UNIVERSITY
OF TWENTE.

Enhancing Team Coordination through Interpersonal Neural Synchrony achieved with Neurofeedback Training

Anastasia Veliev

Master Thesis

Faculty of Electrical Engineering, Mathematics and Computer Science

Examination Committee:

Dr. M. Poel

Dr. J. van Erp

November 2024

Table of contents

Abstract

Recent studies have indicated that interpersonal neural synchrony (INS) in dyads may be increased through neurofeedback training (NFT), and that there might be a correlation between INS and interpersonal coordination. The current study aims to investigate these recent findings further by examining whether three sessions of INS-NFT can increase the INS and interpersonal coordination between two team members. Two research questions are investigated, namely: *RQ1. Is frontal theta and delta INS-CCorr higher during a third session of INS-NFT than during the first session?*, and *RQ2. Is team coordination higher after three sessions of theta and delta frequency INS-NFT?*. To explore these research questions, four teams of two participants have trained INS-NFT for three sessions of twenty-four minutes. Of these four teams, two teams received real feedback while two other teams received random feedback. Before and after INS-NFT, all teams completed a cooperative interpersonal coordination task in which the goal was to win as many points as possible as a team. Each team could win one point per trial if both team members pressed relatively simultaneously to each other after a stimulus was shown. If the team members did not press simultaneously enough the team would lose one point. The two research questions have been investigated through nine statistical analyses, which did not show a trend towards INS or interpersonal coordination being trainable through INS-NFT. While INS did not increase significantly in the experimental condition, differences between electrode couples suggest that some electrode couples are more trainable than others, with a significant increase of INS in F8xFz and F8xF4 in the delta frequency. Additionally, interpersonal coordination winning rate was significantly higher for the experimental condition than for the control condition. Based on these findings it is recommended for future research to investigate the effect of multi-session delta INS-NFT on INS for electrode couples F8xF4 and F8xFz, and to re-examine the effect of actual and control INS-NFT on interpersonal coordination.

1 Introduction

A team's ability to coordinate their actions effectively has been considered as one of the most important factors in achieving high team performance for decades [71]. Some have even previously described interpersonal coordination as the essence of teamwork [3] as it increases team effectivity [23], which in turn may create more knowledge, minimize errors, promote innovation, save lives, enhance productivity, increase job satisfaction and increase the chance of success [98]. A team's *coordination* ability is commonly defined as "the ability to use strategies and behaviour patterns aimed at integrating and aligning the actions, knowledge, and objectives of interdependent team members, with a view to attaining common goals" [92]. This ability enables unique team members to act as a unified whole instead of as a group of individuals, which enables teams to achieve higher results in situations that require interpersonal coordination [92].

Lately, a number of studies have investigated the phenomenon of *Interpersonal Neural Synchronization* (INS), which can be defined as synchronous brain activity between two or more individuals. While INS has not yet been fully understood, it is known to occur mostly in social situations such as verbal and non-verbal communication [81], shared attention [60], touch by a loved one [66] and during cooperative tasks [112]. It is speculated to play an important role in relationships and mutual understanding, as correlations are demonstrated between INS and prosocial behaviour [31], empathy [48], group/partner affinity [24, 77], attentiveness to social cues [68] and positive affect [113]. INS is also found to be related to a number of phenomena that are important for teamwork, such as cooperation [112], feelings of cooperativeness [45], feelings of being "part of the same team" [5], egalitarian reward allocation [130] and team efficiency [7]. Additionally, it has been demonstrated that INS is related to higher performance of team members but not to higher performance in groups of individuals [90].

Interestingly, a significant increase in INS has also been found during or after tasks requiring cooperative interpersonal coordination through fNIRS [20] and EEG [119]. The purpose of INS during such tasks remains unclear, but it has been shown to be nearly absent when the same task is performed individually or competitively [20, 64]. Some authors speculate INS can be seen as a marker of shared intentionality (i.e., joint action with shared goal) [33]. However, a recent study by Wang et al. (2020) demonstrated that children with severe symptoms of autistic spectrum disorder have significantly lower INS in the right Superior Frontal Cortex (SFC) than dyads of children with less severe symptoms, even though both groups experienced shared intentionality [122]. These findings suggest that INS could be dependent on the dyad's capacity to use theory of mind or communicate non-verbally, as ASD is known to impair these abilities ([122, 87]). Additionally, the children with less severe symptoms had higher task scores in the cooperative coordination task, suggesting covariation between INS and cooperative coordination.

While these results suggest a possible correlation between INS and cooperative interpersonal coordination, this relationship has yet not been thoroughly investigated. One of the main goals for the current study is to investigate the relationship between INS and interpersonal coordination further, and to examine whether INS may have a causal effect on interpersonal coordination within a team of two individuals. The methodology that is used for this purpose is to increase INS in dyads through multiple sessions of Neurofeedback Training (NFT) and to compare the interpersonal coordination in a task before and after INS enhancement. It is only recently that studies have suggested that it is possible to increase INS through NFT in a single session [77] or through multiple sessions in pigeons [127], which makes this methodology a new and relatively unexplored field of research. The current study is the first study to date to increase INS in human dyads through multiple sessions of NFT. Therefore, the second goal for this study is to examine whether NFT is a reliable method for increasing INS in human dyads, and how much training is required to achieve higher INS.

As this study is performed for the purpose of exploration of a new field of research, the goals for this study are translated into research questions. The following research questions will be investigated (Further elaborated on in Section 4, Research Questions).

> RQ1. Is frontal theta and delta INS-CCorr higher during a third session of INS-NFT than during the first session?
>
> RQ2. Is team coordination higher after three sessions of theta and delta frequency INS-NFT?

This thesis is structured as follows. Section 2 gives an overview on all the background information that is necessary in order to understand and read further sections of the paper. This information is especially useful if the reader does not have a background in neuroscience. In section 3 the most relevant state-of-the art literature is presented, which creates a foundation for section 4, in which the two research questions will be further elaborated on. Section 5 informs about the methodology used in the experiment, and presents the reasoning behind the methodology. The next section, section6, presents the experimental results which are discussed in section 7. Finally, the conclusion will give a brief summary of the results and conclude the report in section 8.

2 Background

In this section, background information will be presented about the topics that are related to this study. First, interpersonal coordination will be discussed. What is interpersonal coordination, why is it needed, what influences it? These are questions that will be answered in the first subsection. The second subsection will contain information about the basics of neuroscience: the anatomy of the (social) brain, neuroimaging techniques and their applications and neural frequencies. In the third section an introduction will be given on the topic of Interpersonal Neural Synchrony or INS, which includes information about hyperscanning and similarity measures. Finally, the fourth section will discuss the basics of Neurofeedback Training and how it can be applied.

2.1 Interpersonal coordination

Social situations often require some form of coordination of actions, emotions, thoughts and physiological processes by different actors: be it for playing an instrument, during teamwork or even when walking the same street. This type of alignment is often referred to as "interpersonal coordination", which can be defined as "the spontaneous rhythmic and temporal coordination of actions, emotions, thoughts and physiological processes between two or more individuals" [70]. It has been argued that there are two main types of interpersonal coordination: mimicry and interpersonal synchrony [70]. *Mimicry* can be defined as the practicing of the same behaviour as another individual in a relatively short period of time after it occurred, and *interpersonal synchrony* refers to all other types of interpersonal coordination. These two types differ from each other by (i) timing of the coordination (which is critical to synchrony, and less so in mimicry) and (ii) in mimicry, behaviour is similar or even identical, whereas in interpersonal synchrony behaviour may even be complementary (such as in turn-taking during conversations).

In the current study I will refer to interpersonal coordination as "the ability to use strategies and behaviour patterns aimed at integrating and aligning the actions, knowledge, and objectives of interdependent actors, with a view to attaining common goals" [92]. By this definition, the interpersonal synchrony of action, knowledge and objectives that arises from interpersonal coordination is not accidental, but instead deliberate in order to attain common goals [70]. This type of coordination is essential in teamwork, as by definition teams work toward shared and valued goals [96].

A team's ability to coordinate is partly dependent on a number of factors like the team's cohesion and trust [85] and shared knowledge [2]. Shared knowledge is the degree of common or complementary knowledge among team members which ensures that team members can describe and explain knowledge to each other as well as predict each other [34]. Failure to establish a shared understanding of a situation can result in impaired teamwork and negative outcomes [97]. The cohesion between team members represents the degree of willingness to work for the group, perceptions of group attractiveness, and the integration of the group as a whole [85]. While team coordination seems to improve with better team cohesion and trust, the opposite also seems to be the case - the relationship is reciprocal [85].

According to various studies in team cognition, team coordination depends on explicit and implicit coordination between team members [92]. While explicit coordination consists of communicative feedback processes, implicit coordination takes place when one person anticipates the actions and needs of the others and changes their behaviour accordingly [92]. Both implicit and explicit coordination are pivotal drivers of team performance [97].

Importantly, there can exist different types of interpersonal coordination. When speaking about interpersonal coordination of joint actions (see subsection 2.1.1. Related concepts) one can speak of multiple individuals performing the same action, or individuals performing complementary actions. The latter is called *complementary interpersonal coordination* [97].

2.1.1 Related concepts

Joint action is commonly defined as "any form of social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment" [100]. Within business environments teams commonly participate in joint actions, as it is common to share a particular goal (e.g. finishing a report) and coordinate the work required. It has been suggested that successful joint action mainly depends on three abilities: the ability to share representations, the ability to predict actions of the other collaborator and to integrate predicted effects of one's own and others' actions [100, 118]. As observing and responding could be not quick enough for some situations, predicting the actions of others and the corresponding consequences can be necessary, which is followed by adaptation of their own actions in accordance. This process facilitates fast and accurate interpersonal coordination

There are multiple theories for how actions of others may be predicted by the human brain. It is generally assumed that the mirror neuron system (MNS) is important for action prediction. Not only does it increase activity when observing an action, it has also been demonstrated to activate before a predictable action [69]. Action prediction is also often considered to work according to the predictive processing framework; that is, through a generative model that constantly updates its expectations about upcoming actions according to previous experiences. In line with this theory, it has been previously found that the prediction of an action affects the way it is being perceived [117].

Joint attention is when people experience the same thing at the same time, knowing that they are experiencing it together [111]. Importantly, joint action requires joint attention, but not vice versa [109].

The term *theory of mind* (ToM) describes both the ability to understand and predict the behaviour of other people by making inferences about their mental states, their intentions, feelings, expectations, beliefs or knowledge, and to cognitively represent one's own mental states [62]. Another term for ToM is mentalizing, and behaviourally it is commonly associated with empathy [12]. Healthy individuals are able to read the affective states of others from their body movements, body posture, gestures, facial expressions, and action performance [117]. It has been demonstrated before that the perceived affective states of others may have an influence on how quickly participants respond during computer tasks in which the goal is to press as quickly as possible [43].

2.2 Anatomy of the human brain

The brain can be divided into three parts with roughly the same functionalities: the cerebrum, cerebellum and brainstem (see figure 2.2.1.) [103]. The *cerebrum* is the largest part of the human brain and is generally associated with higher order brain functions such as generation of thoughts, movements, emotions and motor functions. The outermost layer of the cerebrum is made up of neural tissues known as the cerebral cortex. The cerebrum can be divided into two hemispheres and four lobes, which are grouped by functionality: the frontal lobe, parietal lobe, occipital lobe and temporal lobe. The frontal lobe is involved in inhibition, reasoning, personality, emotions, problem solving, motor development, planning, (parts of) speech and movement [103]. The parietal lobe is involved in the feeling of sensations (e.g., pain, touch), sensory comprehension, recognition and perception of stimuli, orientation and movement. The occipital lobe is crucial in visual processing, and finally the temporal lobe is involved in recognition of auditory stimuli, speech, perception and memory [103].

The *cerebellum* is mostly associated with motor control, sensory perception and motor coordination, and is crucial for voluntary muscle movements, fine motor skills, posture and balance regulation. Finally, the *brainstem* is located at the bottom of the brain and connects the cerebrum to the spinal cord. The brainstem is associated with basic and vital functions such as breathing, consciousness, movements of the eyes and mouth, the relaying of sensory messages, heartbeat, blood pressure and hunger [103].

The regions that are of particular interest for the current study are the mentalizing system or the theory-of-mind network and the mirror neuron system (see figure 2.2.2.). The Mirror Neuron System (MNS) plays an important role in tasks involving movements, such as imitation and coordination/joint tasks [121]. The ToM network is mainly important in tasks pertaining to the inferences of oneself or others' intentions or thoughts [121].

2.2.1 Theory-of-mind network (ToM)

The ToM network is a name for a selection of brain regions that are supposed to be related to the ability of ToM. It has also been referred to as the mentalizing network/system. The brain regions that are commonly referred to as the ToM network are: (i) the medial prefrontal cortex, the temporoparietal junction (TPJ) and (ii) the anterior cingulate (ACC), anterior insula (AI),

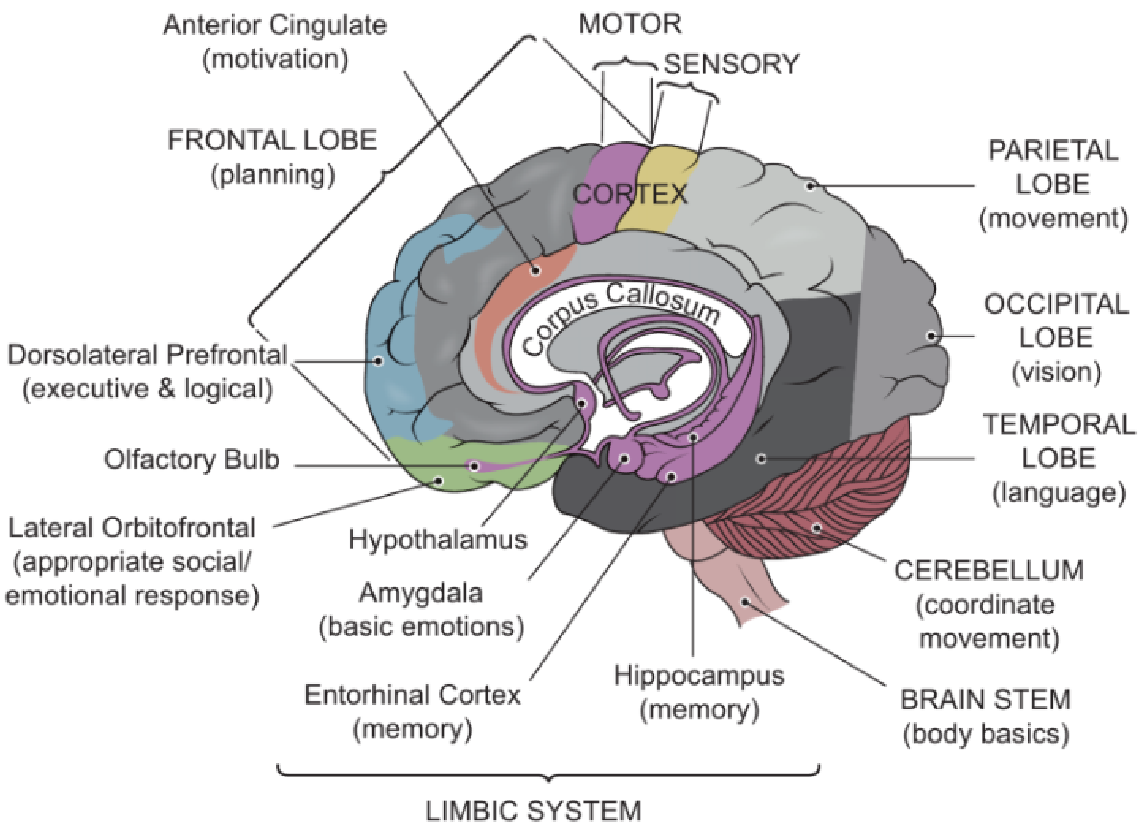


Figure 2.2.1: Anatomy of the brian [1].

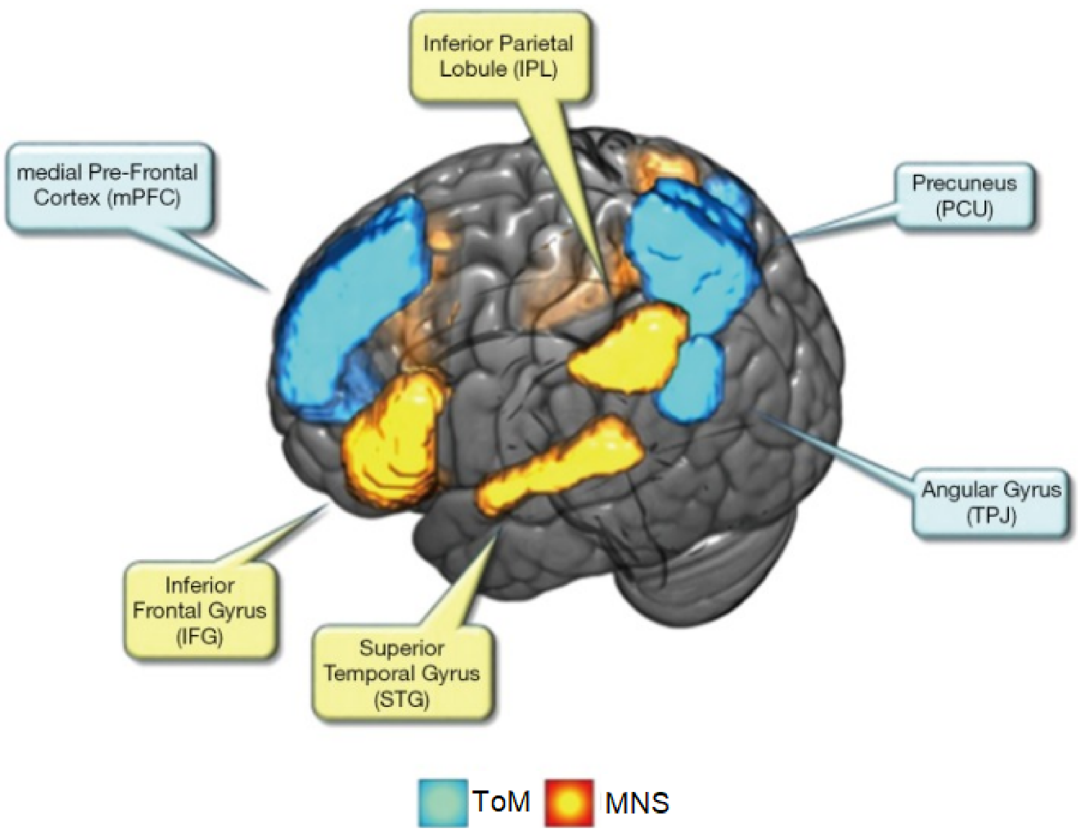


Figure 2.2.2: The ToM Network (ToM) and the mirror neuron system (MNS), both involved in social interactions [121].

precuneus, the posterior cingulate cortex, lateral areas of the middle temporal lobes, the superior temporal sulcus (STS) and the temporal poles [99]. The latter (ii) will be discussed in this subsection, while the former (i) will be explained through additional subsections.

The AI and the ACC are associated with emotional awareness, or the ability to perceive feelings of oneself [42]. The precuneus seems to be important for the experience of agency and the processing of self-relevant material and the representation of others' thoughts [99]. The temporoparietal regions are involved in the representation of the thoughts of others and are active during verbal and nonverbal ToM tasks [99]. The temporal regions around the STS contain mirror neurons that are activated during behavioural imitation and learning as well as in recognition of intentional movements.

2.2.2 Mirror Neuron System (MNS)

The Mirror Neuron System (MNS) is a frontoparietal sensorimotor network recruited during the observation (and mostly also execution) of actions, and is considered to be essential for understanding others' overt actions from low-level behavioural inputs. Factors that may influence activity in the MNS can be observed in figure 2.2.2.1. [52]. Like with the ToM, different studies define the MNS differently in terms of brain regions. The core regions that are often included in the definition are the premotor cortex, inferior frontal gyrus, and inferior parietal lobule (see figure 2.2.2.2) [52]. Other regions that are also commonly studied are the primary motor cortex, the supplementary motor area and the superior temporal gyrus. Some parts of the MNS have been shown to contain "mirror neurons", which discharge when an individual performs a given motor act and when the similar actions of another person are observed.

Both the Inferior frontal gyrus and the inferior parietal lobule are related to language, motor and sensory detection [121]. The Superior Temporal Gyrus can provide additional visual information to the inferior parietal lobule [46] which allows movements to be executed together with the inferior frontal gyrus. This activation in the latter two regions allows for the manipulation of potential actions through information such as the goal of the action [121]).

2.2.3 Prefrontal Cortex (PFC)

The Prefrontal Cortex (PFC) consists of many subregions that are associated with ToM and decision making [16]. It is also related to the perceivement of similarities between oneself and another [74]. One area that is particularly known for making this distinction is the medial PFC, which is involved in distinguishing between the self and others, and also between reality and deception. It is particularly important in the attribution of saliency and error monitoring [51]. The dorsal part of the medial PFC is especially involved in mentalizing tasks, and tends to be more activated when individuals reflect on others as compared to themselves. One interesting feature is that its activation increases when appraising the intention of others with respect to the implications for one's own well being [25]. It has been found in a study by Wittmann et al. that the confusion between one's own performance with that of others is related to activity in Brodmann area 9, which is a part of the dorsolateral and medial PFC [125]. The dorsolateral PFC is also associated with social cognition, and in particular with theory of mind [51]. The ventromedial PFC is thought to be related to the recognition of facial expressions [126]. The Left PFC is found to play important roles in cognitive control [67]. Finally, the orbitofrontal cortex (OFC) is thought to be one of the most important brain areas for emotional assessment and emotional regulation. It is known to be related to the adjustment of emotional expression through evaluation of emotional salient stimuli and regulation of the subjective emotional experience [91]. The OFC is also associated with inhibitory control. Often, the OFC is considered to be part of the ventromedial PFC, but this seems to vary between studies. For a more thorough literature review, see [25].

