1 Participants

Twenty-one participants (15 male and 6 female), 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) and have no past of brain injuries and no neurological disorders. Finally, all participants signed an informed consent. Participants were quasi-randomly (by order of enrollment in the study) assigned to the three experimental groups. Each group consisted of 7 participants (5 male and 2 female).

3.2 Experimental Conditions

This experiment used three experimental groups based on levels of vividness in VR: 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 using textures, shadows, and reflections (Table 1).

The virtual environments were developed using the Unity game engine (Unity Technologies, San Francisco, CA). 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

Table 1 Differences in the level of vividness

	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

new details, modifying textures and shadows, and using more elaborate 3D models (see Fig. 1).

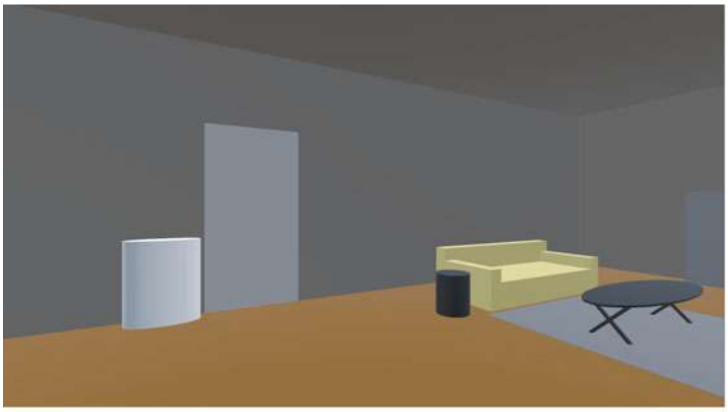
3.3 VR CAVE

For delivering the VR feedback, a CAVE was used. The NeuroRehabLab CAVE has a configuration of three orthogonal walls and a floor (Fig. 2). It uses a Kinect v2 sensor for tracking, thus enabling motion parallax effects and body interaction through the KAVE plugin, developed for the integration of Unity applications with CAVE systems (Gonçalves and Bermúdez 2018). The feedback consisted of an object changing color. We chose this type of feedback because it is often used in literature for upper-Alpha neurofeedback training (Escolano et al. 2011; Hanslmayr et al. 2005; Zoefel et al. 2011). In the Low vividness environment, the object was a cylinder while in the other two environments it was the light from a lamp.

The color scheme ranged from a highly saturated red to a highly saturated blue. The color changed according to the upper-Alpha ratio. Red and blue values indicated the upper-Alpha ratio above or below the THRESHOLD value. Respectively; the full saturated red corresponded to an upper-Alpha ratio greater than or equal to the MAX value; the full saturated blue corresponded to an upper-Alpha ratio less than or equal to the MIN value; the closer the upper-Alpha ratio was to the THRESHOLD, the whiter the color became (Fig. 3).

3.4 EEG Acquisition

For EEG acquisition, the Enobio 8 (Neuroelectrics, Barcelona, Spain) system was used. Enobio, is a wearable, wireless EEG sensor with 8 EEG channels and a triaxial accelerometer, for the recording and visualization of 24-bit EEG data at 500 Hz. The spatial distribution of the electrodes followed the 10–20 system over the locations F3, F4, C3, Cz, C4, Pz, O1, O2 (as shown in Fig. 4). Enobio connects via Bluetooth to the Neuroelectrics Instrument Controller (NIC) software, for visualizing real-time EEG



(a) Low vividness.

(b) Medium vividness.

(c) High vividness.

Fig. 1 VR feedback level of vividness. **a** Low vividness: Low geometric complexity, Smooth surfaces, No shadows/reflections; **b** Medium vividness: Higher geometric complexity, Limited object textures, No shadows/reflections; **c** High vividness: High geometric complexity, High-resolution textures, Soft shadows/reflections

(a) NeuroRehabLab CAVE outline.

(b) participant using real-time NF to control the VR feedback.

Fig. 2 **a** NeuroRehabLab CAVE outline, **b** participant using real-time neurofeedback to control the VR feedback (lamp light) inside the CAVE

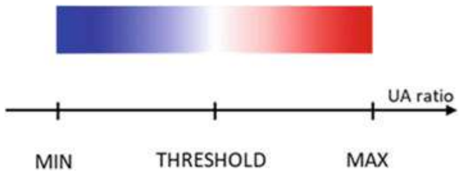


Fig. 3 Color scheme. The color changed from blue to red according to the upper-Alpha (UA) ratio. A highly saturated red corresponded to a high upper-Alpha relative amplitude. Participants' task was to make the color as red as possible, in order to increase their upper-Alpha relative amplitude

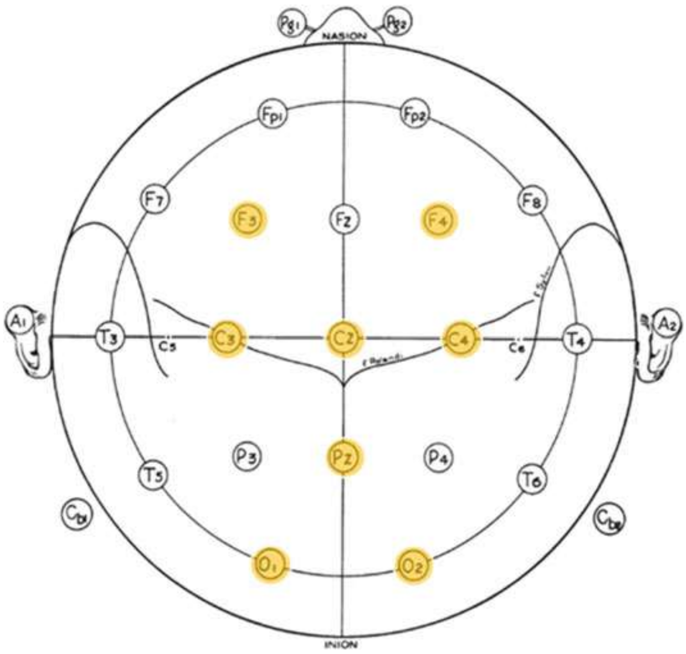


Fig. 4 Electrode location in the 10–20 layout over the locations F3, F4, C3, Cz, C4, Pz, O1, O2

while streaming the data via the Lab Streaming Layer (LSL) protocol[1] to a dedicated computer. LSL was used to send raw data to the OpenVibe platform (Renard et al. 2010) for real-time EEG processing before sending it to the application used for the VR feedback.

[1] https://github.com/sccn/labstreaminglayer/.

3.5 EEG Feedback Parameter

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). The absolute EEG amplitude has large inter-individual differences owing to influences of many factors (such as anatomical and neurophysiological properties of the brain, cranial bone structure, and electrode signal quality) (Nan et al. 2015). Furthermore, additional confounding factors across sessions could result from changes in the time of day (Aeschbach et al. 2001; Vourvopoulos et al. 2017), mood or spontaneous cognitive activity (Laufs et al. 2003). Simple ratios between EEG band amplitudes are commonly used in neurofeedback protocols as relative measures are less sensitive to differences from these uncontrolled factors that modulate EEG amplitudes (Nan et al. 2015). Hence, in order to ensure comparability across participants and sessions, we used the upper-Alpha relative amplitude as a feedback parameter. The upper-Alpha relative amplitude was defined to the analyzed frequency band (upper-Alpha: 10–12 Hz) amplitude relative to the EEG band amplitude from 4 to 30 Hz (Nan et al. 2012; Wan et al. 2014). For brevity, we will refer to the upper-Alpha relative amplitude as an upper-Alpha ratio.

$$UA_{relative\,amplitude} = \frac{UA(10{-}12\mathrm{Hz})_{amplitude}}{EEG(4{-}30\mathrm{Hz})_{amplitude}} \tag{1}$$

3.6 Experimental Design

3.6.1 Protocol

Participants received neurofeedback training session on five consecutive days (except weekend days), from Day 1 to Day 5. On Day 1, before the start of the neurofeedback training, participants signed an informed consent form and provided some basic demographic information (i.e., age, gender). Then they performed three working memory tests (Pre-tests): the digit span test and N-back tests (in the 2-back and 3-back versions). After that, they started the neurofeedback training session. The same neurofeedback procedure was repeated from Day 1–5, 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 neurofeedback session, each participant completed an additional neurofeedback session (Transfer session) and repeated the same working memory tests performed on Day 1 (Post-tests). The transfer session consisted in the same neurofeedback training of the previous sessions, but with a different type of feedback (Fig. 5).

During the neurofeedback 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

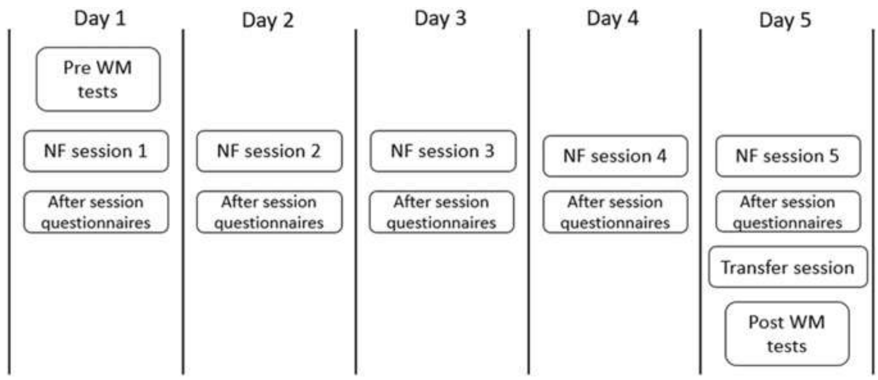


Fig. 5 Overview of the experimental procedure

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. Participants were instructed not to move their head during the neurofeedback session to avoid interference with the signal acquisition.

3.6.2 Procedure

Each session was composed of three blocks: a resting Baseline block and two neurofeedback Training blocks. The Baseline block consisted of a 5-min 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 color of the object was fixed to white, didn't change). Moreover, the Baseline stage before training was used as a familiarization stage to ensure each participant had enough time to explore in the virtual environment.

Next, two Training blocks followed Baseline, with each block lasting 5 min, with a 2 min break in-between (Fig. 6). During the Training blocks, participants tried to modulate their brain activity in the desired direction. They were instructed to make the color as red as possible. No other instruction or suggestion about strategies was given since effective mental strategies vary among individuals (Nan et al. 2012). Moreover, they were not allowed to keep their eyes closed, because Alpha activity naturally increases with eyes closed.

3.7 Subjective Measures

Besides assessing the effect of vividness on neurofeedback learning, we measured the effect it could have on a subjective measure of presence. First, we used the

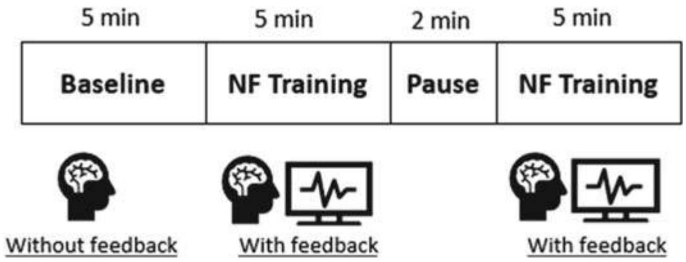


Fig. 6 Neurofeedback session structure

Slater-Usoh-Steed (SUS) questionnaire, that aims at measuring presence in immersive virtual environments (Slater et al. 1994). SUS questionnaire was composed of 5 questions, each on a 1–7 scale where the higher score indicates greater presence. The overall score was computed as the mean value from responses to the five questions.

In addition, we used a component of the Intrinsic Motivation Inventory (IMI) in order to measure the perceived competence. Participants answered 6 questions, rating on a scale from 1 to 7 how much they felt competent during the task. The overall score was the mean of the rating of each question.

Finally, we assessed the perceived workload for every session with the NASA Task Load Index (TLX) (Hart and Staveland 1988). 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). The original version of the NASA-TLX requires a weighting process of the six sub-scales in order to obtain the overall score of the questionnaire. We used one of the most common modifications of the NASA-TLX, the Raw TLX, in which the overall task load index is obtained by averaging the rating of each subscale.

3.8 Working Memory Measures

Two commonly used tests for working memory assessment are the Digit Span test and the N-back test (Ma et al. 2017). For this, 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 (see Table 2).

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

Table 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

3.8.1 Digit Span Test

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 trials 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 both the sequences of a given length correctly. The same holds for the backward task, except for the fact that the participants listen to the sequence of numbers 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. 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 (Choi et al. 2014).

3.8.2 N-Back Task

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 (Fig. 7). 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. We measured performance in the