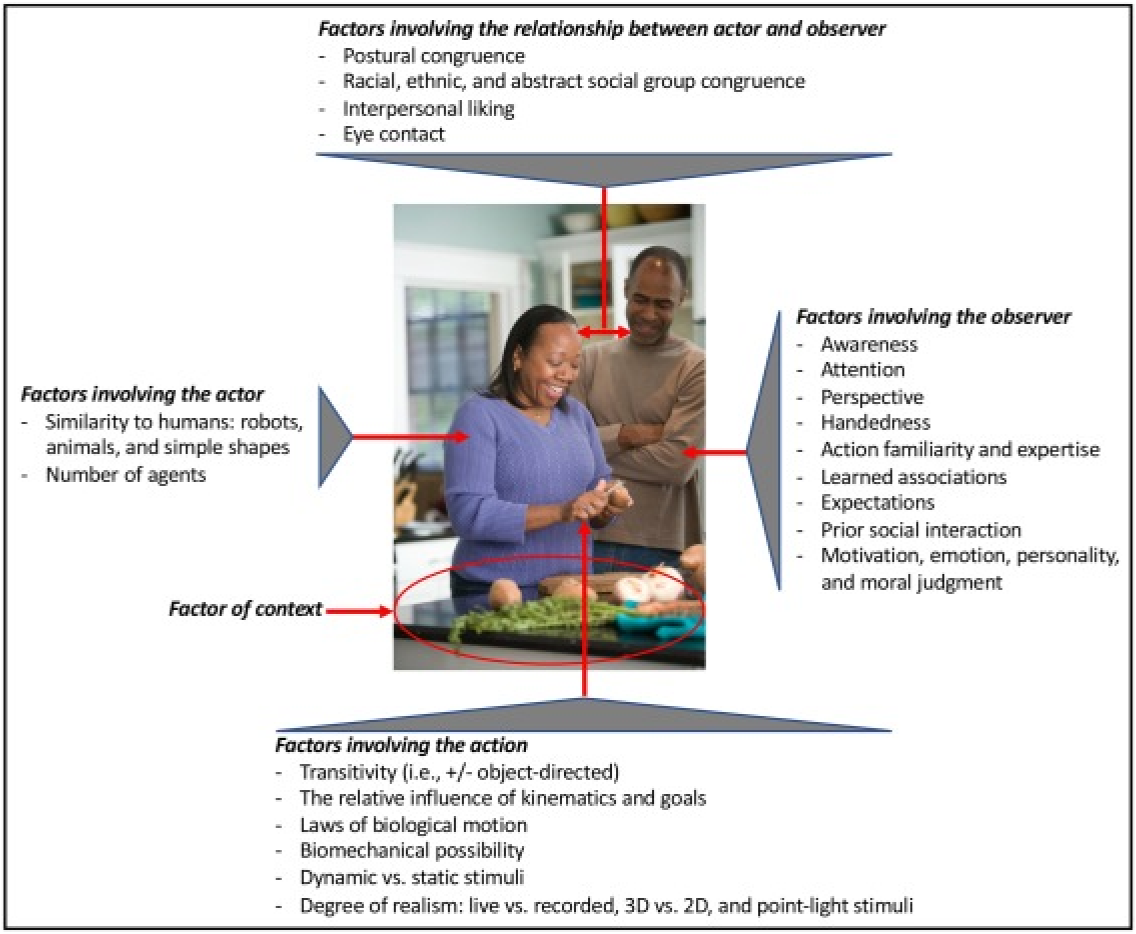


Figure 2.2.2.1: Factors that may influence activity in the MNS, grouped into five categories: factors involving the relationship between actor and observer, factors involving the actor, factors involving the observer, factors of the context and factors involving the action [52].

Figure 2.2.2.2: Core regions of the MNS: the inferior parietal lobule, especially the supramarginal gyrus and adjacent intraparietal sulcus (yellow); the ventral premotor cortex (green); the posterior inferior frontal gyrus (red). Other regions that are thought to be involved in the MNS are the primary motor cortex (blue) and supplementary motor area (magenta) [52].

2.2.4 Temporoparietal Junction (TPJ)

The Temporoparietal Junction is a commonly studied region in ToM research as it is related to mentalizing processes and the understanding of others' intentions or emotions by their gestures, behaviours and facial expressions [16]. It is commonly theorized that the TPJ allows attention to shift from the self to focus on the perspective of others (Steinbeis 2016). The TPJ also seems to be involved in overcoming self-centeredness, rendering behaviour prosocial [106]. There are some results that suggest hemispheric specializations for the TPJ, but the results are somewhat contradictory.

2.3 Neuroimaging techniques

There are various neuroimaging techniques that can be used for Neurofeedback Training (See subsection 2.6) and Hyperscanning (See subsection 2.5). The following subsections will introduce three of the methods: Functional Magnetic Resonance Imaging (fMRI), Functional Near-Infrared Spectroscopy (fNIRS) and Electroencephalography (EEG).

2.3.1 Functional Magnetic Resonance Imaging (fMRI)

Functional Magnetic Resonance Imaging (fMRI) is a hemodynamic measurement technique which measures the changes in the contrast between oxygenated (O2Hb) and deoxygenated (HHb) haemoglobin concentrations in the brain which is called the BOLD (Blood-oxygen-level-dependent) response [21]. Haemoglobin is a molecule in the blood that carries oxygen from the lungs to all other parts of the body. It also has different magnetic properties depending on the amount of oxygen that it carries, which can be measured through fMRI [4]. This is an indirect measure of neuronal activity and increases and decreases in the BOLD response are often assumed to reflect increases and decreases (deactivation) in neural activity [76]. Following an impulse of activity, the BOLD signal takes several seconds to rise and 20 to 25 seconds to fall, leading to a low temporal resolution for fMRI (see figure 2.3.1.1.) [4]. An example of the BOLD response in a fMRI scan can be viewed in figure 2.3.1.2.

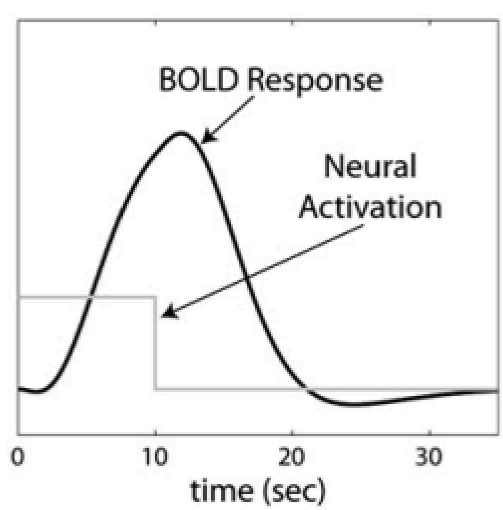


Figure 2.3.1.1: Temporal comparison of the BOLD response and neural activation [4].

The MR scanner uses superconducting electromagnets to produce a static, uniform magnetic field of high strength (approximately 3 Tesla is commonly used). The static field does not produce an MR signal by itself: it requires radiofrequency coils that generate magnetic pulses [4]. By turning on such magnetic pulses changes the magnetization alignment of protons within the magnetic field.

When the pulse is turned off, the protons return to their original equilibrium alignment, which releases energy detected by the coils as the raw MR signal.

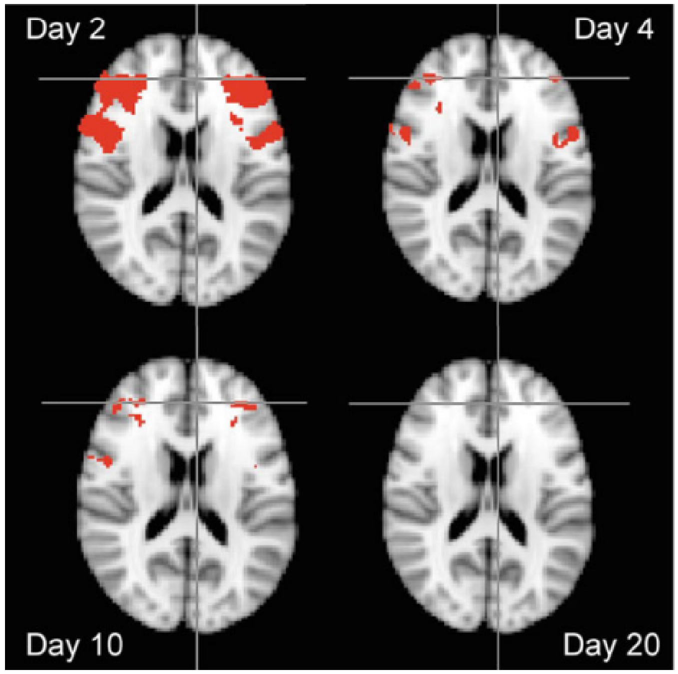


Figure 2.3.1.2: Measures of the BOLD response on four different days of conducting the same task. The BOLD response can be seen in red [4].

One of the most important advantages of fMRI is its spatial resolution, which is usually around 3mm. The disadvantages of fMRI include the limited mobility of the participants (as it requires laying still in a machine, see figure 2.3.1.3), the relatively low temporal resolution and the costs and availability of the machines.

2.3.2 Functional Near-Infrared Spectroscopy (fNIRS)

Functional Near-Infrared Spectroscopy (fNIRS) is an optical, noninvasive neuroimaging technique that allows the measurement of brain tissue concentration changes of oxygenated (O2Hb) and deoxygenated (HHb) haemoglobin following neuronal activation. This is achieved by shining NIR light (650–950 nm) onto the scalp with a light-emitting diode (LED). Due to the relative transparency of the biological tissue within this NIR optical window, light will reach the brain tissue [86]. This light will be reabsorbed by a photo-detector, and the changed characteristics of the light can inform about the BOLD response. This technique is an effective but 'indirect' imaging technique as it assumes that neural activation and vascular response are tightly coupled. Indeed, there are various studies that show a linear relationship between neural activity and hemodynamic responses. Therefore, an increase in cerebral blood flow is generally assumed to be associated with cerebral activity, making the temporal and spatial correlation between cerebral blood flow and task a determinant of cerebral function [86].

fNIRS can be utilised in two ways. The most common and commercially available technique uses fNIRS source and detector optical fibers (or optodes) which are distributed uniquely at various locations on the head and at fixed source–detector separations. Each source–detector separation represents a measuring channel providing topographical representation of the distribution of the changes in concentration of O2Hb and HHB over the cortical surface. Since recently similar fibreless fNIRS can also be utilised, which are more lightweight and robust to movement. The second fNIRS

Figure 2.3.1.3: An fMRI Scanner (Orellana, 2015).

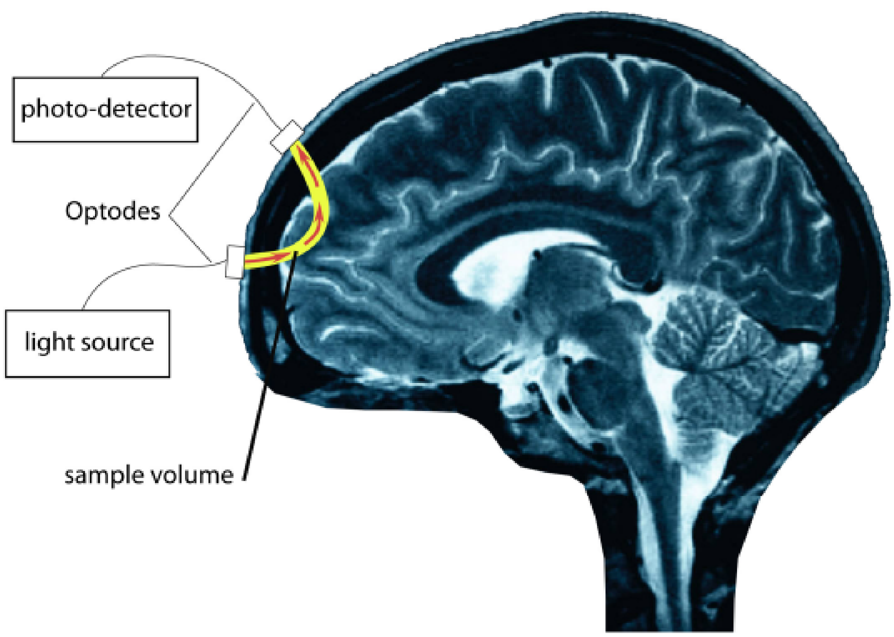


Figure 2.3.2.1: Two optodes in an fNIRS system. The NIR light is guided to the head by an optode. A photo-detector, placed 2 to 7 cm from the head, collects the light when it leaves the head [86].

technique is called diffuse optical tomography (DOT) and utilises denser arrays and overlap of channels.

fNIRS is best suited for measuring superficial brain areas with a low spatial resolution (1 cm) with a temporal resolution lower than that of EEG. The fNIRS systems that are being used today are easy in application and are characterised by the small size, portability, and reliability.

2.3.3 Electroencephalography (EEG)

Electroencephalography (EEG) is a commonly used neuroimaging technique that reads scalp electrical activity generated by brain structures. It is one of the oldest and most commonly used brain-imaging techniques at this moment [34]. It distinguishes itself from other measuring techniques by being non-invasive, relatively quick to use, having a relatively high temporal resolution (i.e. signal per time ratio) and being usable for most tasks in a laboratory environment. EEG is based on the measurement of electrophysiological signals that originate mostly from pyramidal neuron cells. Because the head is a conducting medium, the simultaneous activity of thousands of cells can be measured through EEG electrodes on the scalp [72]. Due to EEG having not a very high spatial resolution compared to other neuroimaging methods, it is hard to determine the exact location from where a certain signal originates. It is also not suited for the measurement of activity in deeper brain structures, and is only restricted to activity in the cerebral cortex. An EEG response which is aligned with a certain stimulus is called an Event-Related Potential (ERP). An illustration of how EEG signals are being measured can be viewed in figure 2.3.3.1.

2.3.4 magnetoencephalography (MEG)

Magnetoencephalography (MEG) measures the netto contributions of excitatory and inhibitory dendritic postsynaptic potentials through magnetic field changes [88]. One benefit of MEG over EEG is that EEG signals travel through multiple layers of organic matter that may slightly alter the signal. MEG does not have this disadvantage. MEG has a relatively good temporal resolution and a spatial resolution slightly better than EEG. The downside of MEG is the low mobility, as participants have to be seated in a machine (see figure 2.3.4.1.).

2.3.5 Preliminary Conclusions Neuroimaging Techniques

For the current project, it was chosen to use EEG as the preferred neuroimaging technique. This is due to several reasons; (i) Currently EEG is the only neuroimaging technique that is being used for neurofeedback training of interpersonal neural synchrony (see section 3, Recent Work); (ii) EEG materials can be set-up relatively quickly, which is important for the duration of the study; (iii) The availability of EEG materials and (iv) the relatively good temporal resolution of EEG.

The following subsections will go more in depth about EEG signal processing and how to use EEG as a neuroimaging technique.

2.4 Neural activity related to EEG

The neural activity that is measured through EEG originates mostly from the pyramidal neurons in the cerebral cortex [72]. All neurons have three basic parts: the cell body, axon and dendrites (see figure 2.4.1), and are able to communicate with each other due to the connection of axons and dendrites. When neurons are activated local current flows are produced called action potentials (AP). When neurons get activated by neurotransmitters (chemicals) at the end of an axon of other neurons a different type of potential is activated namely postsynaptic potential (PSP).

The electrical activity that can be recorded through an EEG are low frequency summed inhibitory and excitatory PSPs that create electrical dipoles between the cell body and dendrites. Due to APs having a much smaller potential field distribution and shorter time span, APs do not significantly contribute to EEG recordings. The amplitude of an EEG signal typically ranges from about 1 to

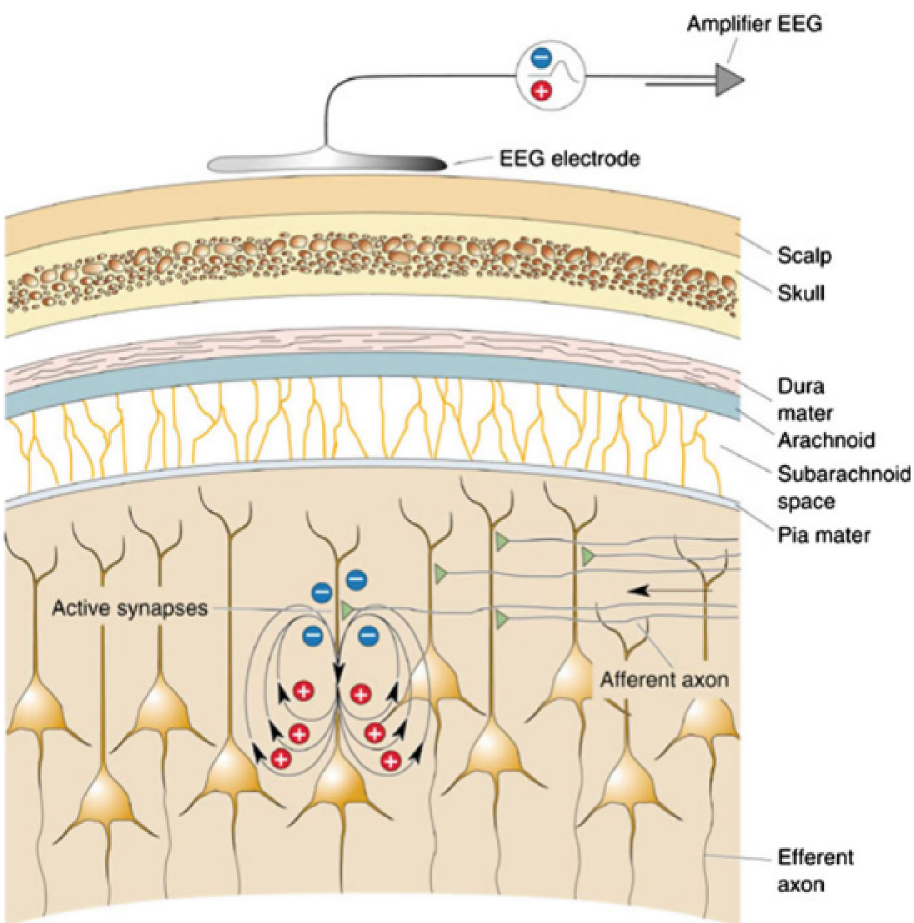


Figure 2.3.3.1: Illustration of the measurement of electrical brain signals through an EEG electrode. The pyramidal brain cells generate small electrical fields, which can only be measured by the electrode if thousands of cells activate simultaneously [103].

Figure 2.3.4.1: A MEG recording session with an eye tracking device [88].

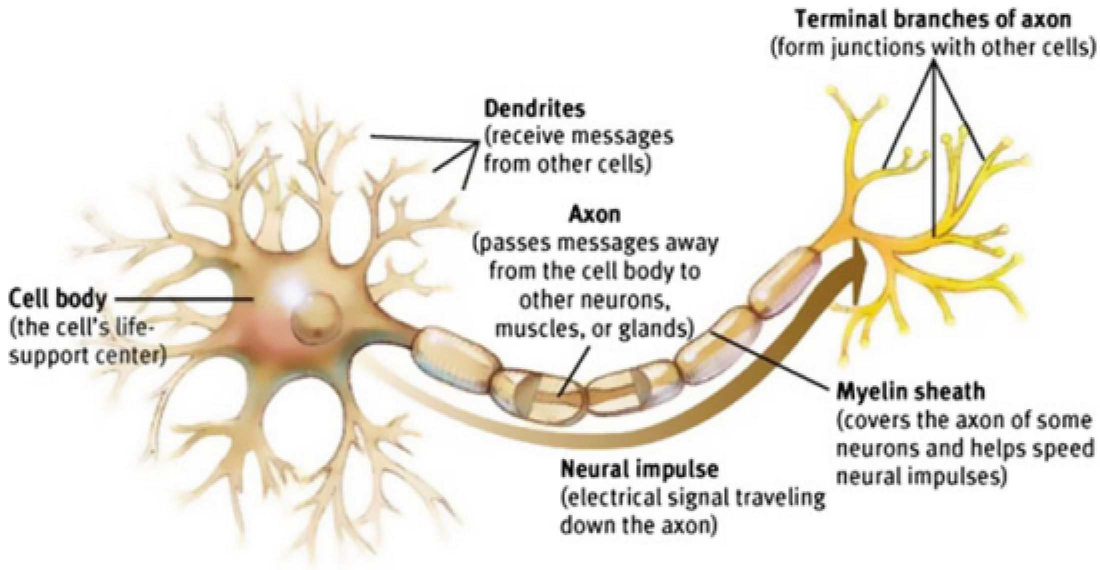


Figure 2.4.1: The cell body, axon and dendrites of a neuron with neighboring neurons [72]).

100 µV in a normal adult. Although PSPs make the largest contribution to EEG signals, other neural processes including calcium and sodium spikes, glial cells, active as well as passive currents, and mono/quadripolar sources may also contribute to some extent [18]. It is also important to note that EEG does not measure clean signals, as signals have to travel through various tissues such as the scalp, skull, cerebrospinal fluid and different parts of the brain which all have different conductivity characteristics and therefore attenuate the current to a different extent.

EEG records these potentials using an array of electrodes, from which topographical maps can be constructed that display the distribution of the scalp potential produced by the active neuronal population at a point in time. If one brain area is active, the potential distribution on the scalp is rather simple and dipolar. However, if several brain areas are simultaneously active, complex patterns of scalp potentials may arise. In the latter case, it may become complicated to deduct the exact source of the brain activity. In general, a priori assumptions are required to read EEGs, preferentially incorporating anatomical, physiological, and biophysical knowledge.