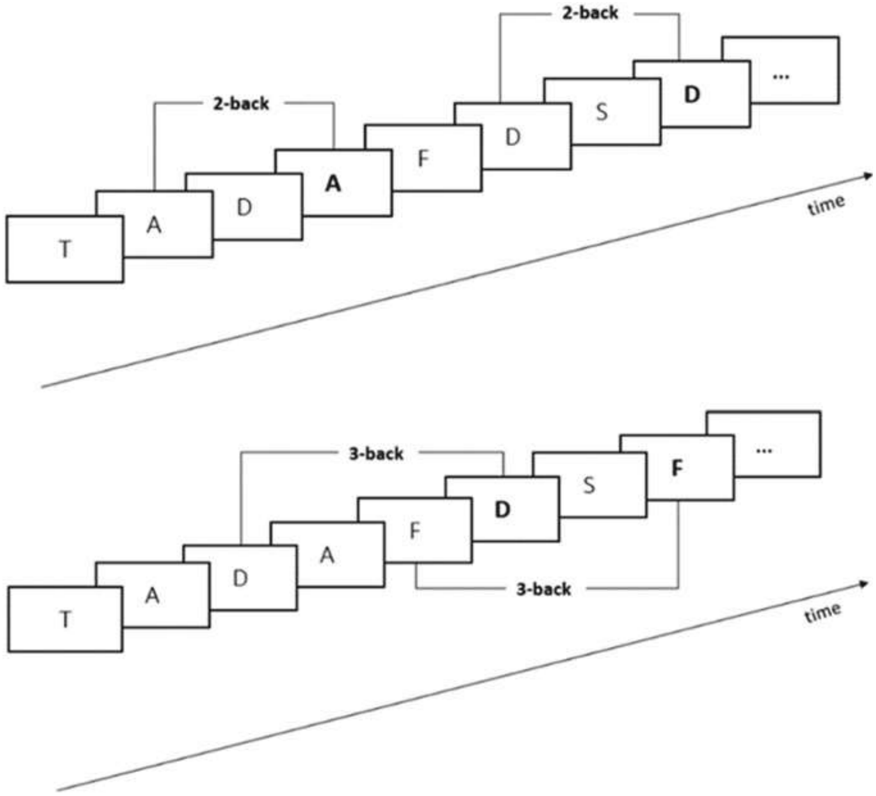


Fig. 7 Examples of 2-back and 3-back tasks. The highlighted letters are Target items, the remaining are Distractors

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 have tested two levels 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:

1. Target accuracy in 2-back
2. Distractor accuracy in 2-back
3. Target accuracy in 3-back
4. Distractor accuracy in 3-back

3.9 Data Analysis

3.9.1 Average UA Relative Amplitude Compared to Baseline

To quantify the changes in the upper-Alpha ratio within a session, we subtracted the average upper-Alpha ratio during the resting baseline from the average upper-Alpha ratio during the training session. This means that any resulting positive value represents enhancement above baseline and any negative value represents falling below the baseline.

For every participant, we computed the average change of the upper-Alpha (UA) relative amplitude (L_1)

$$L_1 = \frac{\sum_{i=1}^{N_{sess}} (mean(UA_{ratio\,training_i}) - mean(UA_{ratio\,baseline_i}))}{N_{sess}} \tag{2}$$

where N_{sess} was the total number of NF sessions, i.e., 5 in our case.

3.9.2 Percentage of Time Above the Threshold

For every session, we considered the percentage of time during which the upper-Alpha ratio was above the threshold, where the threshold was the median value of the upper-Alpha ratio during the corresponding pre-training resting baseline. For every participant, we computed the average percentage of time above threshold (L_2)

$$L_2 = \frac{\sum_{i=1}^{N_{sess}} \%\, time\, above\, threshold_i}{N_{sess}} \tag{3}$$

In order to check how the two measures (upper-Alpha 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.

3.9.3 Average UA Relative Amplitude Compared to Baseline

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

$$UA_{ratio\,increase_i} = mean(UA_{ratio\,training_i}) - mean(UA_{ratio\,baseline_i}) \tag{4}$$

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

3.9.4 Percentage of Time Above the Threshold

For every session, we considered the percentage of time during which the upper-Alpha ratio was above the threshold, where the threshold was the median value of the upper-Alpha 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.

3.9.5 Neurofeedback Transfer

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

Performance during the Transfer session was measured using the same metrics described before: the upper-Alpha relative amplitude compared to baseline, and the percentage of time the upper-Alpha ratio is above baseline.

4 Results

Here we present the impact of vividness of feedback in neurofeedback performance in terms of learning, presence, and how is upper-Alpha activity related to working memory.

4.1 Is Vividness of Feedback Affecting Neurofeedback Performance?

4.1.1 Vividness

By comparing the average upper-Alpha relative amplitude (L_1 index), we observe a tendency for participants in higher vividness groups to have a higher L_1 value (Fig. 8a).

Low vividness group has a median value below 0, meaning that participants did not manage to increase the upper-Alpha relative amplitude successfully. Medium vividness group and High vividness group have a positive median value, thus the partici-

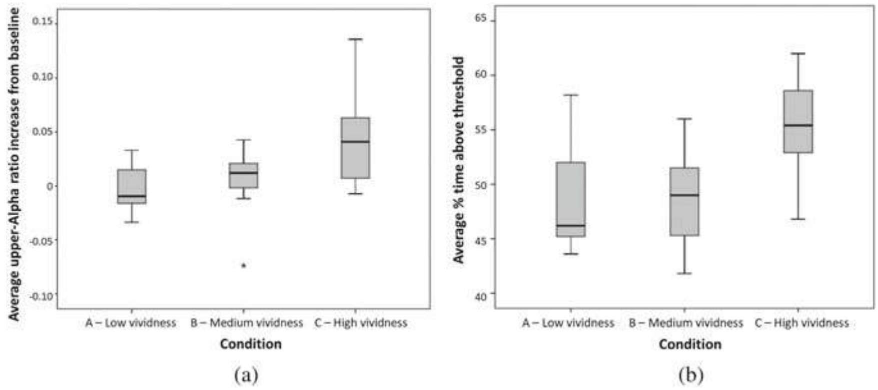


Fig. 8 Effect of vividness in neurofeedback performance. **a** Average upper-Alpha ratio increase from baseline, **b** average percentage of time above threshold

pants managed to modulate the upper-Alpha relative amplitude in the desired direction, with High vividness group performing better than Medium vividness group.

Participants in higher vividness groups tended to better modulate the upper-Alpha 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 upper-Alpha ratio increase between the different groups, $H(2) = 4.839$, $p > 0.05$, although close to significance ($p = 0.089$).

The same tendency is also observed in the percentage of time above the threshold (L_2 index) for participants in higher vividness (Fig. 8b).

During a training session, participants in the High vividness group managed to modulate their upper-Alpha relative amplitude above the threshold level for a longer time than participants in the Medium vividness group. And the same holds for the Medium vividness group compared to Low vividness.

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

Concerning the L_3 index–corresponding to the linear regression slope of the upper-Alpha ratio increase over the five training sessions–the median value is negative for all the groups, suggesting there was not an overall increase of upper-Alpha ratio across sessions. Finally, the Kruskal-Wallis test found no significant difference in the regression slope between groups, $H(2) = 2.494$, $p > 0.05$.

The L_4 index–corresponding to the linear regression slope of the percent time above threshold–has also a negative median value 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, $H(2) = 1.955$, $p > 0.05$.

4.1.2 Correlation Between Learning Indices

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 upper-Alpha relative amplitude and percentage of time above threshold respectively, within a session. While L_3 and L_4 measure the change of upper-Alpha amplitude and percentage of time across sessions. Hence, the results indicate a strong correlation between the two metrics of upper-Alpha, the relative amplitude and percentage of time above threshold. Similarly, an increase in the upper-Alpha amplitude across sessions corresponds to an increase in the 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 a session, achieving a higher increase of upper-Alpha ratio with respect to the baseline level, tended to show a lower increase of upper-Alpha ratio across sessions. Conversely, participants who showed low upper-Alpha ratio increase within a session attained a high upper-Alpha ratio increase across sessions.

4.2 Neurofeedback Learning over Time

Concerning neurofeedback training performance, we analyzed the upper-Alpha relative amplitude increase during the transfer session from the baseline level and the percentage of time above the threshold.

4.2.1 Upper-Alpha Relative Amplitude Increase from Baseline

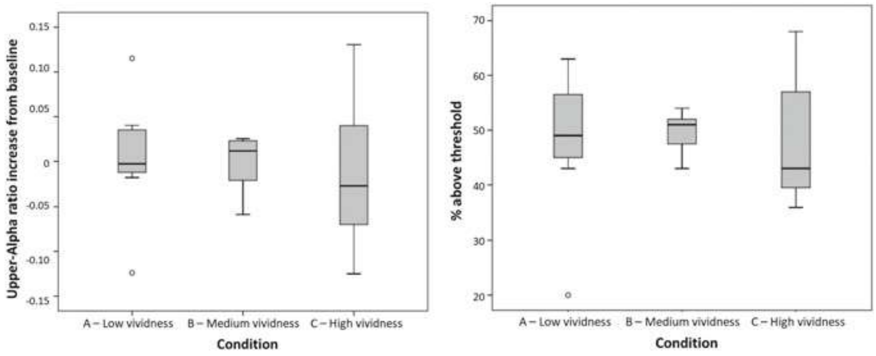
Figure 9a depicts the increase in the upper-Alpha relative amplitude during the transfer session per group. Only for Medium vividness, the median value is above 0, suggesting that during the transfer session participants successfully modulated the upper-Alpha ratio above the baseline level.

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

4.2.2 Percentage of Time Above Upper-Alpha Threshold

The graphical depiction (box plot) of the percentage of time the upper-Alpha ratio was above the baseline level during the transfer session per group, can be found in Fig. 9b.

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



(a) (b)

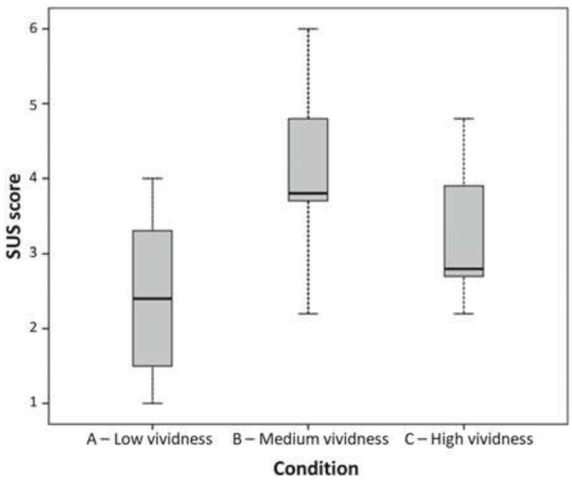
Fig. 9 Neurofeedback learning over time. **a** Upper-Alpha ratio increase from baseline during the neurofeedback Transfer session. **b** percentage of time above threshold during the neurofeedback transfer session

4.3 How Does Vividness Affect Presence?

In terms of presence, participants in the Low vividness group reported the lowest SUS score, but no statistically significant difference was found between groups, $H(2) = 4.954$, $p > 0.05$ (Fig. 10).

Regarding motivation, from current results, we can identify an increasing trend, with participants in higher vividness groups reporting to feel more motivated during neurofeedback training (Fig. 11a). However, a Kruskal-Wallis test showed no statistically significant difference between groups, $H(2) = 3.680$, $p > 0.05$.

Fig. 10 SUS questionnaire score between all three conditions



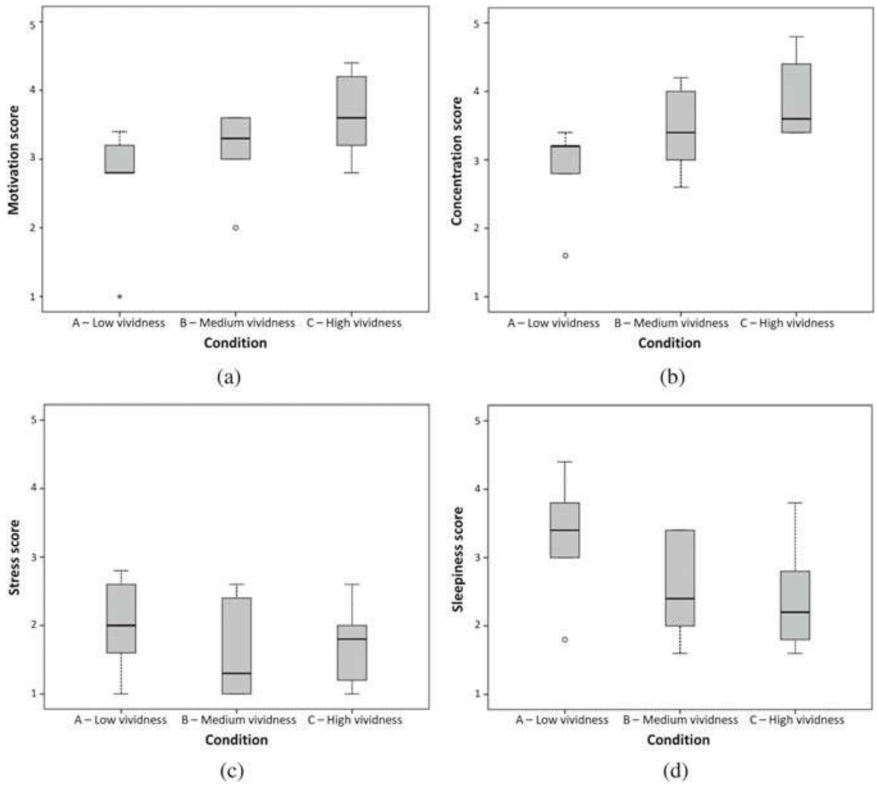


Fig. 11 Post-session survey results. **a** Motivation score per condition, **b** concentration score, **c** stress score, **d** sleepiness score

In terms of concentration, as for motivation, participants in higher vividness groups tended to feel more concentrated during neurofeedback training with respect to lower vividness groups (Fig. 11b), although a Kruskal-Wallis test showed no statistically significant difference between groups, $H(2) = 5.637$, $p > 0.05$.

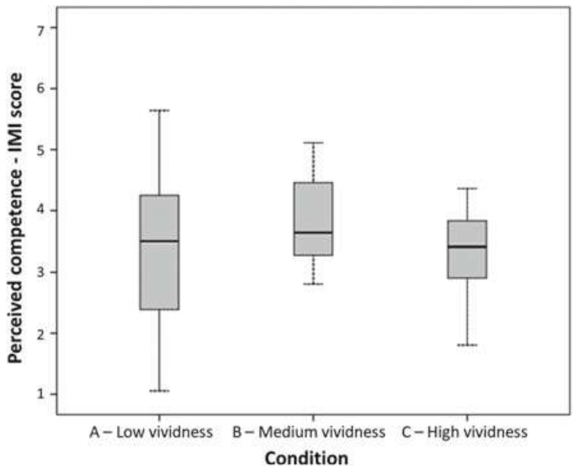
Moreover, participants reported low-stress scores with no statistically significant difference found between groups, $H(2) = 1.085$, $p > 0.05$ (Fig. 11c).