EEG has a very low signal-to-noise ratio, and it cannot be used in its raw form to measure most of the neural processes that relate to a specific event or task [103]. However, it is possible to measure how brain activity changes as a response to a certain stimulus or event. The phase-locked change of EEG activity that is related to stimuli or events is called an event-related potential (ERP), and it is possible to extract time-locked ERPs by averaging techniques. If a change in EEG is not phase-locked to the stimulus onset it is called an event-related oscillation (ERO). As ERPs are phase-locked they can be easily observed in the time domain after averaging across trials.

2.4.1 Neural oscillations

Neural oscillations are the most prominent feature of EEG data, and it has been frequently demonstrated that specific oscillations correlate with particular perceptual, cognitive, motor, and emotional processes. Over the years it has been established that different neurons have different preferred frequencies and that different frequencies dominate in different brain regions. There are some particular frequencies that are studied more often than others, and this subsection will go more in depth about the various frequency types.

Delta (0.5 Hz - 4 Hz)

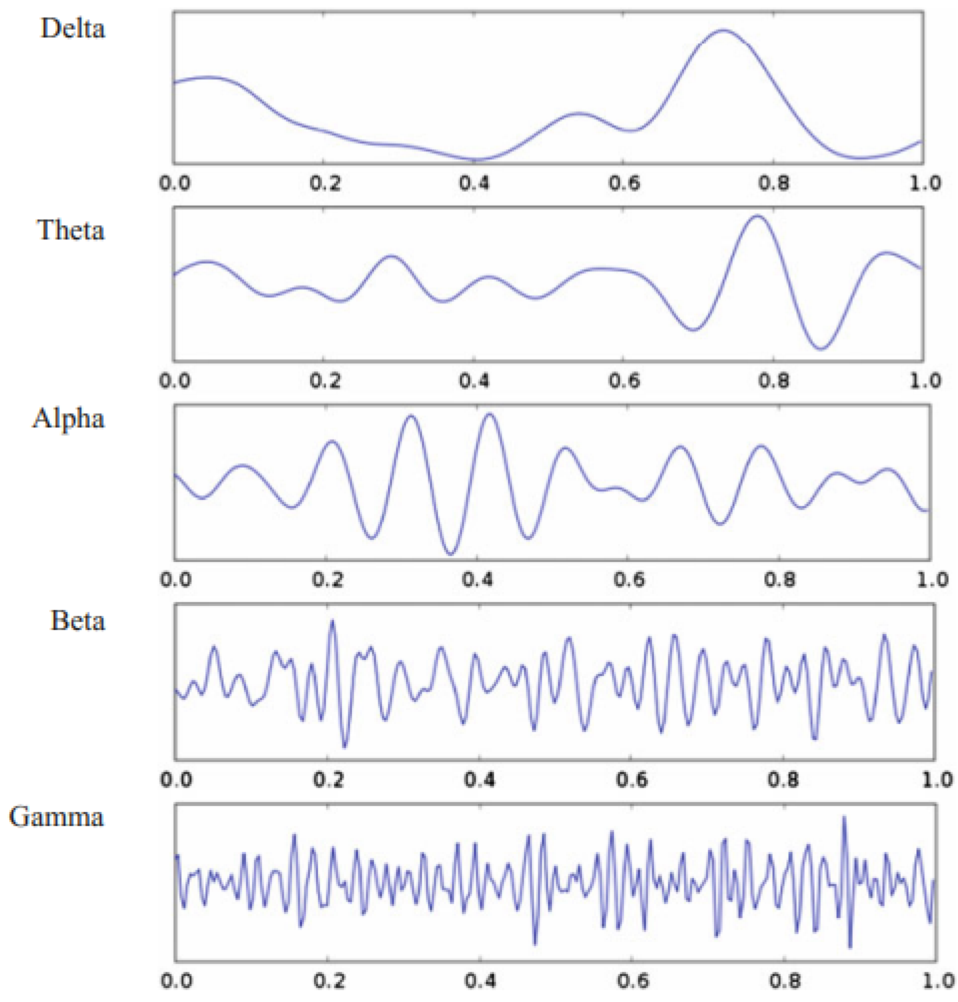


Figure 2.4.1.1: The five most commonly studied EEG frequency types [103].

Delta oscillations are related to deep sleep and unconsciousness. Sources of delta oscillations are often found in the frontal and cingulate cortex and span a wide region of neural networks, possibly in an inhibitory manner [41]. It has been hypothesized that low frequency oscillations of delta and theta ranges are associated with motivational and emotional processes [58]. Typical tasks in which delta can be measured are Go / no-Go tasks, which are known to increase delta when a Go target is detected. Go/No-Go tasks are tasks in which the participant has to inhibit action until he or she detects a Go target, which signals to the participant that they have to perform an action. Delta band activity in the frontal lobe has been related to emotional processes and the motivational system [58]. Moreover, delta and theta bands have been associated with motivational and attentive responses evoked by relevant emotional stimuli [6].

Theta (4 Hz-8 Hz)

Theta is most commonly associated with memory processes [54] but also with emotional states [57] and emotion regulation in the PFC [28]. A common assumption is that cortical theta reflects the communication between the hippocampus and other regions. The hippocampus is a region that is known to serve memory functions and to exhibit oscillations in the theta range [74]. Another assumption is that it may be involved in inhibitory processes in the frontal cortex, as it is commonly observed during inhibition of executive functions.

Interestingly, it has been shown that increased frontal theta has been associated with strategic control and conflict monitoring in social contexts [10]. Increased theta has also been found during empathic processes, such as empathy for pain (Mu, Fan, Mao, & Han, 2008). And like mentioned in the delta section, theta is also associated with motivational and attentive responses evoked by relevant emotional stimuli [6].

Alpha (8 Hz - 12Hz)

This wave promotes focusing, diminishing emotions, thoughts and distractions [104], memory [53]) and attentional processes [55]. Alpha is the dominant frequency in the human scalp EEG of adults, only with the exception of irregular activity in the delta range and below [54]. EEG alpha oscillations can be modulated during sensory stimulation [94]. There are studies to suggest that several distinct mechanisms could produce alpha oscillations, including thalamocortical loops, rhythmically firing pyramidal cells, local interneurons, and interactions of synaptic inputs with different time-constants [18]. There is a lack of convergence on a single cellular mechanism which suggests multiple distinct mechanisms of alpha [18]. Features of alpha oscillations such as amplitude, time-course, and peak frequency can vary as a function of cognitive task, cortical region, and cortical layer. Moreover, alpha is the only oscillation (apart from beta during movement) which typically responds with a pronounced event-related decrease in band power - an 'event related desynchronization' (ERD) - during cognitive tasks [56]. Other frequencies such as gamma, theta and delta typically respond with an increase in band power (ERS) [56].

SMR (Sensorimotor Rhythm, 12 Hz-15 Hz)

The Sensorimotor Rhythm (SMR) is sometimes called low beta. The neuronal mechanism by which SMRs are formed is the interaction between the cerebral cortex and the thalamic ventrobasal complex relay nuclei [14]. This rhythm is related to body movement and concentration and it can be modulated through motor imagery [36].

Beta (13 Hz - 30Hz)

Beta oscillations originate from the somatomotor cortex, the cerebellar system and basal ganglia (BG). It is related to states of wakefulness and mental activity, states of alertness, attentiveness and active concentration [116].

Gamma (30 Hz - 140 Hz)

Faster gamma-band oscillations are associated with cortical activation, attentive processing of information, active maintenance of memory contents and conscious

It has been speculated that slow frequencies may represent cooperation of large neuronal networks in the brain whereas high-frequencies may predominantly reflect the activity of local neuronal populations [41]. Additionally, coherent EEG oscillations in two distant brain regions may reflect the functional cooperation of these two regions [41].

2.4.2 Properties of oscillations

The EEG is a complex and compound signal, consisting of a superposition of multiple signals stemming from different sources [56]. Investigating task-related phases in compound signals is most commonly meaningless, as compound signals get easily distorted due to the superposition with other frequencies. Fortunately, various frequency bands can be segregated through the use of data-driven methods. As the phase of an EEG signal is highly frequency specific and time critical, this is an important step. The frequency specificity of an oscillation can be blurred by averaging over trials due to frequency shifts between trials [56]. Indeed, the frequency of an oscillation is not fixed, and it is known to vary depending on the state of the participant and the task, even within a very short time span.

According to Klimesch [54], there are two types of oscillations that can be measured during an experiment: oscillations measured during rest and during task-specific conditions

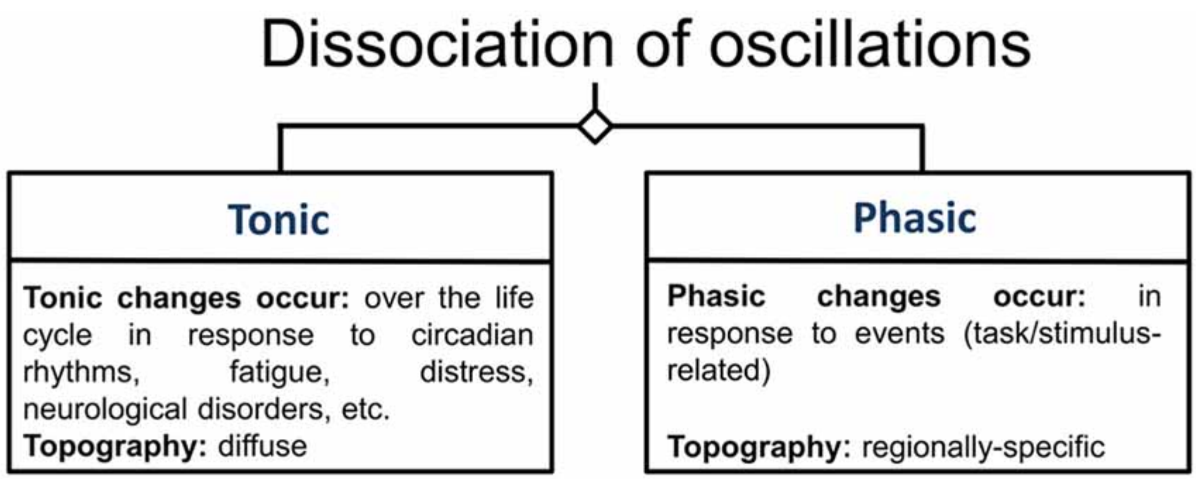


Figure 2.4.2.1: According to Klimesch [54], oscillations can be differentiated according to whether they are measured during rest (tonic) or whether they are related to specific task-conditions or stimuli (phasic).

2.4.3 Artifacts

Signals that originate from non-cerebral origin in EEG data are called artifacts. Artifacts can be grouped into two categories: physiological artifacts (caused by the participant's body) and non-physiological artifacts. The most common artifacts are caused by the excitation of the eyeball's muscles (e.g. blinking), bad contact of the electrode and scalp, swallowing and bad contact of reference electrode and skin (see figure 2.4.3.1.). Other artifacts include scalp perspiration (around 0.5 Hz), (head) movement, heartbeats, power line interference (50 Hz in Europe, and 60 Hz in the United States), Malfunction of any part within EEG recording system and loose wiring or loosening of circuit board connections. The amplitude of an artifact is related to the amplitude of the signal of interest.

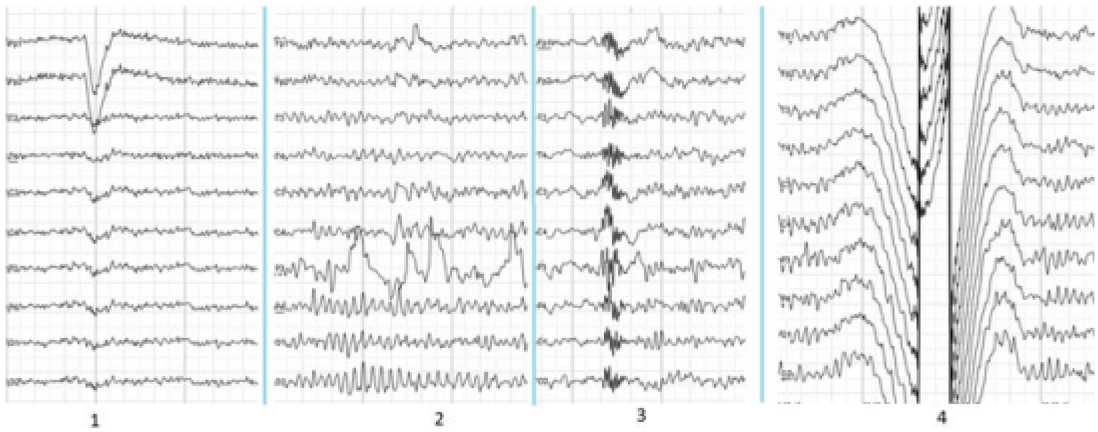


Figure 2.4.3.1: Artifact caused by the excitation of eyeball's muscles. 2. Artifact caused by bad contact of P3 electrode. 3. Swallowing artefact. 4. Common reference electrode artifact [103].

2.5 Processing of EEG data

2.5.1 EEG Electrode placement

EEG electrodes are most commonly placed on the scalp according to the International 10–20 System, which is an internationally recognized method to describe and apply the location of scalp electrodes in the context of EEG studies (see figure 2.5.1.1.). It was developed to maintain standardization across studies in order to facilitate the replication, compilation and comparison of outcomes of clinical or research studies [103].

The system is based on the relationship between the location of electrodes and the underlying areas of the brain. The name of the system refers to the distances between adjacent electrodes, as these are either 10 or 20 percent of the total front–back or right–left distance of the skull. If more precise measures are required, the modified combinatorial nomenclature (MCN) can be used, which fills intermediate sites with extra electrodes halfway between those of the 10-20 system [103].

Both systems use a letter to represent the specific lobe or area that is covered by the electrode: frontal (F), temporal (T), parietal (P), occipital (O) and central (C). The "central" area that is covered by C electrodes consists of central parts of different lobes. Suffixal "Z" electrodes are placed on the midline sagittal plane of the skull (Fz, Cz, Pz and Oz), and are mostly used as reference electrodes. The MCN system uses the numbers 1, 3, 5, 7, and 9 for the left hemisphere and 2, 4, 6, 8 and 10 for the right hemisphere.

The differences in electrical potentials between electrodes constitute channels, and combinations of channels are called montages. The number of channels in a montage is variable and depends on the number of electrodes available and the purpose of the recording. There are two main types of montages: bipolar montages, which are arranged in chains following an anterior-to-posterior or transverse arrangement, and referential montages, in which each channel represents the differential electrode potential of any given electrode to a single chosen reference electrode [103]. Bipolar montages have the advantage of having a lower susceptibility to external or environmental noises, while referential montages can detect both local (near field) and distant (far field) potentials.

2.5.2 Filtering

Often noise is present in EEG data, such as 50 Hz or 60 Hz line noise, high-frequency noises and noises at very low frequencies, which can effectively be reduced by appropriate band-pass filtering [103]. There are four types of filters: low-pass filter, high-pass filter, band-pass filter and band-stop filter (see figure 2.5.2.1). The low-pass filter only keeps signals with low frequencies below a certain

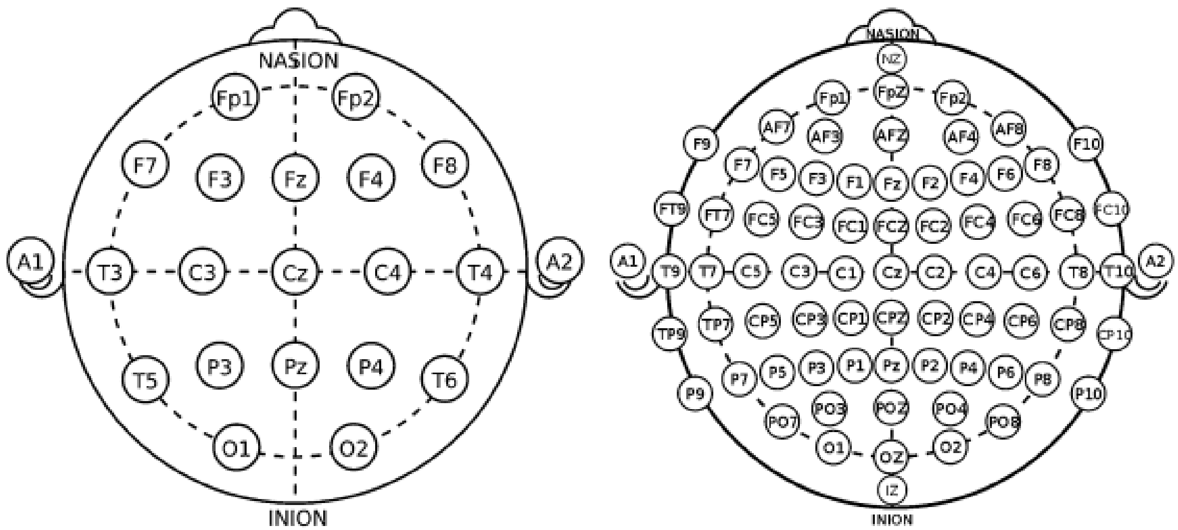


Figure 2.5.1.1: Left: The International 10-20 electrode placement system. Right: The modified combinatorial nomenclature (MCN) [103].

value, and high frequencies greater than the certain value are removed. High-pass filters do the opposite: only signals with frequencies greater than a certain value are kept, while only those with low frequencies below the certain value are removed. Band-pass filters keep only signals between a lower and upper bound, and remove signals below the lower limit or greater than the upper limit. The band-stop filter removes signals with a frequency between the lower and upper bound and keeps all other signals.

When selecting the appropriate filter, the frequency ranges of the artifacts embedded in the EEG recordings should be taken into consideration [103].High-pass filters with a limit of 0.1 Hz can be applied to remove low-frequency drifts, while low-pass filters can be applied to remove high-frequency noise (e.g., interference due to muscular activities). Power line noise (50 Hz in Europe and Asia, 60 Hz in the United States) is commonly removed through a band-stop filter. The frequency range of interest is also important in filter selection. If only alpha frequency is of interest, then a band-pass filter within alpha frequency range is the most appropriate. It is recommended to apply the filtering to the continuous EEG data as the first step of the preprocessing pipeline, and especially before segmenting the continuous EEG data into epochs. If EEG epochs are filtered it will introduce filtering artifacts at epoch boundaries. One methodology to apply filters to EEG data is through a Finite impulse response (FIR) filter, which is suitable in phase-sensitive applications [95].

2.5.3 Referencing

Referencing is an important step in the preprocessing pipeline as it removes noise from the EEG data [103]. Referencing can be done online and offline. Commonly, one of the following references are used: left or right mastoid, linked mastoids, the vertex electrode, average referencing, single or linked earlobes and the nose tip. Referencing is a linear transformation, and any activity present in the reference electrode will be reflected as activity in all other scalp electrodes. Therefore, when choosing a reference electrode, the reference electrodes should be properly placed and should have a good signal, because noise in the reference electrode will turn into noise in the scalp electrodes. When choosing a reference it is also important that the reference electrode is placed far away from the locations of signals of interest so that the reference electrode has little influence on the signals of interest. It should also not be placed too far from the scalp. Theoretically, a reference can be

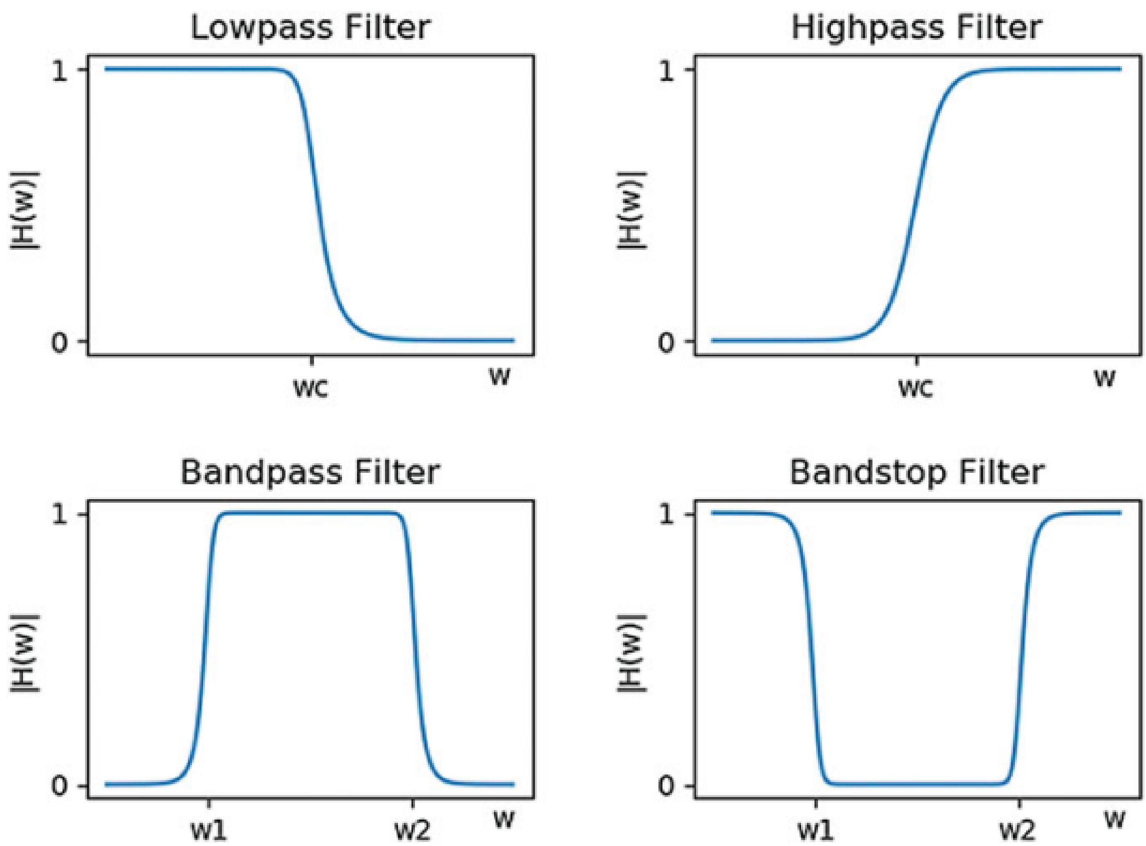


Figure 2.5.2.1: The Lowpass, Highpass, Bandpass and Bandstop filter with wc, w1 and w2 as threshold frequencies and H(w) = 1 entailing pass and H(w) = 0 entailing no pass [103].

placed anywhere on the body or on other objects such as a wall. However, this is not recommended as the goal of referencing is to remove noise which decreases when moving farther from the scalp.

Mastoids (the bone behind the ear) are frequently used as reference electrodes since they are often relatively far away from the locations of signals of interest. Referencing to one lateralized site is generally not recommended, because this will lead to a lateralization bias in the data. The average of the two earlobes is commonly used. Another method that is mostly recommended when measuring a large number of electrodes (e.g., > 100) is average referencing. Average referencing relies on the assumption that the sum of the electric field values recorded at all scalp electrodes (sufficiently dense and evenly distributed) is always 0, and the current passing through the base of the skull to the neck and body is negligible. This assumption is not always correct, but is the most accurate when the number of electrodes is very high. Another technique is the reference electrode standardization technique (REST) which standardizes a reference of scalp EEG recordings to a point at infinity. Importantly, no single reference is perfect and the choice depends on the signal of interest [103].

2.6 Interpersonal Neural Synchrony

Interpersonal Neural Synchrony (INS) is a relatively new topic in the field of Neuroscience, as it originated in 2002 with the introduction of the hyperscanning technique (see subsection 2.6.1. Hyperscanning) [75]. INS is sometimes called Brain-To-Brain Coupling or Interbrain Synchronization, and is characterized by measuring the similarity between interpersonal neural activity. Most commonly the brain activity of different individuals is measured at the same time and compared through a similarity measure. This method is called hyperscanning (see subsection 2.6.1. Hyperscanning). Non-simultaneous interpersonal brain activity comparisons are also being performed, but are less common [19].

The reason that INS has become popular in recent years is that many social phenomena seem to be related to the degree of INS (see introductory section). It has been speculated by many authors that INS may be beneficial for social interactions as there are many articles published about the correlations between INS and some pro-social outcome (e.g. [31, 44]. Other articles try to disprove this argument, demonstrating that synchronization also can lead to more negative forms of social influence such as increasing the likelihood to comply with a partner's request or to aggress towards third persons [59]. Therefore, it has been argued that INS may only enhance relationship-specific cohesion while not necessarily increase pro-social behavior.

While there is an increasing amount of publications on the topic of INS, to this day there is no clear understanding of why INS occurs between individuals. A widely supported theory is that INS depends on shared or joint attention, and occurs when neural oscillations get "tuned" to the temporal structure of surroundings [9]. Temporally aligned entrainment to the oscillatory features of external stimuli is thought to support information extraction from the stimulus and attentional selection of relevant information [9], 2019). However, even if the tasks performed by two people are not identical, INS seems to be significantly higher when working together [124]. It has been believed for a long time that INS is only present during real social interaction [129], but recent work also suggests that INS can be initiated through Virtual Reality [37], online gaming [124] and NFT [77] with the absence of other forms of communication between individuals.

There have also been criticisms of the performed studies in the field of INS. One major criticism is that many studies in INS do not control for spurious synchronization, which is INS caused by random coincidence [132].

2.6.1 Hyperscanning

Hyperscanning is the simultaneous measuring of the brain activity of two or more individuals and is commonly used in INS research [13]. The advantage that this relatively new technique offers is that it allows the investigation of real-time dynamics between two or more interacting brains [21]. It is also useful in intra-brain research, as it can also compare the activity between different structures of an individual brain. However, for the purpose of this study only inter-brain hyperscanning is discussed. The main goal of inter-brain hyperscanning studies is to reveal neural underpinnings of social cognitive functions during real-time social interactions.

Various neuroimaging techniques can be utilised for hyperscanning, but the most common techniques are EEG, MEG, fMRI and fNIRS [21]. The choice of neuroimaging technique depends on the research question and signal of interest. The current study will focus on EEG hyperscanning, which was first reported in Babiloni et al. (2006).

Figure 2.6.1.1: A Virtual Reality/EEG hyperscanning set-up.

Arguably one of the most important steps in hyperscanning research is choosing the right similarity measure. The similarity between brain signals (i.e. the *neural coupling*) is most commonly measured by comparing the difference in phase (phase-coupling) or amplitude (amplitude-coupling) [101]. Hyperscanning paradigms that measure phase-coupling compute the phase-alignment of different signals in a particular frequency. When measuring amplitude-coupling, the temporal co-modulation of the amplitude (or power) of neuronal oscillations is measured. Both amplitude and phase-coupling reflect neuronal interactions, but may also regulate these interactions by temporally aligning distant processes associated with fluctuating oscillations [101]. The following subsection will discuss some of the most common amplitude- and phase-coupling methods.

2.6.2 Similarity measures

Neural coupling measures can be differentiated by the oscillation property that they measure (mostly amplitude or phase) and also by the relationship between signals: *functional connectivity* and *effective connectivity*. Functional connectivity refers to the statistical dependence between oscillations and is commonly measured by a type of correlation measure. Effective connectivity refers to a directed causal interaction between oscillations, and this type of method can be used to estimate causal links between brains and brain regions. Obviously, it depends on the research question whether functional or effective connectivity measures should be used. As only functional connectivity measures are of interest for the current study (measuring correlation instead of dependence), only functional connectivity measures will be discussed in the current section.

2.6.2.1. Phase-coupling measures

Phase-based similarity measures calculate the consistency of phase difference between two signals. The most widely used phase-based measure for hyperscanning studies has been the Phase Locking Value (PLV) [61] due to it being relatively well suited for capturing the rapid flow of information between people in social situations.

2.6.2.2. Phase Locking Value (PLV)

The Phase Locking Value (PLV) is a commonly used phase-based similarity measure originally defined by Lachaux et al. [61]. It measures how two signals from different brains are phase-locked (or have a similar consistency of phase difference) in an observed time window. The PLV is estimated by: ,

$$PLV_n \ = \frac{1}{N}\left|\sum_{k=1}^{N} e^{i} \ (\phi_{(t,k)} - \psi_{(t,k)})\right| \tag{1}$$

where N is the number of trials,$\phi_{(t,k)}$, is the phase on trial, n at time t, in channel ω and $\psi_{(t,k)}$ in channel ψ. The PLV$_n$ varies between 0 and 1 where 1 indicates perfect phase locking and 0 indicates no phase locking. PLV is known to suffer from the "common source" problem. This problem arises due to nearby electrodes measuring activity from the same (common) sources, giving rise to spurious correlations between time series [107]. A similar algorithm that does not suffer from the common source problem is the Phase Locking Index (PLI). However, as the compared electrodes in inter-brain hyperscanning are from different individuals and thus locations, the common source problem is non-existent and there is no reason to use PLI instead.

This form of the PLVn is based on the phase difference across trials, which makes it only suitable for event-related paradigms. However, there is a variant of the equation (2) which involves averaging the instantaneous phase differences over time within a single trial, making it a measure of the intra-trial consistency of the phase difference between channels:

$$PLV_t \ = \frac{1}{T}\left|\sum_{n=1}^{T} e^{i} \ (\phi_{(t,n)} - \psi_{(t,n)})\right| \tag{2}$$

where T is the number of time points.

2.6.2.3. The problem of covariance

PLV and many other phase-based similarity measures suffer from another, well-known problem. This problem becomes clear when examining figure 2.6.2.1. It can be observed that the four pairs of clocks in the image are in a similar phase to each other, and if measured through for example PLV, the similarity score would suggest that the clocks move in synchrony. However, the relationships between the pairs of clocks are very different for the pairs A, B, C and D. In the case of A, the clocks have a reciprocal relationship, due to which they swing in phase. In B, the clocks have a common external influence that makes them swing in phase. In C, there is a causal relationship with the red clock influencing the phase of the blue clock. In D, the similarity of phase is due to coincidence.

This suggests that simply observing a consistent phase relationship between two oscillators (e.g. clocks or brain activity) does not necessarily mean that they are in the same condition of reciprocal information exchange. In conclusion, PLV cannot distinguish between coincidental phase synchronization and true phase synchronization, which means that it does not necessarily imply covariance between two signals [13].

2.6.2.4. Circular Correlation Coefficient (CCorr)

The Circular Correlation Coefficient (CCorr) is different from many other power-based similarity measures as it does not suffer much from the problem above (i.e., not being able to distinguish between coincidental phase synchronization and true phase synchronization). The reason for this is that CCorr measures the circular covariance of differences between the observed phase and the expected (i.e., mean) phase instead of the phase difference. As EEG signals are not "perfect"

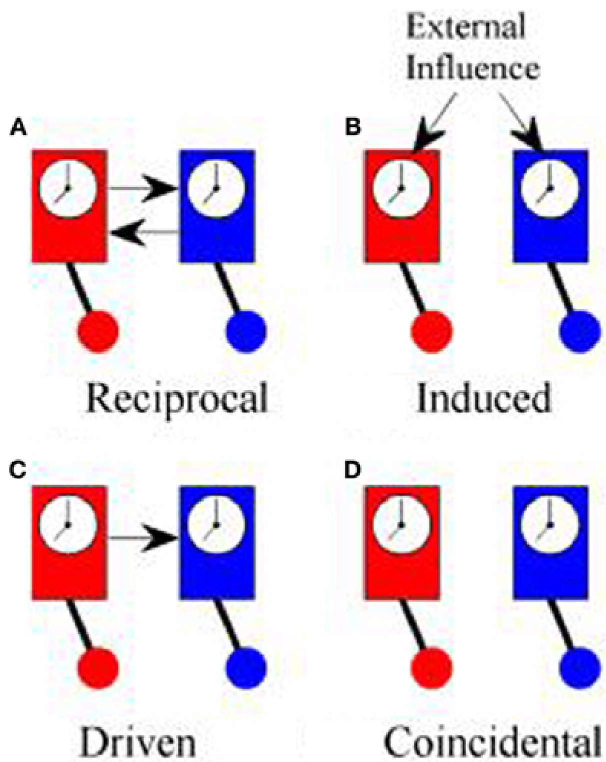


Figure 2.6.2.1: Four pairs of phase synchronization that are due different relationships. A. the clocks have a reciprocal relationship, due to which they swing in phase. B. the clocks have a common external influence that makes them swing in phase. C. there is a causal relationship with the red clock influencing the phase of the blue clock. D. the similarity of phase is due to coincidence [13].

oscillations and tend to differ slightly at every oscillation, the difference (or variance) between the mean and actual phase also tends to differ. Using CCorr, this difference is measured for two channels which are compared to each other in terms of covariance. In the case of two unrelated channels, the phase variance will not co-vary and the CCorr will be close to zero. If the channels have related oscillations, the phase variance will co-vary and the CCorr would be closer to 1.

It is clear that the CCorr measure is more insensitive to the problem of coincidental phase synchronization than PLV, as CCorr measures the covariation of phase variance while PLV solely measures the phase difference between channels. This results in PLV being worse at discriminating between related and unrelated signals than CCorr.

The CCorr measure is a direct parallel to the Pearson Product Moment Correlation Coefficient for circular data and is given by:

$$CCorr_{\phi,\psi} = \frac{\sum_{k=1}^{N} \sin(\phi - \overline{\phi}) \sin(\psi - \overline{\psi})}{\sqrt{\sum_{k=1}^{N} \sin^2(\phi - \overline{\phi}) \sin^2(\psi - \overline{\psi})}} \tag{3}$$

Where ϕ and ψ represent the phase angles for channels 1 and 2 respectively. Ccorr values can range from 0 to 1. CCorr measures the phase variance of two oscillators, which co-vary over time if these are related. In the case of two unrelated channels, the phase variance will not co-vary and the CCorr will be zero.

In addition to the insensitivity to coincidental phase, CCorr is also relatively insensitive to changes in the marginal distributions of deviations from the expected phase [13]. The reason for this is that for oscillatory data such as the EEG, the phase is approximately uniformly distributed, which entails that the mean is not defined. Any arbitrary phase angle may be defined as the mean without a large consequence according to Burgess (2013). However, it has been recently argued that the CCorr measure is likely to fluctuate for EEG data as a result of the variability of expected phase due to shifts in time window onset [133]. This is illustrated in figure 2.6.2.2, which shows the CCorr measure (red, figures E and F) and an adjusted CCorr measure (blue, figures E and F) plotted for the movement trajectories of two participants (green and purple, figures A and B) in a study by Zimmermann et al. (2023). The CCorr and adjusted CCorr measures had been plotted for two conditions: individual, unrelated movement (left) and synchronized, related movement (right). Figures C and D show the actual phase of movement, and E and F show the correlation index in CCorr and adjusted CCorr. It can be seen in sub-figures E and F that CCorr is more likely to fluctuate when the time window is shifted than the adjusted CCorr meaure.

The adjusted CCorr measure by Zimmermann et al. can be calculated as:

$$P_c = \frac{(R_{\phi-\psi} - R_{\phi+\psi})}{2\sqrt{\sum_{k=1}^{N} \sin^2(\phi - \overline{\phi}) \cdot \sum_{k=1}^{N} \sin^2(\psi - \overline{\psi})}} \tag{4}$$

with

$$R(a \pm b) = |\sum_{k=1}^{N} e^{i(a \pm b)}|$$

Where ϕ and ψ represent the phase angles for channels 1 and 2 respectively, and the values range from 0 to 1.

The authors also compared adjusted CCorr to CCorr in terms of accumulated change over time for the same dataset, which can be seen in figure 2.6.2.3. The results in CCorr and adjusted CCorr (y-axis) are visualized for individual and synchronized movement (x-axis) and for two

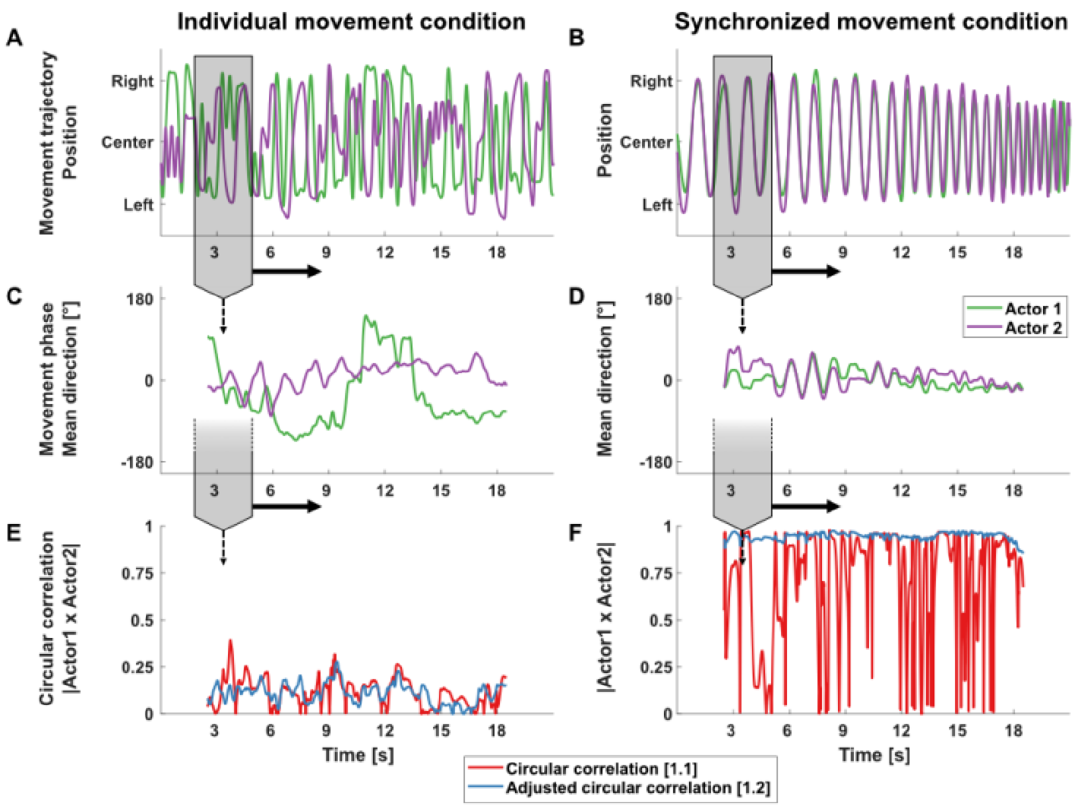


Figure 2.6.2.2: The evolution of the Circular Correlation Coëfficient (E and F, red) and the Adjusted Circular Correlation Coëfficient (E and F, blue) for individual movement (left) and synchronized movement (right), with time on the x-axis (in seconds). The Adjusted CCorr measure is created by Zimmerman et al. to correct for the variance in expected phase of CCorr. A and B: The movement trajectory of two participants (green and purple) on the y-axis (right, center, left). C and D: The mean or expected movement phase on the y-axis (in degrees) for two participants (green and purple). E and F: The correlation index is plotted on the y-axis (CCorr in red, adjusted CCorr in blue). The time window at which the correlation index is calculated is of a length of 5 seconds, and it changes its offset over time. It appears as that Adjusted CCorr is less likely to fluctuate than CCorr [133].

conditions: variable time window onsets (figure 2.6.2.3., left graph) and variable epoch lengths (figure 2.6.2.3., right graph). The accumulated change in adjusted CCorr was significantly smaller than the accumulated change in CCorr, which could be seen for both conditions and also in artificial datasets [133].

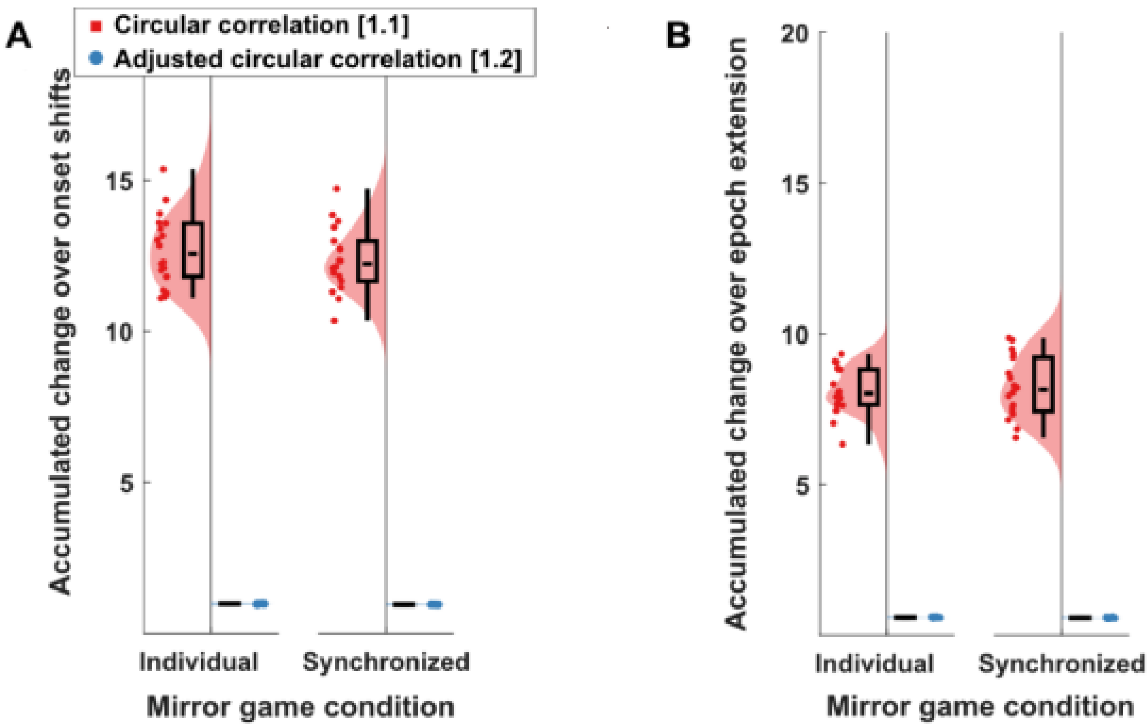


Figure 2.6.2.3: The accumulated change in the circular correlation coëfficient (red) and the adjusted circular correlation coëfficient (blue) on the y-axis, plotted for two conditions (individual movement and synchronized movement) for two different variations (A and B) of the same EEG dataset. A: This dataset consists of 512 epochs with a length of 1 second. For this variation, one sample (at a sampling frequency of 256 Hz) was shifted for each epoch. It can be seen that the accumulated change in CCorr is significantly larger than the accumulated change of the adjusted CCorr for both conditions (individual and synchronized movement). B: The dataset consists of 512 epochs with a variable epoch length. The first epoch starts at an epoch length of 1 second, and the epoch length increases by one sample per epoch. The longest epoch is three seconds. The accumulated change in CCorr is significantly larger than the accumulated change of the adjusted CCorr for both conditions [133].

2.6.2.5. Coherence (COH)

Coherence (COH) is the traditional Fourier-based method of connectivity and the Welch estimate of coherence. It is given by:

$$COH_{x,y} = \frac{\left| \frac{1}{N} \sum_{k=1}^{N} Y_k(f) X_k^*(f) \right|}{\sqrt{\frac{1}{N} \sum_{k=1}^{N} X_k(f) X_k^*(f) \cdot \frac{1}{N} \sum_{k=1}^{N} Y_k(f) Y_k^*(f)}} \tag{5}$$

where X_i (ω) denotes FFT of the kth segment of the time series $x(t)$ at frequency f and * indicates the transpose and complex conjugate. COH values range from 0 to + 1, and represent the cross-power spectral density. However, it assumes that the covariance between signals is stationary throughout an epoch, which is not always the case .