For sleepiness, we observe a pattern between the groups with participants in Low vividness group feeling drowsier during neurofeedback training sessions (Fig. 11d). A Kruskal-Wallis test showed no statistically significant difference between groups, $H(2) = 2.397$, $p > 0.05$.

In terms of perceived competence, there is no major differences between groups, with no statistically significant difference between groups, $H(2) = 0.831$, $p > 0.05$ (Fig. 12).

Finally, in terms of perceived workload, as reported through NASA-TLX, we observe a tendency to increase in higher vividness compared to lower vividness

Fig. 12 IMI Perceived competence score

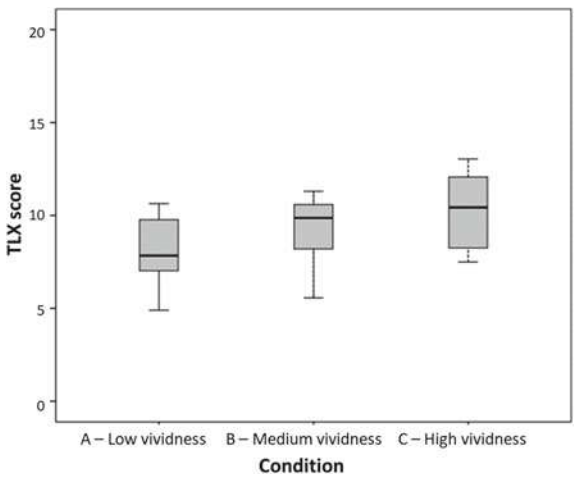


groups (Fig. 13). However, no statistically significant difference was found in the TLX score between groups, $H(2) = 2.753$, $p > 0.05$.

4.4 Is Upper-Alpha Activity Related to Working Memory?

The Digit Span test results showed that 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). Moreover, a Spearman's rank correlation between forward and backward Digit Span increase and the indices of neurofeedback learning showed no

Fig. 13 TLX perceived workload score



statistically significant relationship (Fig. 14). In the N-back test, results showed that 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 (Fig. 15).

A Spearman's rank correlation between the indices of neurofeedback learning and the performance measures in the N-back test revealed a statistically significant relationship between the Distractor accuracy increase in the 3-back test and the neurofeedback learning index L_3 ($r = 0.641$, $p < 0.01$) and the neurofeedback learning index L_4 ($r = 0.639$, $p < 0.01$) as well (Table 3). This relationship suggests that, when the upper-Alpha ratio or the percentage of time above threshold increased across the neurofeedback sessions, it corresponded to an enhancement in the Distractor accuracy in the 3-back test.

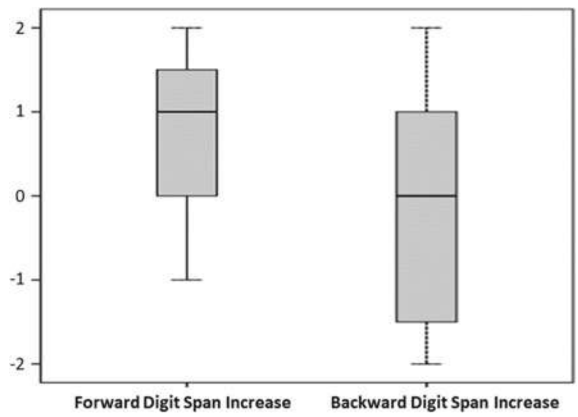


Fig. 14 Differences in Digit Span increase from pre to post-test

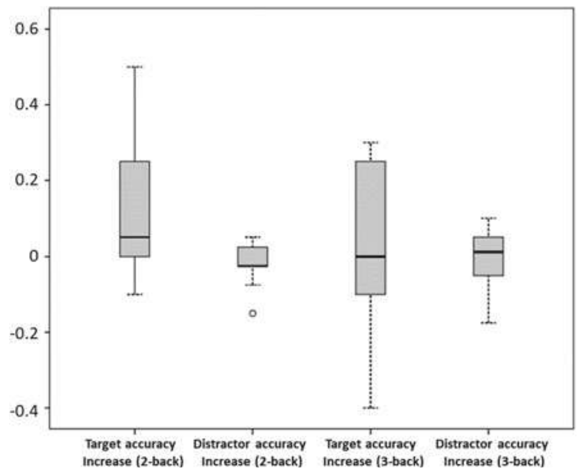


Fig. 15 2-back and 3-back test results

Table 3 Relationship between the indices of neurofeedback and performance measured

	Target accuracy increase (2-back)	Distractor accuracy increase (2-back)	Target accuracy increase (3-back)	Distractor accuracy increase (3-back)
L3—Slope UA ratio increase from baseline	−0.315	0.244	0.054	0.641*
L4—Slope % time above threshold	−0.374	0.111	0.027	0.639*

*Significant correlations ($p < 0.05$)

5 Discussion

In the present study, we manipulated the vividness level of the virtual environment used for providing neurofeedback training. Our aim was to assess the effect of different levels of vividness on presence levels during neurofeedback training as well as on the neurofeedback training outcome. Furthermore, we evaluated the effect of upper-Alpha neurofeedback training on working memory performance.

Concerning the effect of vividness on neurofeedback training performance, the literature suggests that the immersive properties of virtual environments facilitate neurofeedback learning. We hypothesized that a more vivid (thus more immersive) virtual environment would imply better neurofeedback performance. In terms of neurofeedback performance, we measured it using two metrics: the increase of upper-Alpha 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 current results, it emerges that participants in higher vividness groups tended to perform better within a neurofeedback session, in terms of both performance metrics than participants in lower vividness groups. Specifically, participants in the High vividness group attained better neurofeedback performance within a training session compared to participants in Medium vividness; the same holds, in turn, for participants in the Medium vividness group, who showed an improved neurofeedback training performance compared to the Low vividness group. 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), it is consistent with a positive effect of vividness on neurofeedback training, in accordance with our hypothesis. It is important to notice that participants in Low vividness group failed to attain control on their upper-Alpha relative amplitude, showing no increase in the upper-Alpha ratio with respect to the baseline level. We would have expected that every group was able to modulate the upper-Alpha 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 upper-Alpha self-regulation skill—it was monotonous and boring—thus not engaging compared to the higher vividness virtual environments. This could have made participants tired and reduced their dedication to the neurofeedback task (Berger and Davelaar 2018). This seems to confirm the importance of vividness and the advantage of a highly vivid virtual environment.