2.6.2.6. Wavelet Transform Coherence (WTC)

Wavelet coherence, also known as Wavelet Transform Coherence (WTC), is a phase-based similarity measure that is commonly used when comparing signals from fNIRS [20]. It is probably the most commonly used method in fNIRS-based hyperscanning [21]. The reason for this popularity is likely that this similarity measure is invariant to interregional differences in the hemodynamic response function (HRF) [79].

It measures the cross-correlation between two time series as a function of frequency and time. WTC can be thought of as the local correlation between two time series [20]. The wavelet coherence at time t and frequency f can be estimated through:

$$R_n^2(s) \;=\; \frac{\left|S(s^{-1}W_n^{XY}(s)\right|^2}{S(s^{-1}\left|W_n^X(s)\right|^2)\cdot S(s^{-1}\left|W_n^Y(s)\right|^2)} \tag{6}$$

where S is a smoothing operator and $\left|W_n^X(s)\right|^2$ is the wavelet power of which $W_n^X(s)$ is the local phase [35]. $R_n^2(s)$ estimates a similarity value between 0 and 1.

2.6.2.7. Amplitude-coupling measures

A downside of phase-based similarity measures is that they assume that the connectivity is instantaneous and at the same frequency. Power-based similarity measures do not have this constraint. This method also ignores the temporal structure of the data.

2.6.2.8. Spearman coefficient

The Spearman coefficient measures the correlation between the power of two signals and is estimated by:

$$r = \frac{\sum_{t=1}^{N}(\phi_t - \overline{\phi})(\psi_t - \overline{\psi})}{\sum_{t=1}^{N}(\phi_t - \overline{\phi})^2 \sum_{t=1}^{N}(\psi_t - \overline{\psi})^2} \tag{7}$$

The numerator represents the sum of variables ϕ times variable ψ at each time point

(or trial) t after subtracting the mean of each variable, and the denominator represents the variance of each electrode.

2.7 Neurofeedback Training

In the last decades Neurofeedback Training (NFT) research has been expanding from only being used for prevention or treatment of mental illnesses to improving healthy individual's performances. The objective of NFT is to provide the users with operant control of their brain activity and thereby gain control over processes usually not available for conscious regulation [123]. Different behavioural changes have been observed to result from NFT, and it has been used as a therapy for various mental conditions like Attention Deficit Hyperactivity Disorder (ADHD) and depression, reportedly giving good results. NFT is always used in real-time.

An overview of typical neurofeedback procedures can be viewed in figure 2.7.1. The process of NFT starts with the measurement of brain activity of a participant or patient. This can be done through various neuroimaging techniques, such as fMRI, EEG, fNIRS and MEG. A particular time window of this signal is then processed and some type of feedback is calculated, which depends on the goal of the training. Features from a set of sensors, such as the power at a frequency window, can be classified as multivariate patterns of activity (MVPAs). The MVPA classifier is used to calculate the feedback and this feedback is given to the participant through audio, visuals, haptics or electrical stimulation. As NFT is commonly used in children, it is made more engaging by using games as a medium. In NFT games, the feedback is commonly gaming progress if the neural signal is above the desired threshold.

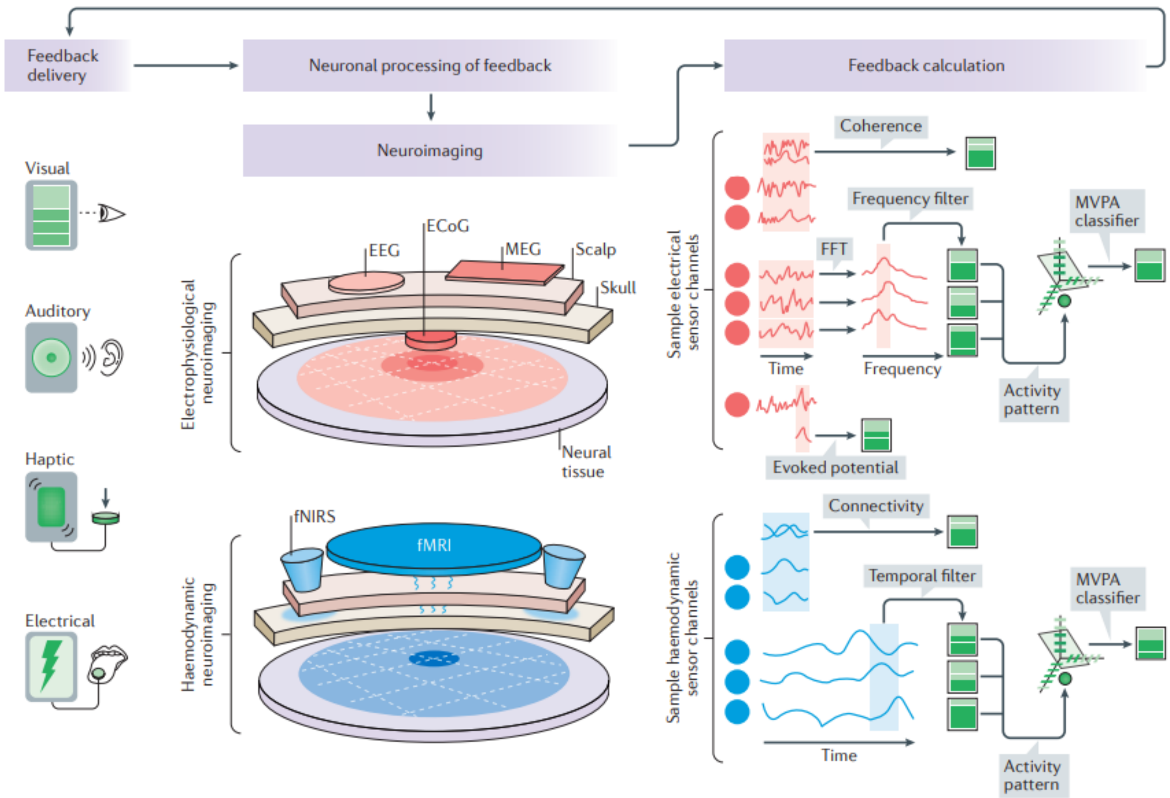


Figure 2.7.1: An overview of most NFT paradigms. Brain activity is measured through neuroimaging, after which feedback is calculated. This feedback is given to the participant through visual, auditory, haptic or electrical cues and the cycle repeats again [102].

2.7.1 Types of brain activity modulation

EEG NFT is one of the most commonly used NFT paradigms, and it enables individuals to manage their brain oscillations. One of the most common goals is to increase the amplitude (or amplitude squared, i.e. power) of a particular frequency in a particular cortical location. Other goals include increment of intracortical neural synchronization. This type of NFT has also been used to make structural changes in gray matter volume and white matter connectivity. One study showed an increase in both one week after beta NFT of frontal and parietal regions. Interestingly, these structural changes were associated with increased visual and auditory attention.

fMRI-based neurofeedback is commonly used to increase or decrease activity in distinct cortical and subcortical regions of interest (ROIs), which has been used to modulate behaviour. It has also been used to increase so-called 'functional connectivity', or correlated activation between two brain regions. One study that demonstrated this enhanced the functional connectivity between two distant neural networks, which was still significantly enhanced after two months. Another study increased top-down connectivity from the dorsomedial PFC to the amygdala, which enhanced the subjective emotional valence ratings from these participants. Another methodology that is closely related to fMRI and EEG- NFT is the combination of these techniques. EEG-fMRI NFT is an emerging approach that has been used in several studies, such as a study by Zotev et al., which used this approach to increase emotion self-regulation [134].

There are studies that have demonstrated an effect of NFT on single cell activation and other types of neural activity. One such study showed that monkeys could be trained to voluntarily increase or decrease the firing rate of neurons in the frontal eye field through auditory feedback and juice rewards. Another study trained mice to enhance spike-related calcium signals measured with two-photon imaging in motor and sensory cortices.

2.7.2 Oscillations and NFT

One of the most common NFT paradigms is to upregulate (or downregulate) brain activity of particular frequencies through EEG. These are some of the most common protocols:

2.7.2.1. Delta

Delta NFT is used to alleviate headaches, traumatic brain injury, learning disorders and hard and sharp contraction of muscles. It is also used to reduce concerns and improve sleep.

2.7.2.2. Theta

Theta upregulation treatments are used to reduce anxiety, depression, daydreaming, distractibility, emotional disorders, and ADHD symptoms.

2.7.2.3. Alpha/Theta

Alpha/theta training is one of the most popular protocols for stress reduction. It is also used to treat depression, addiction and anxiety, or the enhancement of creativity, relaxation, musical performance and healing from trauma. The goal of this training is to increase the theta to alpha ratio. This treatment is typically done with the eyes-closed and with auditory feedback.

2.7.2.4. Alpha

Alpha upregulation is one of the oldest NFT protocols, going back to as far as the late 1960's [82]. A distinction can be made between lower and upper alpha frequencies as it has been demonstrated that these frequencies behave differently with different cognitive demands [29]. It has also been argued that alpha does not have a fixed frequency range, but it is rather different for each individual [54]. Upper alpha upregulation has been used to improve cognitive abilities such as memory performance [29].

2.7.2.5. Beta

Beta training is typically used to improve focus and attention, improve reading ability and to increase school performance. It has been suggested that beta training also can improve computational performances and cognitive processing while reducing worries, over-thinking, obsessive compulsive disorder, alcoholism, insomnia, fatigue and stress.

2.7.2.6. SMR

It has been demonstrated that upregulation of the SMR may decrease reaction time [26] and increase working memory performance in older adults [14]. Moreover, it has been used for the treatment of epilepsy and ADHD [108].

2.7.3 NFT Feedback

The feedback in a NFT protocol is an important feature, yet little is known about which feedback stimuli are preferred over others. It has been established that for the training of alpha/theta, calm auditory stimuli are preferred as the eyes of the participants must often remain closed during the training. In-game feedback is often used in the training of children in order to keep them engaged and motivated (see figure 2.7.3.1) [47]. Other than these restrictions, there are few limitations to the use of feedback. It is essential however that the feedback is easy to interpret in order to not confuse the participant. For this reason, often clear, simple and binary feedback is used, such as the growth of a circle figure if the feedback is positive, and shrinkage if the feedback is negative.

In order to stimulate learning during NFT the threshold value of a particular neural activity is often slowly increased. It has been reported that setting a low initial threshold value is more effective for learning than setting a higher threshold value [78]. There is no golden standard for how much this threshold should increase per trial, but an increase of 3 percent has commonly been reported [49]. Failure to reach the threshold commonly decreases the threshold again to keep participants engaged.

Figure 2.7.3.1: Five games that have been used for NFT feedback [47].

2.7.4 Duration and number of sessions in EEG Neurofeedback Training

There are no clear guidelines in terms of how many sessions should be needed for EEG NFT, and no guidelines for how long each session should last. Past studies have utilized NFT sessions that lasted from minutes to hours. However, there is some consensus that the sessions should not be too short (i.e. a few minutes) as this may be insufficient to allow learning to occur, and not too long (approximately one hour or longer) as it may fatigue the participants.

Overall, it can be determined from the literature that NFT for the purpose of performance enhancement in healthy participants requires less training sessions than for clinical practices. A literature study by Rogala et al. (2016) compared the methodologies from 86 studies that used theta, alpha and beta NFT to enhance performance. The article reports an average training duration of 3.8 to 7.7 sessions for performance enhancement NFT studies [93]. Furthermore, an average interval between training sessions of 2.4 to 3 days was found.

2.7.5 Critiques of NFT

While NFT has been commonly practiced for both clinical practices as well as performance enhancement, the neural mechanisms behind NFT are yet not fully understood. Indeed, there are some concerns for how rapidly NFT has evolved while outpacing a proper understanding of the phenomena behind it. This is in part due to the improper use of control conditions. Not only have control conditions rarely been used in the past, it is also challenging to select a suitable control condition. For example, random feedback may elicit learned helplessness in participants, which is known to lead to passive behaviour and frustration [27]. Learned helplessness occurs when users learn that nothing they did had any effect on success [93]. Naturally, this may lead to a decrease of performance enhancement. Another type of commonly used control condition is training a different frequency band than in the experimental condition. In some cases this method is not suitable as the effects of training that frequency may not be verified, or that they may have unethical consequences (e.g., making symptoms worse in ADHD) [27].

Importantly, NFT is not guaranteed to work, and is even known to not work in approximately 30 percent of all participants [123]. Some individuals are simply not able to learn from NFT. Therefore, participants are commonly screened at the initial stages of training in order to determine whether they are able to learn from NFT. Interestingly, the success rate of being able to up-regulate or down-regulate oscillation amplitude seems to partly depend on the frequency that is being trained (see figure 2.7.5.1.) [93]. Rogola et al. (2016) looked at 86 studies that either up-regulated or down-regulated the theta, alpha or beta frequency during 5 or more NFT sessions, and investigated whether there is a difference between the success rates of these different training paradigms. They found that the success rate of the theta and alpha frequencies are significantly higher than in the beta frequency.

2.7.6 Preliminary Conclusions

Some conclusions can be drawn from the information that has been presented in the current section. Firstly, it can be deduced from 2.1. that interpersonal coordination (by the definition used in the current study) is strongly dependent on timing and shared knowledge. Secondly, from subsection 2.2. It can be deduced that the main areas of interest are the Mirror Neuron System and the Theory of Mind Network. Thirdly, 2.3 suggests that while multiple neuroimaging techniques can be used for the purpose of this study, the most suitable is EEG. Fourthly, subsection 2.4.3. Illustrates that there are multiple similarity measures that can be used to calculate INS values, but that the adjusted CCorr is favoured for the current study as it is insensive to coincidental phase while also being more resistant to the variability of EEG data than CCorr [133].

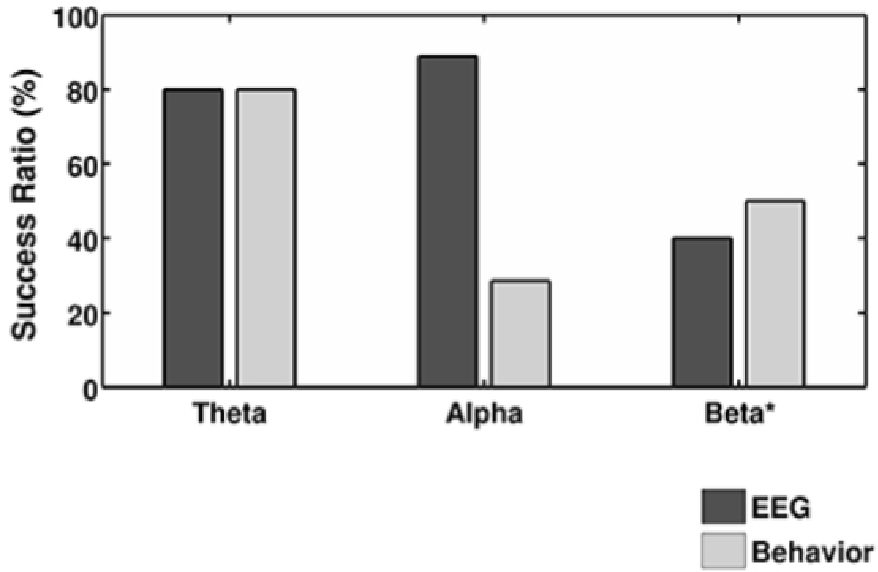


Figure 2.7.5.1: The success ratio of theta, alpha and beta training in up- or down-regulation of amplitude (dark gray) and the succes ratio of behavioural modification (light gray) [93].

2.8 Inclusion criteria

The research papers that were used for the literature research were identified by (i) a search of web resources (Google Scholar) using the following keywords and combinations thereof: Interpersonal Neural Synchronization, Interpersonal Neural Synchrony, inter-brain synchrony, brain-to-brain synchrony, Interpersonal Neuronal Synchrony, Neurofeedback, Neurofeedback Training, EEG, fMRI, fNIRS, MEG, Team Performance, Team Efficiency, Cooperation, Coordination, Team Coordination, Theta, Delta, Alpha, Gamma, Beta, SMR, Theory of Mind, Team, Task, Emotions, Hyperscanning, Similarity, Coherence, CCorr, PLV, PLI, Wavelet Transform Coherence ; (ii) an examination of the reference lists of the retrieved articles; (iii) the Thesis database from the University of Twente. From the collected database 125 articles were selected.

3 Recent work

The current section will discuss the most relevant recent literature in the field of INS. This literature serves as the building blocks for the current study, and the most important conclusions will be listed at the end of the section (see subsection 3.6., preliminary conclusions).

3.1 INS during interpersonal coordination tasks

Multiple recent studies have suggested that there is a correlation between cooperative coordination tasks and INS. The pioneering work was by Cui et al., who showed that cooperative coordination tasks induced an increased amount of INS in the superior frontal cortex (measured through fNIRS) when compared to a competitive task ([20]). This finding was replicated and expanded through multiple studies, showing a correlation between interpersonal coordination and INS in the dorsolateral prefrontal cortex and left TPJ [80], frontopolar cortex and dorsal medial prefrontal cortex [89], left PFC [120] and the interior and middle frontal gyrus of each hemisphere of the PFC. Studies also showed a mediating effect of the gender [17], relationship [89], self-other overlap [31], severity of autistic spectrum disorder symptoms [130], inducement of pain [120], task difficulty level [84] and shared feelings of happiness [64]. The correlation between INS and cooperative coordination tasks was also found in different tasks in which coordination was complementary ([63, 33, 77]) and without the physical presence of the partner [124].

3.2 Requirements for INS during interpersonal coordination tasks

At this time, it can only be speculated why INS occurs during cooperative interpersonal coordination tasks. It has been demonstrated multiple times that competitive or individual coordination tasks increase INS to a much lesser extent than cooperative coordination tasks [73]. It has also been demonstrated that shared intentionality alone can not predict INS - at least not for all related brain regions, and if it is defined as the combination of joint action and a shared goal. In a study by Wang et al. (2020) two groups of children with different severities of autistic spectrum disorder (ASD) were compared in terms of INS during an interpersonal coordination task with their parents [122]. The dyads with more severe symptoms of ASD showed lower INS in the superior frontal cortex during the task than children with less severe symptoms, although both groups shared actions and a common goal [122]. The researchers also compared INS at the left lateral PFC, in which no difference was found between the different groups.

Another important aspect is that physical joint action is not necessary for INS to occur - at least no greater action than visual inspection [109]. Indeed, Szymanski et al. demonstrated that joint visual inspection during a cooperative searching task also can increase PLI-INS between two team members. In this experiment, the goal of the participants was to count the total number of objects in an image as quickly and accurately as possible, both individually and with one partner. The results suggest that both the response time (i.e. how quickly dyads selected the number of objects) as the response accuracy (i.e. whether the number of objects is correct or not) were found to be negatively correlated with PLI. These results suggest that while INS may not be related to all forms of team performance, it may be related to the temporal dimension - i.e., more effective time management in team coordination. In any case, this study also presents the suggestion that INS during cooperative coordination is not dependent on physical actions, as it can also be present during joint searching. Szymanski et al. argue that this joint searching is an example of joint attention without joint action - however, this depends on the definition of joint action, as searching can also be considered an “action”.

3.3 Related oscillations of Interpersonal Coordination

Previous studies that investigated INS through EEG in the context of cooperative coordination mainly report a correlation at the theta frequency. Other frequencies that have been suggested are alpha and delta ([7, 119]). Indeed, according to many reports, slow oscillation INS is considered to play a larger role in interpersonal coordination than higher frequency INS [77].

An example of a study in which INS has been investigated during interpersonal coordination tasks is a study by [7]. During this study participants performed joint actions by simultaneously pressing two different keys when confronted with a Go stimulus and withholding action when confronted with other stimuli (see figure 3.3.1). From the EEG data it was found that delta and theta power-based COH were increased significantly between participants while alpha and beta were not.

Another study investigated effective alpha (8-13 Hz) and theta (3-7 Hz) connectivity during three phases of a dyadically performed flight simulation [112]. They found that INS increased for both alpha as theta during the parts of the simulation that demanded high coordination (see figure 3.3.2.). Namely, “taking off” and “landing” required more coordination than “cruising”.

A study by Wang et al. (2020) elaborated on these findings, investigating both alpha and theta phase-COH during the cooperative and competitive green button pressing task [119]. While both alpha and theta-INS increased during both tasks, only theta-band COH increased significantly in the cooperation task when compared to the competitive task (see figure 3.3.3.). The authors argue that increased INS in the alpha-band can possibly occur due to the attentive load that the tasks

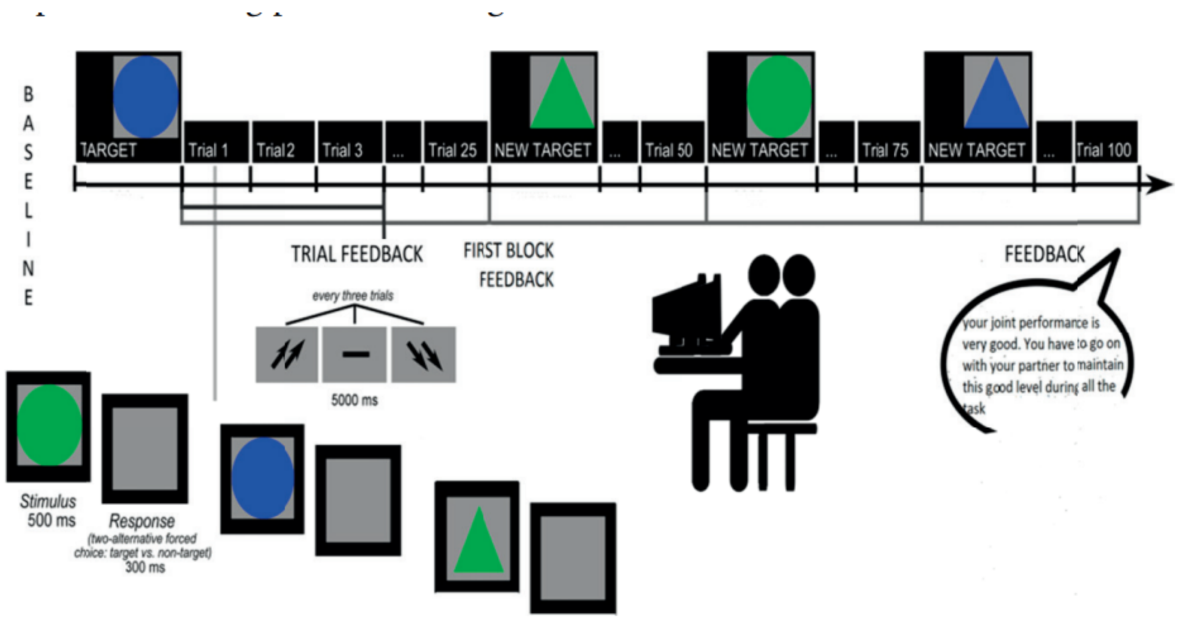


Figure 3.3.1: Experimental procedure of Balconi et al. (2018). Here, participants performed a joint action task in which the goal was to press two different keys simultanously after a "Go" stimulus, while withholding action when confronted with other stimuli [7].

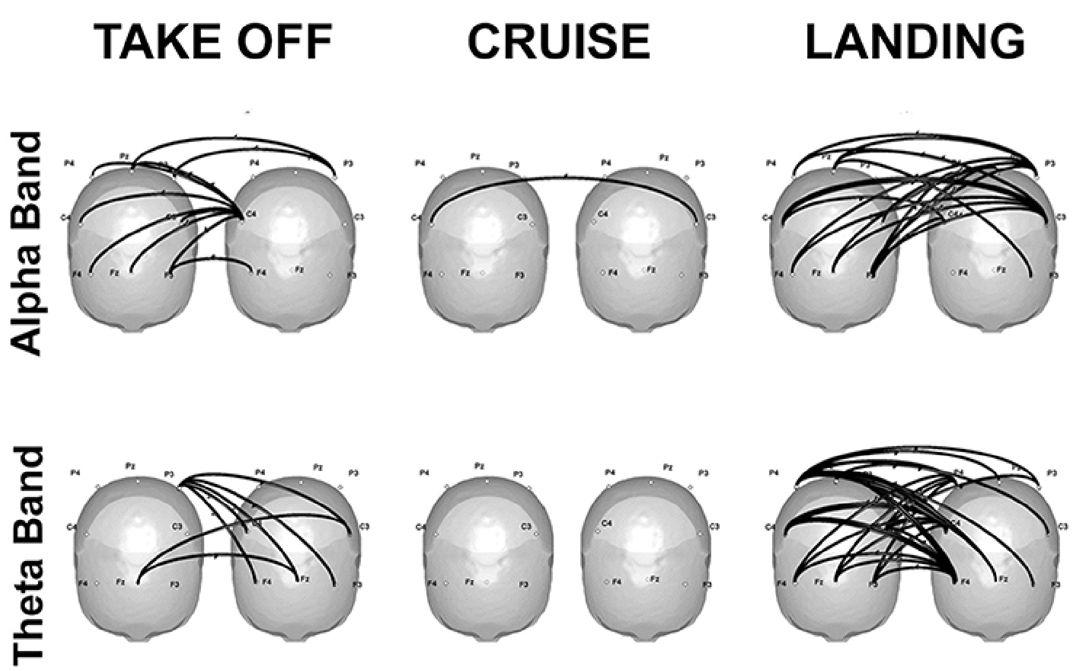


Figure 3.3.2: Statistically significant INS-connectivity patterns (seen as black arrows) between individuals during take-off, cruise and landing in the alpha (8-13 Hz) and theta (3-7 Hz) bands [112].

require, as alpha is known to be associated with attention. As both the competitive as well as the cooperational tasks require a similar attentional load in order to be "ready" for the Go-signal, alpha INS did not differ significantly between tasks.

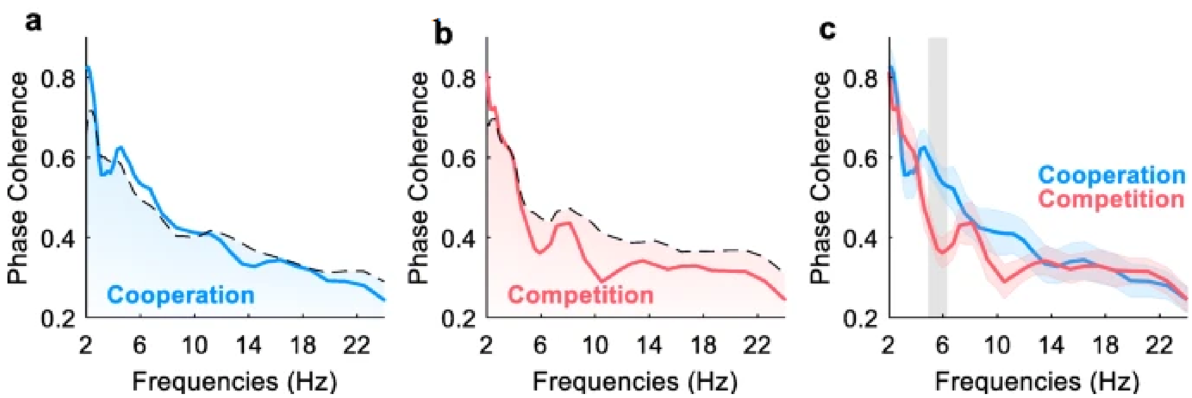


Figure 3.3.3: a.The mean Phase-COH during cooperation. The solid line represents the true value and the dashed line represents the cutoff of statistical significance (ps < 0.05, FDR corrected) b. The Phase-COH during competition. The solid line represents the true value and the dashed line represents the cutoff of statistical significance (ps < 0.05, FDR corrected) c. The Phase-COH during cooperation (blue) and competition (red) [119].

3.4 Related INS studies in Theory of Mind

Theory-of-Mind is a concept that is less straightforward than interpersonal coordination. One type of task that is assumed to be related to ToM is the Prisoner's Dilemma task, as in order to gain as many points as possible it is required of the participant to predict the actions of their opponent. In the Prisoner's Dilemma, participants must choose to either cooperate or to defect, but the amount of points gained depends on the choice of the partner. If both participants choose to cooperate, both gain approximately half of the points that can be obtained. If one chooses to defect while the partner cooperates, the person defecting gets the largest amount of points, while the partner gains nothing. If both choose to defect, both get a very small amount of points (nearly nothing). In a study by Hu et al. [45] EEG was measured during the Prisoner's dilemma task, and PLV was calculated for each trial. The authors also manipulated the likability of cooperation by making it less favourable to cooperate (altering the points given for cooperation). The results from the study showed a significant increase in theta PLV in the medial PFC and alpha PLV in the centro-parietal cortex for the condition in which it was more favourable to cooperate. Furthermore, PLV also covaried with cooperation choices, as it was increased during cooperation. These results suggest that a cooperative mindset is more related to INS than a more self-centered mindset.

In a different type of study, [8] investigated the effect of 'interpersonal tuning', i.e. the capacity to understand and mirror others' feelings, on INS during leader-employee performance reviews. In this study, the leader was asked to review the performance of their actual employee either through quantitative ratings or through words. As quantitative leader-employee ratings are known to have a negative effect on the employee's performance (arguably due to it being perceived as a threat) it was hypothesized that a description with words would lead to higher INS than a quantitative rating. The results showed higher delta- and theta-band COH over the bilateral frontopolar cortex for the qualitative approach than for the quantitative approach, which was also related to an increase of self-perceived emotional tuning, agreement on content, and interpersonal cooperation by both leader and employee.

It is also important to note that there have been an extensive line of studies investigating the correlation between mutual attention and INS [38]. Some researchers claim that INS may increase mutual attention, facilitating more attunement and greater allocation of attention to the significant interaction in order to increase its potential gains [38]. This claim has been supported by a number of studies, showing that INS has an influence on the level of a learner's attention to their instructor's vocal behavior and to the interaction [83] and the effectiveness of teaching [22].

3.5 Neurofeedback Training in INS research

To date, only a handful of studies have been conducted to research the trainability of INS through NFT, of which all-but-one used single session NFT in human trials and one performed multi-session training in pigeons. The latter study was conducted by Yang et al. (2020), who trained three groups of three pigeons to increase PLV-INS in the gamma-band [127]. The pigeons were trained to adjust their brain activity shortly after receiving a visual cue by a LED light by reaching the threshold value of synchronization (see figure 3.5.0.1.). The feedback was food if the threshold was reached, and no food when the threshold was not reached. The training lasted for approximately three months per group and the training was given to the pigeons every day for twenty minutes. The results suggest that PLV-INS can be significantly increased through NFT in pigeons. Interestingly, the training also enhanced intra-brain connectivity in the pigeons.

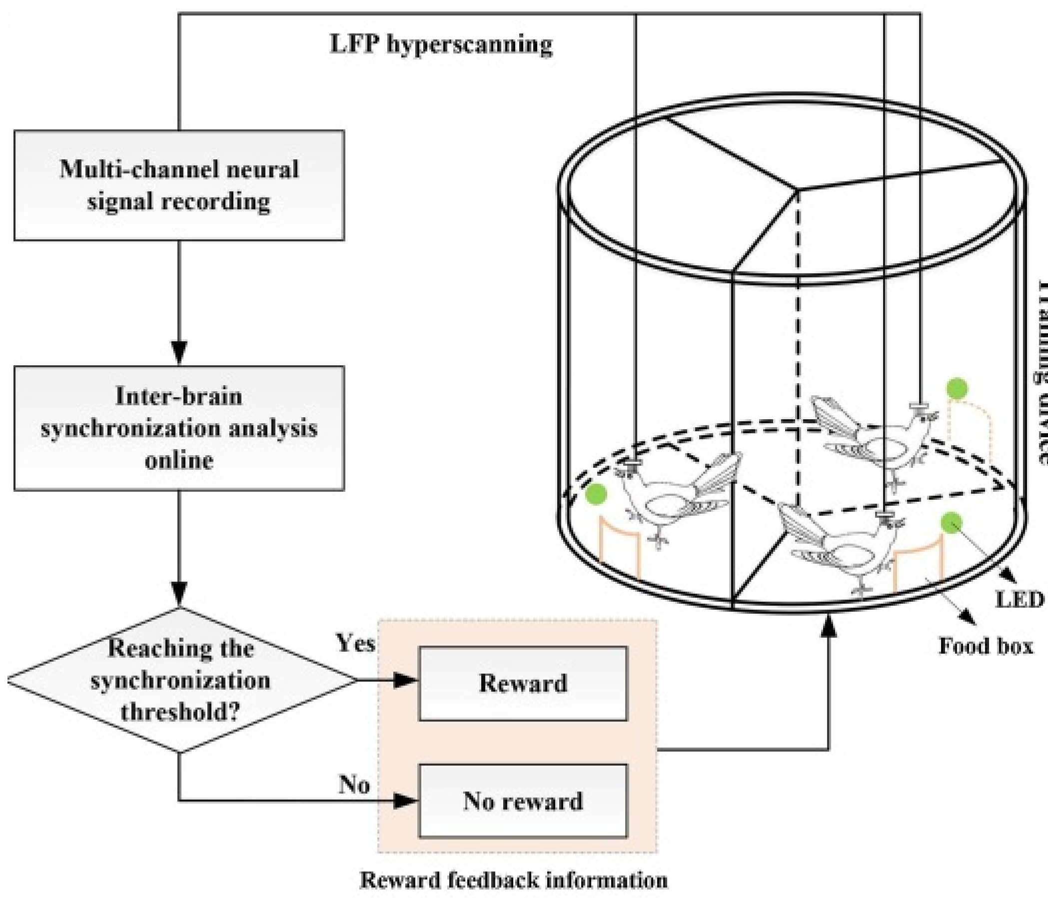


Figure 3.5.1: The experimental set-up used in [127]. Three pigeons had to increase PLV-INS above threshold to receive a food reward [127]..

Müller et al. (2021) was one of the first and most thorough studies that investigated the effects of single-session INS-NFT [77]. In this study, delta and theta INS were increased between two participants through two different tasks (see figure 3.5.0.2). INS was calculated through the absolute coupling index (ACI), which is a phase-based similarity measure. In the ball task the goal was to make the two circles (or balls) overlap, which could be achieved through an ACI value of above a certain threshold. The balls moved towards each other if the ACI value was increased. For the pendulum task, the goal was to make two pendulums swing in synchronization. Each of the pendulums represented the instantaneous phase angles of each participant.

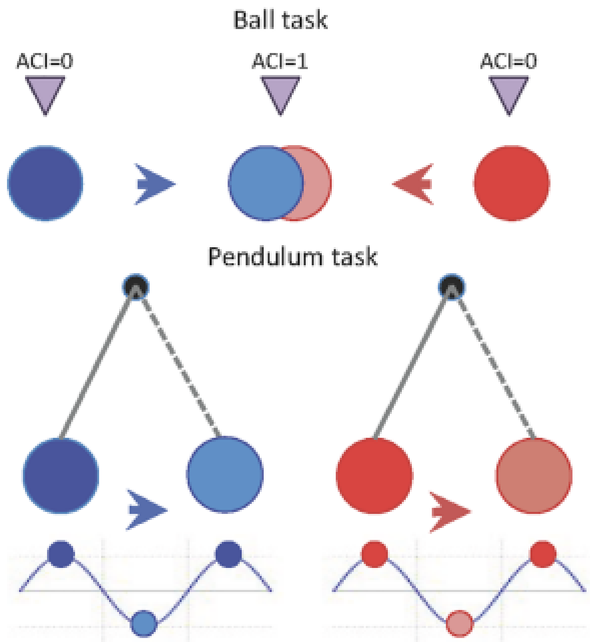


Figure 3.5.2: The ball task (upper image), dependent on ACI, and the pendulum task (lower image), dependent on the instantaneous phase angles of each participant [77]

Interestingly, there were three within-subject neurofeedback conditions used in this experiment. One was a 'default' neurofeedback experimental condition, for which the threshold for synchrony was not very easy to reach. A different neurofeedback condition was added with a much lower ACI-INS threshold called the 'fake' condition. In this condition, it was much easier for the dyads to reach the INS threshold and to remain in synchrony. The third condition was a type of control condition, where inverted feedback was given to the participants. In this condition - called the 'inverse' condition - the dyads received negative feedback when they reached the threshold for synchronicity and positive feedback when they were out of sync. The results for the three groups, two tasks and four frequencies (delta, theta, alpha and beta) can be seen in figure 3.5.0.3. The most important finding was that the mean delta and theta INS were significantly higher during the NFT trials than during baseline conditions. This was not the case for alpha and beta, which were not trained. Interestingly, there were no significant differences in delta and theta INS between the ball and pendulum task.

[15] recently made an application that can be utilized for INS-NFT, which uses a task similar to the ball task from [77]. This application can train the COH, CCorr, PLV, and three other measures (power correlation, imaginary coherence and envelope correlation). The goal of the task is to make two heads overlap, which move towards each other when the similarity measure between two sets of electrodes is increased. While this application has been validated through the Mutual Wave Machine, the task has not been used before in a NFT study.

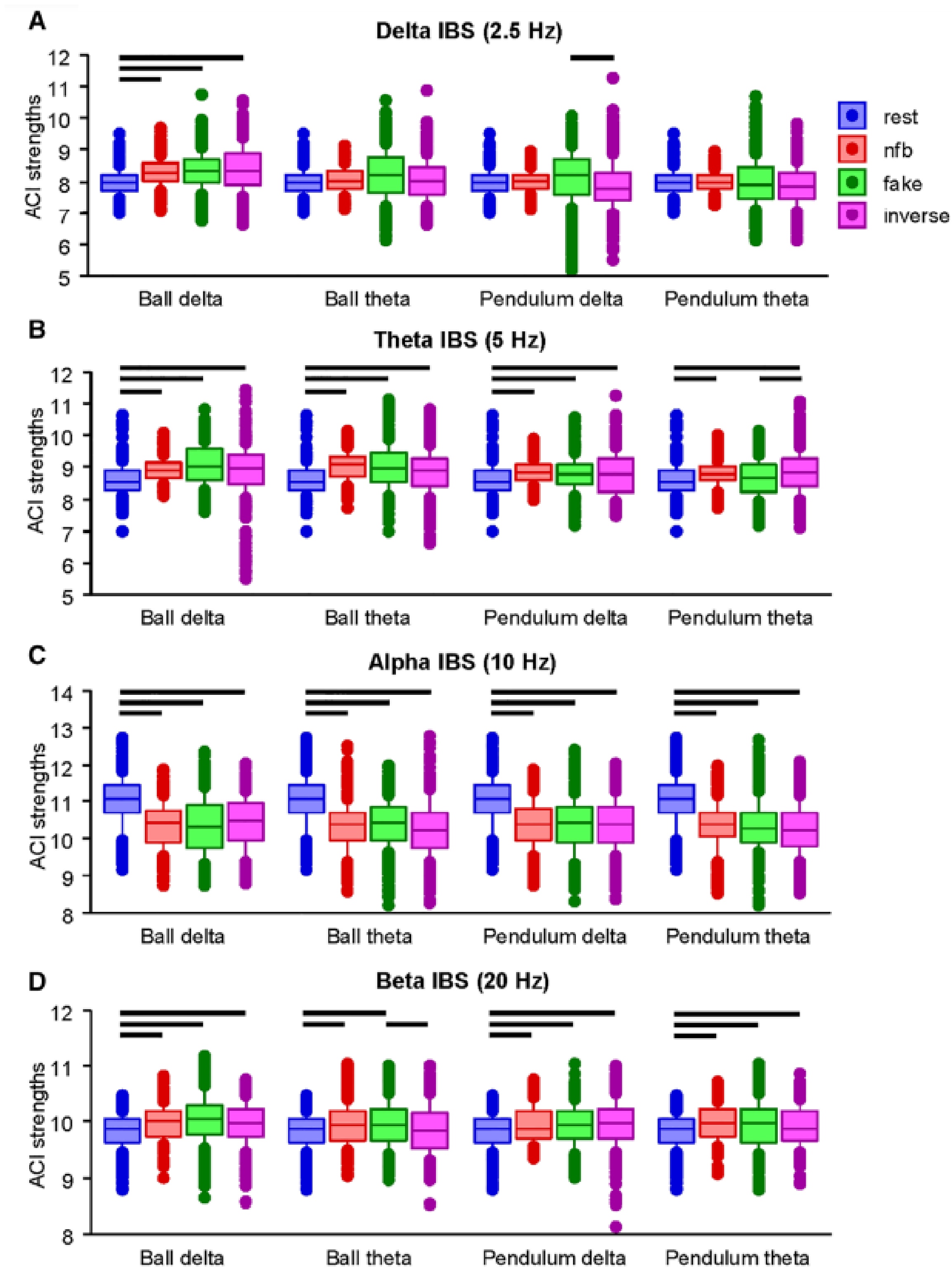


Figure 3.5.3: Mean results from the Neurofeedback group (nfb), 'fake' group (fake) and the 'inverse' feedback group (inverse) compared to baseline (rest). The y-axis represents the strength of ACI, a phase-based INS value. The results are listed separately for both tasks (ball and pendulum) and for both frequency bands (delta and theta) [77].

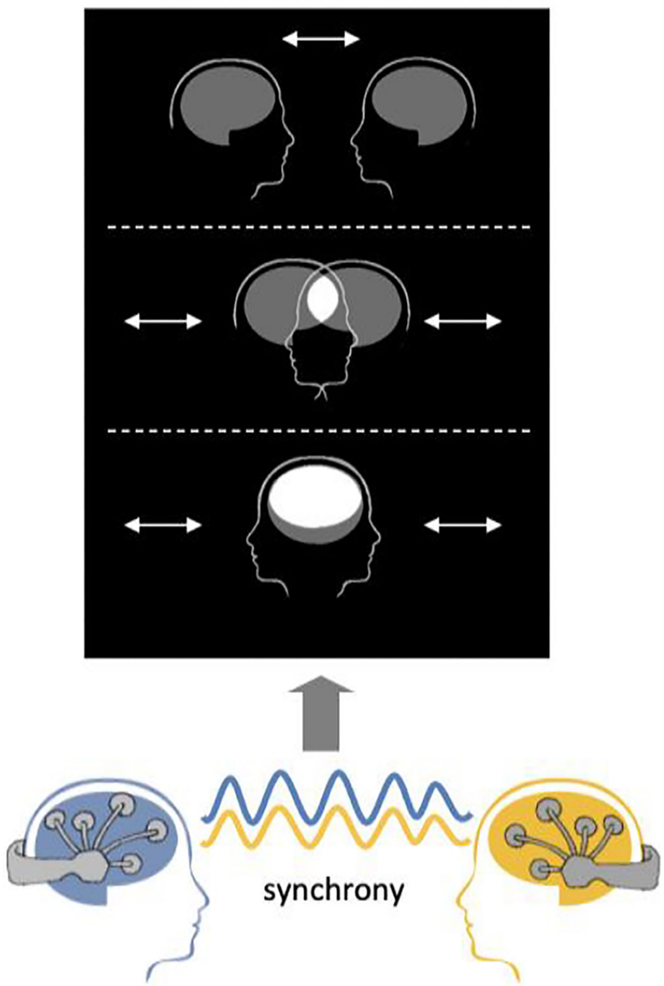


Figure 3.5.4: The task ulitised in [15]. The proximity of both heads represents a similarity measure (PLV, CCorr, COH, power correlation, imaginary coherence, envelope correlation), and a higher similarity would decrease the distance between the two heads [15]

3.6 Preliminary Conclusions

The current section has given an overview of the most relevant literature in the field. From this literature, several conclusions can be drawn which can be used to formulate hypotheses to help examine the research questions.