While evidence of neurofeedback learning within a session was found, there was no improvement in neurofeedback performance across sessions (neither in upper-Alpha 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 neurofeedback session, each composed of two 5-min training blocks, for a total of 50 min of neurofeedback training. Studies in which significant upper-Alpha learning across sessions was found used a neurofeedback procedure with longer sessions (25 min) (Escolano et al. 2011; Zoefel et al. 2011) and/or with a higher number of sessions (about 10) (Hsueh et al. 2016; Kober et al. 2016), resulting in at least double the neurofeedback training time than in our experiment. Furthermore, it has been shown that significant neurofeedback changes across sessions are usually found when comparing between the first and the later sessions, with no significant differences identified for the intermediate sessions. This suggests that, in the early stage of the neurofeedback training, changes across sessions cannot be detected. This is in agreement with our study, where a relatively short (50 min overall) neurofeedback training might be the reason why an enhancement of neurofeedback performance across sessions was not found. Furthermore, we noticed a negative relationship between the performance measure within a session (ability to up-regulate the upper-Alpha relative amplitude in a session) and across sessions (ability to enhance the upper-Alpha relative amplitude across sessions). This indicates that participants who showed a low increase of upper-Alpha ratio within a session tended to enhance neurofeedback performance across sessions. Finally, yet importantly, no significant difference was found in neurofeedback performance during the neurofeedback Transfer session between groups. The transfer session aimed to assess how the ability acquired during the neurofeedback 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 neurofeedback transferability.

From self-reported measures, it is shown that the vividness of the virtual environment had no statistically significant effect on subjective presence response, as measured with the SUS questionnaire. Even though not significant, we could notice that presence levels tended to be higher for Medium and High vividness groups compared to Low. This is in-line with findings in the literature (Slater et al. 1994; Usoh et al. 2000), and seems to confirm that subjective presence response increases with higher 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 complex 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., no texture vs. some textures) 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, there was a clear trend in both motivation and concentration tended to increase with greater levels of vividness. Participants in higher vividness groups reported to feel more motivated and focused on the task during neurofeedback training compared to participants in lower vividness groups. Furthermore, the results relative to the sleepiness variable showed that participants in the Low vividness group tended to feel drowsier during neurofeedback training compared to participants in Medium and High vividness group. The Low vivid training scenario made participants feel bored and lose interest in the neurofeedback training. As previously hypothesized, this could explain the fact that participants in group A did not achieve successful results in upper-Alpha 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 did not affect the sense of mastery in executing the neurofeedback task. Although non-significant, there was an increasing trend in workload results, showing that the perceived workload increased with a greater level of vividness.

Regarding working memory, there is evidence in the literature that upper-Alpha enhancement training has the effect of improving working memory performance (Escolano et al. 2011; Laufs et al. 2003; Zoefel et al. 2011). The hypothesis that an increase in upper-Alpha activity is correlated with increasing working memory 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 neurofeedback performance across neurofeedback training sessions. Specifically, it has been shown that an increase in the upper-Alpha ratio or in the percentage of time above threshold across neurofeedback sessions corresponded to an increase in the Distractor accuracy in the 3-back test.

6 Conclusions

The objectives of this study included first, the investigation of the effect of vividness in VR in terms of neurofeedback performance and subjective user experience, and second, the effect of upper-Alpha neurofeedback training on working memory performance.

From current results, we have been able to identify that vividness of feedback is affecting neurofeedback performance, showing in all performance metrics in the Medium and High vividness groups performed better than Low vividness. Moreover, highly vivid training scenarios had a positive effect on user's motivation, concentration, and reduced boredom. Nonetheless, we did not observe any learning effects across sessions for any of the groups. Finally, our results show that upper-Alpha neurofeedback training is an effective procedure to improve working memory performance, showing a positive correlation of upper-Alpha with working memory performance. This is also in-line with the findings of prior studies, indicating that also vivid VR feedback could possibly affect working memory training outcome.

7 Limitations and Future Work

Current limitations include the relatively small sample size per experimental group and also the short period of training time (5 sessions) per participant. There is evidence in the literature that a longer neurofeedback practice may be 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 neurofeedback transferability and on the improvement of neurofeedback performance across sessions. For future study, a prolonged neurofeedback training is necessary including a follow-up assessment for detecting the long-term effects. Importantly, future research should consider investigating further immersive factors for effects on neurofeedback performance and subjective response measures. Besides vividness, other variables such as extensiveness, proprioceptive matching, and inclusiveness could be examined, holding the potential for significant effects on neurofeedback outcomes. Finally, as wearable sensors become more ubiquitous in human-computer interaction, we aim to investigate further how we can gather unobtrusively ecologically-valid data in a CAVE-VR environment through the use of a wearable-EEG prototype in the shape of commercial smart glasses (Vourvopoulos et al. 2019a).

References

Aben B, Stapert S, Blokland A (2012) Front Psychol 3. ISSN1664-1078. https://doi.org/10.3389/fpsyg.2012.00301. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3425965/

Aeschbach D, Postolache TT, Sher L, Matthews JR, Jackson MA, Wehr TA (2001) Neuroscience 102(3), 493–502. ISSN 0306-4522. https://doi.org/10.1016/S0306-4522(00)00518-2. http://www.sciencedirect.com/science/article/pii/S0306452200005182

Baddeley A (1992) Science 255(5044), 556–559. ISSN 0036-8075. 1095–9203. https://doi.org/10.1126/science.1736359. http://science.sciencemag.org/content/255/5044/556

Bailenson JN, Yee N, Blascovich J, Beall AC, Lundblad N, Jin M (2008) J Learn Sci 17(1):102–141

Berger AM, Davelaar EJ (2018) Neuroscience 378:189–197. ISSN 0306–4522. https://doi.org/10.1016/j.neuroscience.2017.06.007. http://www.sciencedirect.com/science/article/pii/S0306452217304050

Berger H (1933) p 555. ISSN 1433-8491 (Electronic) 0940-1334 (Linking). https://doi.org/10.1007/bf01814320

Bucho T, Caetano G, Vourvopoulos A, Accoto F, Esteves I, Badia SBU, Rosa A CD, Figueiredo P (2019) Comparison of visual and auditory modalities for upper-alpha EEG-neurofeedback. Berlin, Germany. ISBN 978-1-5386-1311-5

Cho B-H, Ku J, Jang DP, Kim S, Lee YH, Kim IY, Lee JH, Kim SI (2002) Cyberpsychol Behav 5(2), 129–137. ISSN 1094-9313. https://doi.org/10.1089/109493102753770516. https://www.liebertpub.com/doi/abs/10.1089/109493102753770516

Choi HJ, Lee DY, Seo EH, Jo MK, Sohn BK, Choe YM, Byun MS, Kim JW, Kim SG, Yoon JC, Jhoo JH, Kim KW, Woo JI (2014) Psychiatry Invest 11(1), 39–43. ISSN 1738–3684. https://doi.org/10.4306/pi.2014.11.1.39. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3942550/

Dingler T, Agroudy PE, Le HV, Schmidt A, Niforatos E, Bexheti A, Langheinrich M (2016) IEEE MultiMedia 23(2):4–11 (2016). ISSN 1070–986X. https://doi.org/10.1109/MMUL.2016.31

Enriquez-Geppert S, Huster RJ, Herrmann CS (2013) Int J Psychophysiol 88(1), 1–16. ISSN 0167-8760. https://doi.org/10.1016/j.ijpsycho.2013.02.001. http://www.sciencedirect.com/science/article/pii/S0167876013000330