Firstly, it can be concluded from subsection 3.1. that there is some correlation between cooperative coordination tasks and INS, which can be detected in the superior frontal cortex, dorsolateral prefrontal cortex, left TPJ, frontopolar cortex and the left PFC [20, 80, 89, 120]. Secondly, it can be concluded that shared intentionality alone cannot predict INS (in the superior frontal cortex), and that physical joint action greater than visual inspection is not required for INS to occur [109]. Thirdly, Interpersonal Coordination is mostly related to INS in the theta frequency, but also to alpha and delta [7, 119]. Fourthly, ToM is also mostly related to theta-INS, with also large correlations for the delta-band and some results for alpha [45, 8]. Finally, there are results to support that INS can be trained through NFT, at least for the delta and theta frequency. There are also proper results in animal-trials for multi-session INS-NFT [127]. The following section will discuss the implications of these results for the current study more thoroughly.

4 Research Questions

In the previous section (section 3, Recent Work) multiple studies have been presented that demonstrated a positive relationship between INS and interpersonal coordination between team members. However, this relationship has yet not been researched thoroughly enough to explain whether a possible effect may be due to causation or correlation. The current study investigates whether there may exist a causal relationship between INS and interpersonal coordination, specifically that an increase of INS would entail an increase of interpersonal coordination (RQ2). This is investigated by increasing frontal delta and theta INS (as CCorr similarity) through three INS-NFT sessions, and measuring interpersonal coordination before and after training. Frontal theta and delta have been chosen as both of these frequencies in this location are implicated in interpersonal coordination as well as ToM when synchronized between two persons (see section 3, Recent Work) [20, 80, 89, 120]. Additionally, these frequencies have also been successfully trained before in a different neurofeedback study [77]. The adjusted CCorr similarity measure was chosen due to its insensitivity to coincidental phase while also being more resistant to the variability of EEG data than CCorr (see subsection 2.4.2., Similarity measures).

As NFT has not yet been established as a valid method of increasing INS, the current study also investigates the effectiveness of INS-NFT on increasing INS. For this purpose it is tested whether three sessions of INS-NFT have an effect on INS-CCorr (RQ1).

The following two research questions are formed:

> RQ1. Is frontal theta and delta INS-CCorr higher during a third session of INS-NFT than during the first session?
>
> RQ2. Is team coordination higher after three sessions of theta and delta frequency INS-NFT?

The proceeding sections of this study will examine these research questions further by presenting the performed experiment through the methodology, results, discussion and conclusion. Through these sections it is investigated whether there exists an empirical trend which may help answer RQ1 and RQ2. Due to a limited time frame and the exploratory nature of the study, the current research utilises a small sample size with little statistical power. However, the investigation of the research questions is also thorough, as multiple sub-hypotheses are tested through statistical analyses. Additionally, this research presents a foundation for further related research in the field to build and improve upon.

5 Methodology

The goal of the current study is to examine whether three INS-NFT sessions can increase frontal delta and theta INS-CCorr (RQ1) and whether this training would lead to an increase of interpersonal coordination between team members (RQ2). The methodology for the experiment is described in the current section, which is structured as follows. The current section starts with a subsection on the design of the experiment, which includes participant selection, the conditions and experimental set-up. The second subsection will discuss the procedure of the experiment, including the questionnaires, pre-experimental tasks, interpersonal coordination task and NFT task. The final subsection will discuss the measures, including data acquisition and statistical analysis.

5.1 Design

The design of the study is between as well as within subjects, which will be clarified in subsection 5.1.2, Conditions.

5.1.1 Participants

Eight participants between the ages of eighteen and sixty-five have been recruited through the University of Twente in exchange for credits and through acquaintance. These eight participants are divided into four teams of two participants (dyads). Out of the four dyads, two dyads are placed in the control condition and two in the experimental condition. All four dyads consist only of male participants as there is empirical data to suggest that dyads consisting of both a male and female could initiate significantly higher INS in frontal regions than same sex dyads [17]. Another reason is to minimise bias, as some studies report gender differences in INS [17, 65].

The exclusion criteria for participation are: (i) any severe mental disorder that could affect results due to atypical (social) cognition (Panic disorder, (severe) obsessive-compulsive disorder, major depressive disorder, bipolar disorder, dysthymia, Borderline Personality Disorder, Antisocial Personality Disorder, psychotic disorder, Post-Traumatic Stress Disorder), (ii) any severe neurological disorder that could affect results due to atypical (social) cognition (e.g. Dementia, Parkinson, Epilepsy, Alzheimer's disease), (iii) use of medication that could impair reaction time, (iiii) physical impairments that could impair reaction time and (iv) uncorrected impaired vision.

5.1.2 Conditions

The current study is conducted within as well as between subjects. The between subject experiment is conducted by forming two groups of equal size, one experimental group and one control group. Both groups of dyads receive mostly the same treatment, with the only difference between the groups being that dyads in the control group always receive random feedback during theta and delta INS-NFT training sessions. This is a valid methodology to use as a control condition for NFT and was chosen due to the minimal chance of influencing the results of the control group [105]. Using the control condition allows for the measuring of the difference between actual training and placebo. In this experiment the control condition is used for (I) investigating the effect of INS-NFT (actual and random) on INS-CCorr, and (II) investigating the effect of INS-NFT (actual and random) on interpersonal coordination scores.

The within-subject experiment is conducted by (i) comparing the INS-CCorr before and after three sessions of INS-NFT, and (ii) comparing the interpersonal coordination scores before and after three sessions of INS-NFT.

5.1.3 Experimental Setup

For each individual participant the experiment lasted for a total of four hours distributed over five days, which may have been ideally two or three days apart. Within a period of two months each participant received three training sessions of INS-NFT, two sessions of the team coordination task, one session of alpha NFT and one session of the single person coordination task (see figure 5.1.3.1. for a flowchart of the experiment overview).

The period in between the introductory session (i.e. *first* session) and the neurofeedback training and team coordination task may have been any time period between one day and two months, as it is not part of the INS-NFT training schedule.

The setup of the experiment is as follows. On the first day of the experiment each participant is screened for whether they are able to be trained through NFT, as not all people are (see 5.6.6.). This is done through a short alpha up-regulation NFT task, which is explained in section 5.2.2, Pre-experimental procedures. On the same day, the mean response time (RT) of participants is measured in order to form groups of participants with different mean RTs to ensure that interpersonal coordination is necessary to complete the interpersonal coordination task. After two participants are found that differ significantly in RT and are of the same sex, a team is formed. The first two teams that have been formed were placed in the experimental condition, and the last two teams in the control condition.

On the second day each dyad participated in an interpersonal coordination task and an INS-NFT task (more information on these tasks will follow in subsection 5.2., Procedures). On the thirth and fourth day only an INS-NFT session was performed, and the experiment concluded on the fifth day with an interpersonal coordination task. The reason for why three sessions of NFT were used to train each dyad is to train the dyads with as many sessions as possible while still being able to recruit participants, as finding participants who would participate in more sessions would not have been feasible for the current thesis. The current amount of sessions is expected to yield a result, as the experiment by Müller et al. did increase INS in the delta and theta band significantly in one session (in comparison to the baseline period and to other frequency bands) [77].

After each session the participants filled in a NASA Task Load Index questionnaire (see subsection 5.2.1., and were asked about their feelings about the training or coordination task session.

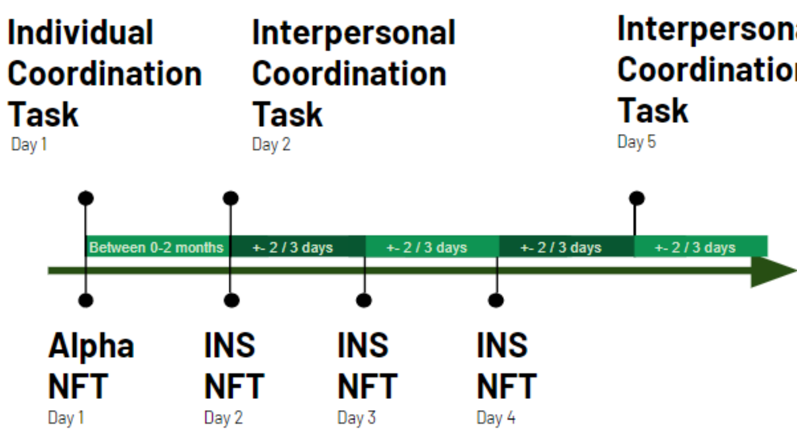


Figure 5.1.3.1: Experiment overview.

5.2 Procedures

5.2.1 Questionnaire

Before the participants started the experiment they were asked to give their consent through a physical consent form. Furthermore, they were asked about possible exclusion criteria and recent alcohol, drugs and caffeine consumption, which are known to have an effect on RT as well as learning [114].

After each NFT session a questionnaire was given to each individual participant to assess mental workload during the training (the NASA Task Load Index, see Appendix B). In this questionnaire the participant is asked to assess the mental, physical and temporal demand, performance, effort and frustration that has been experienced during the task [40]. This is asked in order to determine whether participants experienced learned helplessness, as this can occur during NFT and can have a substantial effect on the learning process (see subsection 5.2.1, questionnaire).

5.2.2 Pre-experimental procedures (Day 1)

As discussed in section 2.7.5., some participants do not respond to NFT training. These participants are filtered out in the first session through a non-related alpha NFT task. Upregulation of the alpha frequency is one of the oldest NFT paradigms [82] (see subsection 2.7.3, NFT Feedback), and there are no known risks of alpha upregulation tasks for healthy participants. The alpha NFT task consists of a baseline period of 110 seconds and five NFT trials of 90 seconds each. This training schema is based on [30] and on the pilot trials, after which the training time was shortened. During the baseline period, the participants are instructed to relax with their eyes open. After the baseline period the participants receive further instructions about the alpha NFT on the screen, which explain that the goal of the task is to increase alpha amplitude and that this can be achieved by increasing the saturation of a red circle on the screen for as much and as long as possible (see figure 5.2.2.1). This methodology is in line with [30], which demonstrated that this methodology can be used to increase upper alpha, with the electrodes P3, Pz, P4, O1 and O2. During NFT, the participants must increase their alpha amplitude above a certain threshold, which is initially calculated according to the rules in appendix C (see also subsection 5.3.2, online alpha processing). These calculations were found to be the most effective during the pilot trials, as these thresholds were not too easy or too hard for the participants. After the threshold is set, it may increase by one percent if the participant's mean alpha amplitude is above the threshold, and decrease by two percent if the participant's mean alpha amplitude is below the threshold for two times in a row. The participant receives feedback about their alpha amplitude once for every two seconds. These values were also chosen based on the pilot trials and on previous studies. Previous studies provided alpha NFT feedback approximately once per .1 to 5 seconds with a threshold increase or decrease of .1 to one or two percent.

Furthermore, participants are also asked to perform a variation of the default interpersonal coordination task (see subsection 5.2.3.) that can be completed individually (as used in [119]). This task is conducted in order to measure the mean reaction time (RT) of each participant, so that participants with the greatest difference in RT can be matched together in a team (see figure 5.2.2.2). This ensures that interpersonal coordination is necessary to complete the interpersonal coordination task successfully (see 5.2.3.).

The goal of each participant is to press a key (spacebar) as soon as possible in response to a green circle appearing on the screen. If the response is within the time span of 1000 ms, the participant receives a point. The point system is used in order to motivate the participants to respond as quickly as possible. The green circle appears randomly between 0.6 and 1.5 seconds after the start of a trial, which starts by displaying a gray hollow circle.

5.2.3 Interpersonal Coordination task

Team coordination is assessed through the green button-pressing task (see Figure 5.2.3.1.) [20] which will be called the interpersonal coordination task. This task was chosen as proper task completion requires interpersonal coordination, and the task does not require physical proximity, verbal communication or non-verbal communication. The task is also simplistic and frequently used within INS research, as it is known to evoke INS measured by variety of measuring and similarity techniques (see section 3, Recent Work). The reason for why it requires interpersonal coordination can be seen in the definition for interpersonal coordination: "the ability to use strategies and behavioural patterns aimed at integrating and aligning the actions, knowledge, and objectives of interdependent actors, with a view to attaining common goals" [92]. Other authors argue that in

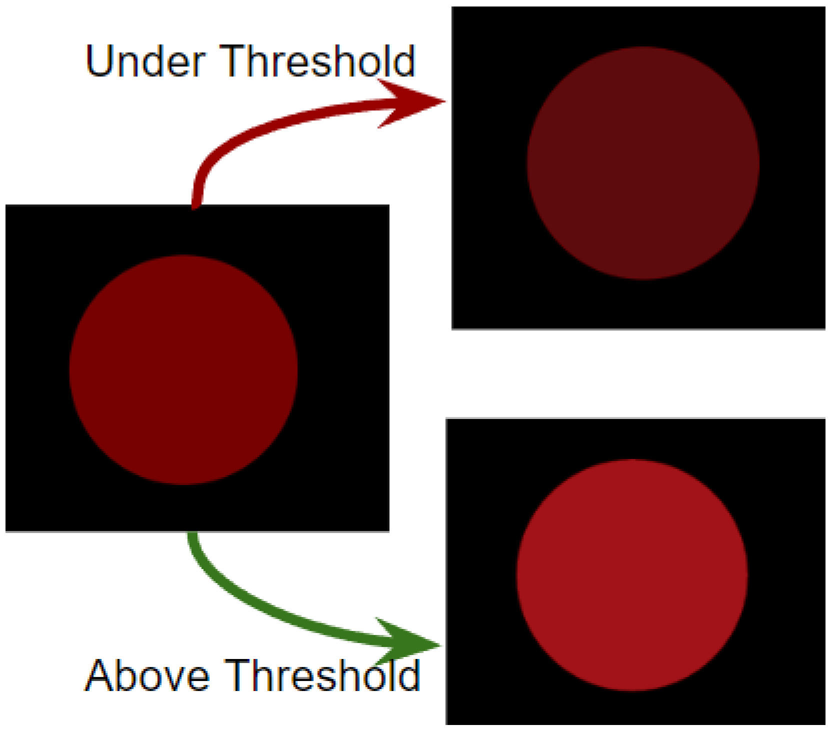


Figure 5.2.2.1: Alpha NFT task. If the mean alpha amplitude is above the threshold, the saturation of the circle increases. If the mean alpha amplitude is under the threshold, the saturation decreases.

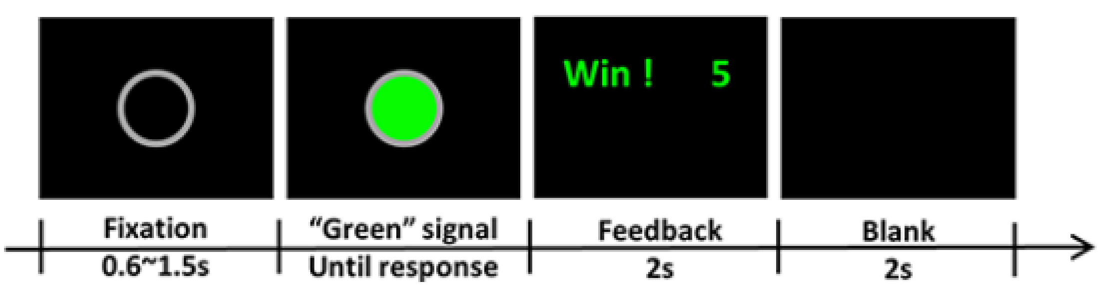


Figure 5.2.2.2: Individual Response Time Task

order to complete the task as well as possible it requires interpersonal coordination, joint action, shared emotion and mentalizing [130]. In order to make it absolutely necessary for participants to align their actions, participants with different mean reaction time (RT) are selected to work together as a team beforehand (see section 5.1.1. Participants). The task's simplicity reduces the chance of having interferences of other variables, and increases the chance of measuring the correct variable.

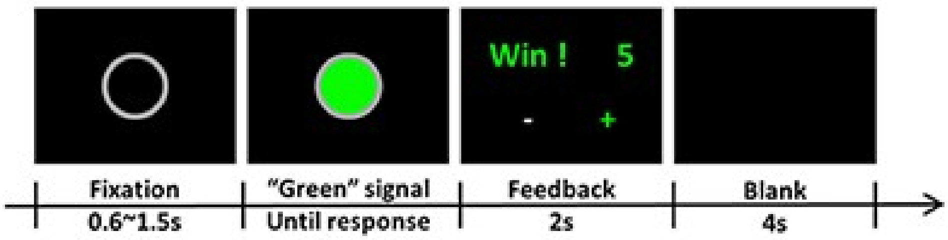


Figure 5.2.3.1: One trial of the interpersonal coordination task.

Each session of the interpersonal coordination task consists of three experimental blocks that are separated by 30 second resting periods (see figure 5.2.3.2.). Each experimental block consists of 20 trials, summing up to a total of 60 trials. The task always starts with a practice block consisting of four trials.

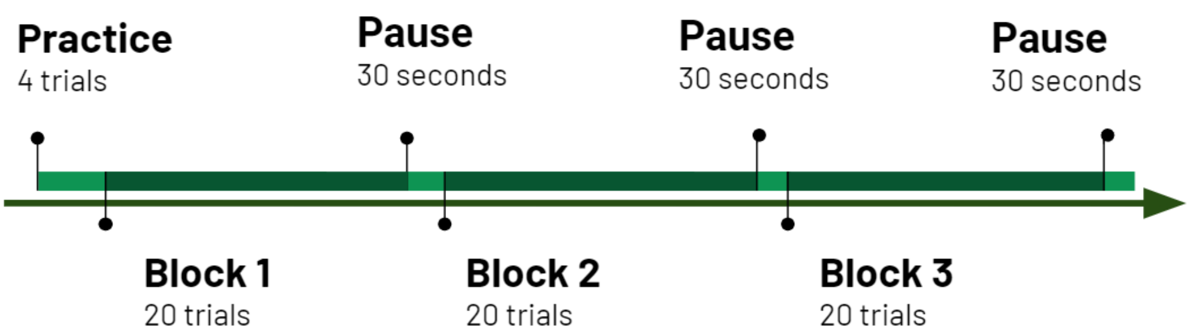


Figure 5.2.3.2: One session of the interpersonal coordination task.

The procedure for the coordination task is as follows. First, the dyads are seated face-to-face in front of two separate computer screens. Sitting face-to-face ensures the participants that they perform the task together, while separate computer screens prevent the participants from imitation [120]. The participants are not allowed to communicate with each other during the task either verbally or non-verbally (through signing) in order to minimise the chance of external interpersonal coordination to occur.

After being seated each participant receives a participant number (either 1 or 2) which they maintain for the duration of the experiment. After this the task starts with a practice block, followed by two experimental blocks (see figure 5.2.3.2.). Each trial starts with a hollow gray circle in the center of the computer screen, which remains on the screen for an unpredictable time interval between 0.6 and 1.5 seconds. After this period of time the circle changes its colour to green, to which the dyads have to respond by pressing a button on their keyboards as simultaneously as possible. One of the participants is instructed to press the 'o' key, while the other participant is instructed to press the 'a' key. After both team members pressed a key feedback is shown for a duration of four seconds, which consists of the word "win!" or "lost", the cumulative points gathered by the team and the participant number of the quickest responder (e.g. "p1 was quicker"). If the difference between the two team member's response times is smaller than one threshold value, both of them would get one point; otherwise, each of them would lose one point. The threshold was calculated as 1/8 of the averaged response time of the two participants of a given trial. The parameter 1/8 was chosen as it is known as a moderate level of difficulty for this task [64]. After the feedback, an inter-trial interval (black screen, 3.6-4.5 s) is shown, followed by the next trial.

5.2.4 Neurofeedback Training

5.2.4.1. Schedule

Each dyad participated in three sessions of INS-NFT. The duration of INS-NFT training was determined to ideally last two to three weeks with two to three days in between sessions. This training schedule was chosen as it is within the range of average number of sessions and time between sessions (see subsection 2.5.4), and it was also feasible in order to recruit a sufficient amount of participants. Each session starts with a baseline measuring, followed by a block of trials, resting phase and a second block (see figure 5.2.4.1.). During the first block theta (4-8 Hz) is trained for twelve minutes, and the second block consists of delta (0.5-4 Hz) training of the same time frame. Each of these blocks consist of eight trials of 1.5 minutes each. The duration of the resting phase is thirty seconds and baseline measuring was performed for two minutes. This training schedule is similar to the training used in Müller et al. (2021), but there are a few differences. The first difference is that the trials are 30 seconds shorter than in Müller et al., which is solely for the purpose of having a round number. Secondly, there are fewer trials than in Müller et al., which is due to this study not including multiple within-participant NFT conditions and only having one type of training task. Müller et al. (2021) had two types of NFT training tasks, and three experimental conditions per participant.

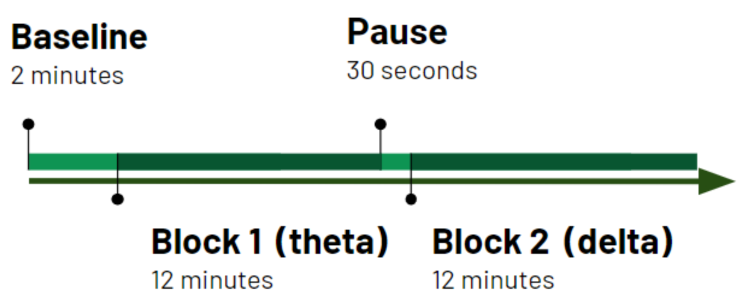


Figure 5.2.4.1: Experimental design of a single session. Block 1 is of Theta training, and block 2 for delta training.