Enriquez-Geppert S, Huster RJ, Herrmann CS (2017) Front Hum Eurosci 11. ISSN 1662–5161. https://doi.org/10.3389/fnhum.2017.00051. https://www.frontiersin.org/articles/10.3389/fnhum.2017.00051/full

Escolano C, Aguilar M, Minguez J (2011) EEG-based upper alpha neurofeedback training improves working memory performance. In: 2011 annual international conference of the IEEE engineering in medicine and biology society, pp 2327–2330. https://doi.org/10.1109/IEMBS.2011.6090651

Freina L, Canessa A (2015) Immersive vs Desktop virtual reality in game based learning

Gonçalves A, Bermúdez S (2018) In: Proceedings of the ACM on Human-Computer Interaction, vol 2 (EICS):10:1–10:15. ISSN 2573-0142. https://doi.org/10.1145/3229092. http://doi.acm.org/10.1145/3229092

Gruzelier J, Egner T (2005) Child Adolesc Psychiatr Clin N Am 14(1):83–104

Gruzelier J, Inoue A, Smart R, Steed A, Steffert T (2010) Neurosci Lett 480(2), 112–116. ISSN 0304-3940. https://doi.org/10.1016/j.neulet.2010.06.019. http://www.sciencedirect.com/science/article/pii/S0304394010007524

Gruzelier JH (2014) Neurosci Biobehav Rev 44:124–141. ISSN 0149-7634. https://doi.org/10.1016/j.neubiorev.2013.09.015. http://www.sciencedirect.com/science/article/pii/S0149763413002248

Halder S, Varkuti B, Bogdan M, Kübler A, Rosenstiel W, Sitaram R, Birbaumer N (2013) Front Hum Neurosci 7. ISSN 1662-5161. https://doi.org/10.3389/fnhum.2013.00105. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3613602/

Hanslmayr S, Sauseng P, Doppelmayr M, Schabus M, Klimesch W (2005) Appl Psychophysiol Biofeedback 30(1), 1–10. ISSN 1573-3270. https://doi.org/10.1007/s10484-005-2169-8

Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index):results of empirical and theoretical research. In: Hancock PA, Meshkati N (eds) Advances in psychology, vol 52 of human mental workload. North-Holland, pp 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9. http://www.sciencedirect.com/science/article/pii/S0166411508623869

Heinrich H, Gevensleben H, Strehl U (2007) J Child Psychol Psychiatry 48(1), 3–16. ISSN 1469-7610. https://doi.org/10.1111/j.1469-7610.2006.01665.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-7610.2006.01665.x

Hsueh J-J, Chen T-S, Chen J-J, Shaw F-Z (2016) Hum Brain Mapp, 37 (7):2662–2675. ISSN 1097-0193. https://doi.org/10.1002/hbm.23201

Hubbard R, Sipolins A, Zhou L (2017) Enhancing learning through virtual reality and neurofeedback: a first step. In: Proceedings of the seventh international learning analytics & knowledge conference, LAK'17, New York, NY, USA. ACM, pp 398–403. ISBN 978-1-4503-4870-6. https://doi.org/10.1145/3027385.3027390. http://doi.acm.org/10.1145/3027385.3027390

Kleih SC, Kübler A (2013) Front Hum Neurosci 7. ISSN 1662-5161. https://doi.org/10.3389/fnhum.2013.00642. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3797970/

Kleih SC, Nijboer F, Halder S, Kübler A (2010) Clin Neurophysiol 121(7), 1023–1031. ISSN 1388-2457. https://doi.org/10.1016/j.clinph.2010.01.034. http://www.sciencedirect.com/science/article/pii/S1388245710000775

Kober SE, Witte M, Ninaus M, Neuper C, Wood G (2013) Front Hum Neurosci 7. ISSN 1662-5161. https://doi.org/10.3389/fnhum.2013.00695. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3798979/

Kober SE, Reichert JL, Schweiger D, Neuper C, Wood G (2016) Effects of a 3d virtual reality neurofeedback scenario on user experience and performance in stroke patients. In: Bottino R,

Jeuring J, Veltkamp RC (eds) Games and learning alliance, lecture notes in computer science. Springer International Publishing, pp 83–94. ISBN 978-3-319-50182-6
Kober SE, Reichert JL, Schweiger D, Neuper C, Wood G (2017). Does feedback design matter? a neurofeedback study comparing immersive virtual reality and traditional training screens in elderly. Int J Serious Game 4(3). https://doi.org/10.17083/ijsg.v4i3.167
Laufs H, Krakow K, Sterzer P, Eger E, Beyerle A, Salek-Haddadi A, Kleinschmidt A (2003) Proc Natl Acad Sci USA 100(19), 11053–11058. ISSN 0027-8424. https://doi.org/10.1073/pnas.1831638100. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC196925/
Lofthouse N, Arnold LE, Hersch S, Hurt E, DeBeus R (2012) J Atten Disord 16(5), 351–372. ISSN 1087-0547. https://doi.org/10.1177/1087054711427530. https://doi.org/10.1177/1087054711427530
Ma L, Chang L, Chen X, Zhou R (2017) PLOS ONE 12(3):e0175047. ISSN 1932-6203. https://doi.org/10.1371/journal.pone.0175047. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0175047
Marchiori E, Niforatos E, Preto L (2017) Measuring the media effects of a tourism-related virtual reality experience using biophysical data. Information and Communication Technologies in Tourism 2017. Springer, Berlin, pp 203–215
Marchiori E, Niforatos E, Preto L (2018) Inf Technol Tour 18(1–4):133–155
Marzbani H, Marateb HR, Mansourian M (2016) Basic Clin Neurosci 7(2):143–158. ISSN 2008-126X. https://doi.org/10.15412/J.BCN.03070208. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4892319/
Morrison AB, Chein JM (2011) Psychon Bull Rev 18(1), 46–60. ISSN 1531-5320. https://doi.org/10.3758/s13423-010-0034-0. https://doi.org/10.3758/s13423-010-0034-0
Nan W, Rodrigues JP, Ma J, Qu X, Wan F, Mak P-I, Mak PU, Vai MI, Rosa A (2012) Int J Psychophysiol 86(1), 83–87. ISSN 0167-8760. https://doi.org/10.1016/j.ijpsycho.2012.07.182. http://www.sciencedirect.com/science/article/pii/S0167876012005478
Nan W, Wan F, Vai MI, Da Rosa AC (2015) Front Hum Neurosci 9. ISSN 1662-5161. https://doi.org/10.3389/fnhum.2015.00677. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4685657/
Niforatos E, Palma A, Gluszny R, Vourvopoulos A, Liarokapis F (2020) Would you do it?: enacting moral dilemmas in virtual reality for understanding ethical decision-making. In: Proceedings of the 2020 CHI conference on human factors in computing systems, pp 1–12
Peper E, Shaffer F (2010) Biofeedback 38(4):142–147
Prime DJ, Jolicoeur P (2009) J Cogn Neurosci 22(11):2437–2446. ISSN 0898-929X. https://doi.org/10.1162/jocn.2009.21337. https://www.mitpressjournals.org/doi/10.1162/jocn.2009.21337
Renard Y, Lotte F, Gibert G, Congedo M, Maby E, Delannoy V, Bertrand O, LÃlcuyer A (2010) Presence Teleoperators Virtual Environ 19(1):3553. ISSN 1054-7460. https://doi.org/10.1162/pres.19.1.35
Robertson GG, Card SK, Mackinlay JD (1993) Computer 26(2):81. ISSN 0018-9162. https://doi.org/10.1109/2.192002
Ros T, Baars BJ, Lanius RA, Vuilleumier P (2014) Front Hum Eurosci 8. ISSN 1662-5161. https://doi.org/10.3389/fnhum.2014.01008. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4270171/
Sherman WR, Craig AB (2002) Understanding virtual reality: interface, application, and design. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. ISBN 978-1-55860-353-0
Slater M, Wilbur S (1997) Presence Teleoperators Virt Environ 6(6), 603–616. ISSN 1054-7460. https://doi.org/10.1162/pres.1997.6.6.603. https://doi.org/10.1162/pres.1997.6.6.603
Slater M, Usoh M, Steed A (1994) Presence Teleoperators Virt Environ 3(2), 130–144. ISSN 1054-7460. https://doi.org/10.1162/pres.1994.3.2.130. https://doi.org/10.1162/pres.1994.3.2.130
Slater M, Khanna P, Mortensen J, Yu I (2009) IEEE Comput Graph Appl 29(3):0272–1716. ISSN 76-84
Subramaniam K, Vinogradov S (2013) Front Hum Neurosci 7. ISSN 1662-5161. https://doi.org/10.3389/fnhum.2013.00452. https://www.frontiersin.org/articles/10.3389/fnhum.2013.00452/full
Teplan M (2002) Meas Sci Rev 2(2):1–11

Toczek YG (2016) The influence of visual realism on the sense of presence in virtual environments. Thesis, 29 Sep 2016. https://research.tue.nl/en/studentTheses/the-influence-of-visual-realism-on-the-sense-of-presence-in-virtu

Usoh M, Catena E, Arman S, Slater M (2000) Presence Teleoperators Virt Environ 9(5), 497–503. ISSN 1054-7460. https://doi.org/10.1162/105474600566989. https://www.mitpressjournals.org/doi/10.1162/105474600566989

Vourvopoulos A, Bermúdez i Badia S (2016) J NeuroEng Rehabil 13(1):69. ISSN 1743-0003. https://doi.org/10.1186/s12984-016-0173-2

Vourvopoulos A, Niforatos E, Hlinka M, Škola F, Liarokapis F (2017) Investigating the effect of user profile during training for BCI-based games. In: 2017 9th international conference on virtual worlds and games for serious applications (VS-Games), pp 117–124. https://doi.org/10.1109/VS-GAMES.2017.8056579

Vourvopoulos A, Niforatos E, Giannakos M (2019a) Eeglass: an eeg-eyeware prototype for ubiquitous brain-computer interaction. In: Adjunct proceedings of the 2019 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2019 ACM international symposium on wearable computers, pp 647–652

Vourvopoulos AT, Jorge C, Abreu R, Figueiredo P, Fernandes J-C, Bermúdez i Badia S (2019b) Front Hum Neurosci 13:244

Wan F, Nan W, Vai MI, Rosa A (2014) Front Hum Neurosci 8. ISSN 1662-5161. https://doi.org/10.3389/fnhum.2014.00500. https://www.frontiersin.org/articles/10.3389/fnhum.2014.00500/full

Wang N, Doube W (2011) How real is really? a perceptually motivated system for quantifying visual realism in digital images. In: 2011 international conference on multimedia and signal processing, vol 2, pp 141–149. https://doi.org/10.1109/CMSP.2011.172

Witte M, Kober SE, Ninaus M, Neuper C, Wood G (2013) Front Hum Neurosci 7. ISSN 1662-5161. https://doi.org/10.3389/fnhum.2013.00478. https://www.frontiersin.org/articles/10.3389/fnhum.2013.00478/full

Yan N,Wang J, Liu M, Zong L, Jiao Y, Yue J, Lv Y, Yang Q, Lan H, Liu Z (2008). ISSN 1609-0985. http://hub.hku.hk/handle/10722/175308

Zimmons P, Panter A (2003) The influence of rendering quality on presence and task performance in a virtual environment. In: IEEE virtual reality, 2003. Proceedings, pp 293–294. https://doi.org/10.1109/VR.2003.1191170.

Zoefel B, Huster RJ, Herrmann CS (2011) NeuroImage 54(2):1427–1431. ISSN 1053-8119. https://doi.org/10.1016/j.neuroimage.2010.08.078. http://www.sciencedirect.com/science/article/pii/S105381191001181X

Floriana Accoto holds an M.Sc. in Computer Engineering from the University of Salento, Lecce, Italy. Her master thesis was focused in assessing the effect of VR elements in Upper Alpha neurofeedback.

Athanasios Vourvopoulos is a Postdoctoral Researcher at the Institute for Systems and Robotics (ISR) of Instituto Superior Técnico, Lisbon, Portugal. He holds a Ph.D. in Computer Engineering and his research is focused on multimodal Human-Machine interaction using Brain-Computer Interfaces and Virtual Reality for neurorehabilitation.

Afonso Gonçalves is an aerospace engineer and informatics engineering Ph.D. candidate at the University of Madeira, in the field of Human-Machine Interaction. His research, at the NeuroRehabLab of Madeira Interactive Technologies Institute, is focused on the study of exergames to promote physical activity in the older age, and virtual reality.

Teresa Bucho completed her master studies in Biomedical Engineering in 2018 at Instituto Superior Técnico (Lisbon, Portugal). Soon after, she began her Ph.D. at the Netherlands Cancer Institute (Amsterdam, the Netherlands), focusing on the development and application of artificial intelligence in medical imaging.

Gina Caetano holds a Ph.D. from Aalto University, with a thesis on multisensory interaction and the human mirror neuron system. She is currently an invited professor and researcher at the University of Lisbon. Her research interests include the application of neuroimaging techniques in clinical and basic neuroscience.

Patrícia Figueiredo has a BSc in Physics Engineering from Instituto Superior Técnico, Universidade de Lisboa (IST), DPhil in Neuroimaging from the University of Oxford and Habilitation in Biomedical Engineering from IST. She is Assistant Professor at IST, and coordinator of the Biomedical Engineering Lab, Institute for Systems and Robotics Lisboa.

Lucio Tommaso de Paolis had a Degree in Electronic Engineering from the University of Pisa (Italy) and is the Director of the Augmented and Virtual Reality Laboratory (AVR Lab). His research interest concerns the applications of Virtual and Augmented Reality and the development of human-computer interfaces in medicine and education.

Sergi Bermudez I Badia is a Senior Researcher and member of the board of directors of the Madeira-ITI, Associate Professor at the University of Madeira, and President of the International Society for Virtual Rehabilitation. Graduated in telecommunications engineering from the UPC (Barcelona) and Ph.D. neuroinformatics from ETH (Zurich).