5.2.4.2. Physical Instructions

For each INS-NFT session the dyads were first asked to sit in front of a computer screen, which were placed approximately 2 meters apart and on opposite sides of a table see (figure 5.4.2.2.). This spacing was used in order to prevent the team members from communicating with each other non-verbally, which is known to affect INS [128]. The dyads were instructed that they are performing the tasks together, and that the goal is to synchronize their brain activity. For the baseline measurement the dyads are instructed that they have to sit as still as possible for two minutes with their eyes open. One minute after the start of baseline measuring the dyads receive feedback that they have one minute of measuring left, and the same happens for the last thirty seconds (the message being "thirty seconds left").

5.2.4.3. Task Instructions

The task that is used for the NFT is based on the ball task from Müller et al., but visualized as the task in [15]. Before the start of the experiment, each individual receives instructions on paper

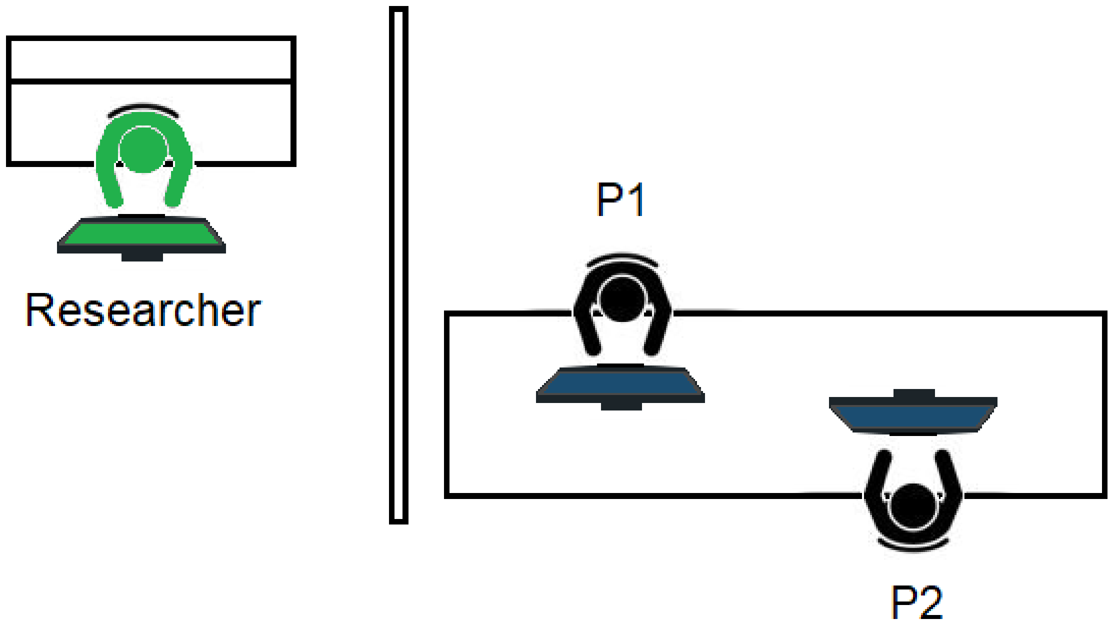


Figure 5.2.4.2: Physical locations of participants and researcher during the experiment

for how to modulate their brain activity within the task. The dyads are instructed that the goal of the task is to make two heads overlap for as long as possible, as for every second of overlap they receive one point. They are also instructed that the overlap of the two heads represents the synchronization of the two participants' brain activity.

The following instructions were physically given to each participant to keep next to their keyboard before and during the INS-NFT task:

Instructions
*The **red** head represents team member 1, and the **blue** head represents team member 2. The two heads move closer together when your brain signals are more synchronized, and move farther apart when your brain signals get less synchronized. When the heads overlap, your brain signals are in synchronization. Your task is to make the two heads overlap for as long as possible, as for every second that you and your team member are in synchronization you both will receive one point. The goal is to get the highest amount of points.*
While there is no single strategy to reach synchronization, relaxation and controlling your breathing may help.

These instructions are based on the instructions given in Müller et al. [77], which listed multiple strategies for increasing INS. These strategies are: relaxation, mental activation, generating thoughts and performing mental calculations. In Müller et al., the participants were asked to fill in a survey after the experiment to find possible correlations between task strategies and INS scores. They found that theta INS seems to be negatively correlated with the estimated capability to influence INS through concentration (for a similar task as used in the current study). Theta PSD was also negatively correlated with influencing the task through concentration or thoughts. Additionally, delta PSD was found to be positively correlated to the estimated capability to influence INS through relaxation.

Due to these results, it was decided to instruct the participants that relaxation or control of breathing may help to increase INS, as these strategies are most likely to increase theta or delta

PSD and/or INS in comparison to other strategies. Importantly, the participants were also told through physical and oral instructions that there is no single strategy for increasing INS, and that they should try out different strategies in order to see which one works best for them. Physical and oral instructions were given before each INS-NFT session, and the physical instructions (on an A4 format paper) were available to the participants during every INS-NFT session.

When the participants press a key to start a trial, an image can be seen of two heads. The goal of this task is to make the two heads overlap for as long as possible, and for each second of overlap the team receives one point. The heads can move closer to each other when the INS-CCorr value is equal to or above a given threshold, while the heads move further apart if INS-CCorr is below the threshold (see figure 5.2.4.2.). The INS-CCorr is calculated every three seconds, after which feedback is given to the dyads through movement of the heads (for details on the calculation of the threshold and CCorr value, see subsection 5.3.4., EEG Neurofeedback data processing (online)) .

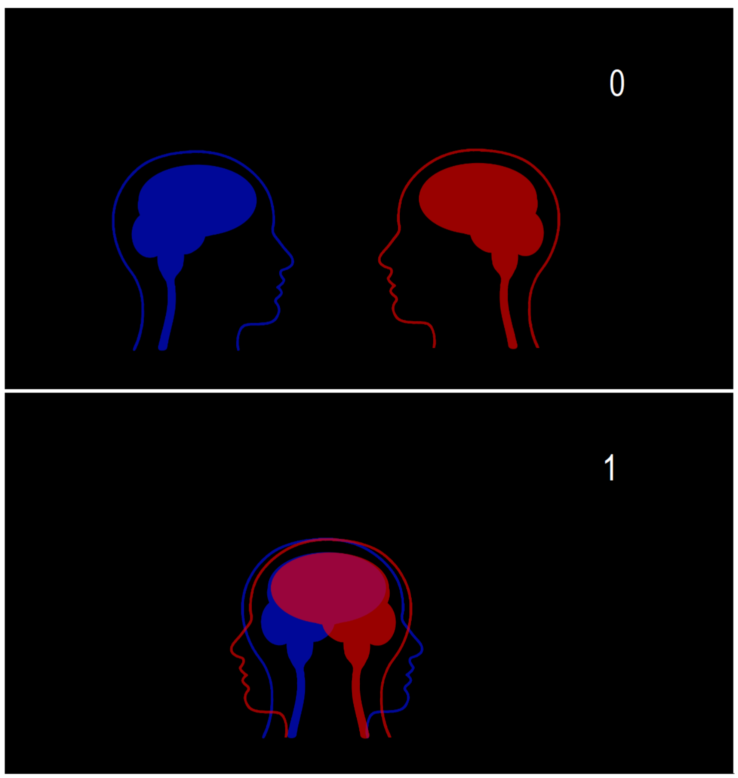


Figure 5.2.4.3: The experimental task. The proximity of both heads represents the CCorr value: higher CCorr entails more overlap of the two heads. Top: asynchronization. Bottom: synchronization (above threshold).

5.3 Measures

5.3.1 EEG data acquisition

At the start of each of the five sessions a 24-electrode ANT Neuro Waveguard EEG electrode cap is placed on each participant according to the 10/20 system (see subsection 2.5.1., EEG Electode Placement) measuring Fp1, Fp2, F9, F7, F3, Fz, F4, F8, F10, T9, T7, C3, C4, T8, T10, P9, P7, P3, Pz, P4, P8, P10, O1 and O2 with the ground electrodes Fpz and AFz. All twenty-four channels are measured during all tasks except for the questionnaires (i.e.: the interpersonal coordination task, INS-NFT, alpha-NFT and single-person coordination task). An eego™ amplifier EE-41x is used for amplification of the EEG signals at a sampling rate of 500 Hz and with impedance below 10kΩ.

During the Interpersonal Coordination task and the INS-NFT task the beginning of each trial is marked within the EEG recording. The time at which a marker is placed is exactly before feedback is given in both tasks.

5.3.2 EEG alpha data processing (online)

During the alpha neurofeedback training in the first session the EEG data from all twenty-four channels is recorded, but only channel P3, Pz, P4, O1 and O2 are used for online data processing during alpha neurofeedback training. These five channels have been selected for this task as it has been demonstrated before that these channels may be trained in alpha neurofeedback training [30].

First, the baseline upper-alpha activity is measured for 110 seconds. The baseline is measured as the mean absolute amplitude of 110 segments of one second which have been individually threshold-filtered at 300 μ V and band-pass filtered at 10-12 Hz. The value of the threshold-filter was determined through pilot trials. After the baseline amplitude has been calculated, an initial threshold is set for the neurofeedback training based on the baseline results (see Appendix C). The value is set in such a manner that the training is relatively easy at the beginning in order to minimize the frustration level.

For the actual neurofeedback training segments of 1.9 seconds are collected, threshold-filtered at 200 μ V and band-pass filtered at 10-12 Hz. If a value from a segment is greater than the threshold-filter or if the data stream is not available, no feedback is given to the participants. After filtering the segment is compared to the threshold value. The threshold value is increased by one percent if the mean segment amplitude exceeds or is equal to the threshold, and decreased by two percent if the mean segment amplitude is lower than the threshold for two segments in a row.

5.3.3 EEG alpha data processing (offline)

The data from twenty-four channels is band-pass filtered at 10-12 Hz and a Common Average Reference (CAR) is used. The data is split into two parts which represent the period before training (first 150 seconds of NFT) and after training (last 150 seconds of NFT). Both parts are converted to non-overlapping epochs of a length of two seconds (starting -.1 before label, ending 2.0s after label) and baseline corrected with the period between -.1s and 0s. The average baseline power spectral density (PSD) of the first part is then compared with the average PSD of the last part of the NFT task. It was determined through the pilot trials that an increase of twenty percent is likely to be considered a trained increase. Thus, participants who managed to increase alpha PSD for twenty percent were considered to be “responders” and were allowed to participate in the research.

5.3.4 EEG neurofeedback data processing (online)

During INS-NFT sessions the EEG data from all twenty-four channels is recorded, but for the generation of the feedback for online data analysis only a selection of EEG channels is used. The data that is used for the generation of feedback during the neurofeedback training task is from seven electrodes, namely: Fp1, Fp2, F3, F4, F7,F8 and Fz. These seven electrodes have been selected due to their relevance to ToM and Interpersonal Coordination [20, 80, 89, 120]. During the whole INS-NFT session the EEG signals are recorded and timestamped as well as annotated with markers directly after feedback is calculated. The amplified signals from both participants of a single team are retrieved separately, after which they are individually threshold-filtered at 200 μ and segmented into individual time segments of three seconds with an overlap of 25 percent. The value of the threshold-filters have been determined through pilot trials. If a value from a segment is greater than the threshold-filter or if the data stream is not available, no feedback is given to the participants. During the first block of each INS-NFT session the segments from both participants are FIR band-pass filtered at 4.0 Hz-8.0 Hz (theta), and during the last block the segments are band-pass filtered at 0.5 Hz-4.0 Hz (delta). After band-pass filtering, a Hilbert transform is used and the adjusted CCorr value is calculated with formula 4 (see section 2.6.2.4., Circular Correlation Coefficient) for each electrode pair (49 pairs in total). The CCorr value is averaged accross electrode pairs. Finally, feedback is given to the participants by comparing the averaged CCorr value to a threshold value. An overview of the pipeline used in INS-NFT online data processing can be viewed in figure 5.3.4.1.

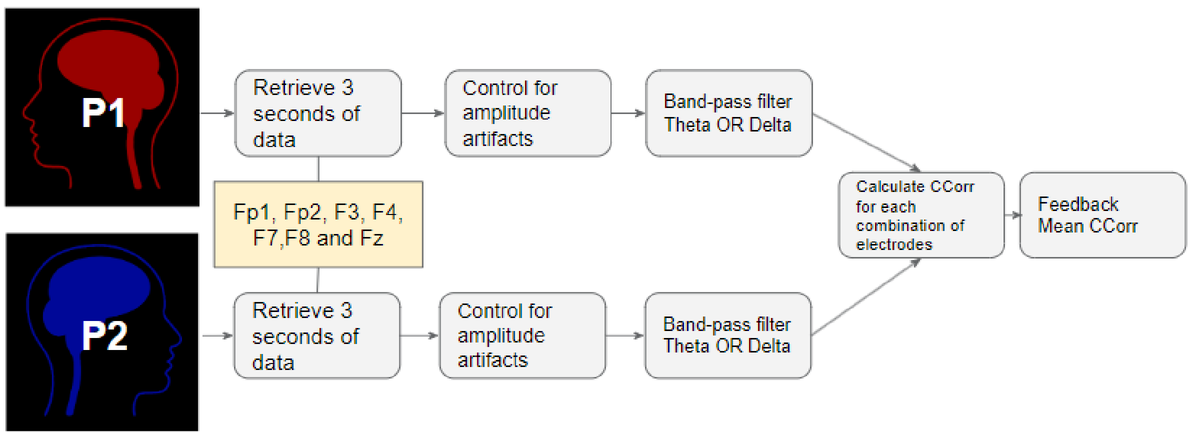


Figure 5.3.4.1: Pipeline for Neurofeedback processing. First, three seconds of data are retrieved. Then, the data gets controlled for artifacts as participants do not receive feedback if data contains artifacts. The data then gets band-pass filtered for theta (4-8Hz) or delta (0.5-4Hz), after which the CCorr is calculated and given as feedback to the participant.

The feedback that is given during INS-NFT sessions is determined by the value of the averaged adjusted CCorr measure and the value of the CCorr threshold at that moment. The threshold value is initially calculated as a percentage of the baseline value (see Appendix D for the used percentage values), which is obtained by measuring the average adjusted CCorr value during an eyes-open baseline period before each INS-NFT training session. This initial threshold value is designed to be relatively easy to reach at the beginning of each session in order to minimize feelings of frustration or learned helplessness. It was determined during pilot trials that participants may experience more frustration and learned helplesness if they fail more at the beginning of a session.

After the initial threshold is calculated the threshold changes at least once for every two instances of feedback. If the averaged CCorr value is equal to or above threshold value, the threshold increases by one percent; if CCorr is below threshold value for two times in a row, the threshold decreases by two percent. These values have been chosen as they are known to support learning (see subsection 2.5.3.), but also due to the effectiveness of this method as tested through the pilot trials.

5.3.5 EEG neurofeedback data processing (offline)

The raw EEG data from twenty-four channels is Band-pass filtered at 0.5 Hz - 45 Hz to correct for power line noise and scalp perspiration and referenced with CAR. Additionally, the data is epoched in non-overlapping chunks of 3.1 seconds (starting -.1 before label, ending 3.0s after label) for each trial which have been marked at the time of feedback during INS-NFT or interpersonal coordination. Epochs that are used for the purpose of amplitude analysis are baseline corrected with the period between -.1 and 0s. Any artifacts are manually removed through visual inspection. The artifact removal procedure can be seen in figure 5.3.5.1., which illustrates approximately the minimal artifact intensity around which eyeblink artifacts (A) and electrode artifacts (B) are removed. The left image (A) shows an epoch with an eyeblink artifact (epoch 76). Any epoch with this amount of deviation from the usual signal would have been removed. Epoch 77 showed less deviation from the signal, and would normally not have been removed. However, as it comes directly after an artifact, this epoch would also be removed. In the right image (B) we see a deviation from the usual signal for electrode F10. If an electrode showed this kind of deviation for too many epochs (approximately one quarter of the recording) it would have been removed entirely. Otherwise, the epochs would be removed as illustrated in 5.3.5.1. B, with the highest deviations (epochs 42,43,44 and 45) being removed and possibly related epochs (41) also being removed.

The methodology for the calculation of INS-CCorr from epochs is identical in both offline and online analysis.

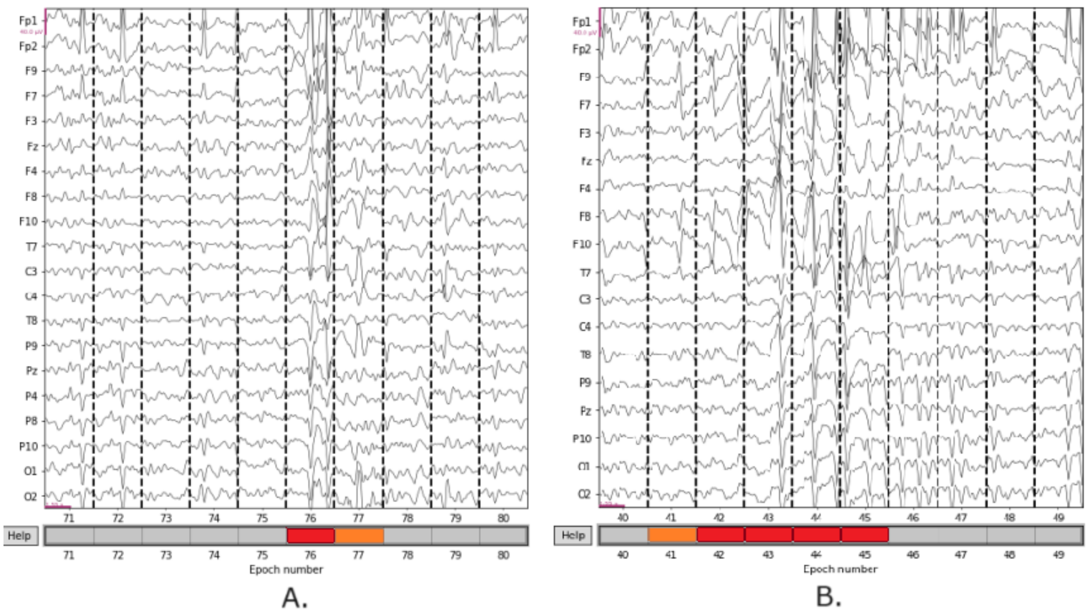


Figure 5.3.5.1: Artifact removal procedure for eyeblink (A) and electrode artifacts (B). Red color coded epochs (76A, 42B, 43B, 44B, 45B) are removed due to large deviation from the usual signal, orange color coded epochs (76A, 41B) are removed due to being related to an artifact. The illustrated artifacts are around the minimal artifact intensity for which removal would apply.

5.3.6 NASA Task Load Index

The NASA Task Load Index consists of six factors that unequally contribute to the overall workload during a task [39]. To calculate the NASA Task Load Index, each of the six factors receives a weight that corresponds to the importance of the factor for the workload during the task. The accumulated score for each factor is then obtained and multiplied by the corresponding weight.

5.4 Statistical Analysis

In order to investigate research questions 1 and 2 as thoroughly as possible, nine sub-hypotheses are formed that are tested through nine statistical analyses. RQ1 (Is frontal theta and delta INS-CCorr higher during a third session of INS-NFT than during the first session?) is split into four sub-hypotheses with one assumption sub-hypothesis, namely:

H1.a. *The Mean CCorr in the theta and delta band is significantly larger for Part 3 of Session 3 than for Part 1 Session 1.*

H1.b. The Mean CCorr in the theta and delta band is significantly larger for Part 3 of Session 3 than for Part 3 Session 1.

H1.c. The Mean CCorr in the theta and delta band is significantly higher for Session 3 Part 3 of the experimental group than in Session 3 Part 3 of the control group.

> *H1.c. assumption: The Mean CCorr in the theta and delta band is not significantly higher for Part 1 of Session 1 of the experimental group than Part 1 of Session 1 of the control group.*

H1.d.: There exists a correlation between CCorr and trial number for the experimental group.

H1.a. is investigated in order to test whether the INS-CCorr in the theta and delta band increases after three sessions of INS-NFT. Part one of session one represents the phase in which almost no neurofeedback training has been performed on the dyads. The opposite is true for phase three of session three, in which maximal neurofeedback training has been given to the dyads in this phase of the training schedule. During one training session of a single frequency the dyads perform the INS-NFT for twelve minutes. Here, "part one" represents the first four minutes of this session, "part two" the second four minutes of training, and "part three" represents the last four minutes.

This hypothesis can be tested through a paired, one sided t-test. Under the assumption that the two populations are normally distributed, a student's t-test can be utilized. Otherwise, a Welch's t-test can be used in unequal population variances but normal distribution, or a permutation test without such normal distribution.

H1.b. suggests that there is a difference in mean CCorr between the first and the third session. "Part three of session one" represents the last part (or last four minutes) of session one. During this phase the dyads have been training for one session already, and when comparing this phase to "part three of session three", the difference is that the dyads have had two sessions in between these two phases. This means that this hypothesis assumes that it is significantly more useful to train dyads for three sessions than for one session in order to increase INS-NFT in delta and theta. This hypothesis can be tested through the same statistical tests as H1.a.

H1.c. suggests that the experimental group has a higher INS-CCorr in the theta and delta band than the control group in the final phase (part three session three) of the training. During this phase, the two groups completed maximal INS-NFT, and this hypothesis suggests that the group which did not receive random feedback actually has a larger INS-CCorr in the final training phase. This hypothesis can be tested through an unpaired, one sided student's t-test, Welch's t-test or permutation test.

H1.c. can only be reliably tested if its assumption is tested that there was no such difference between the groups before INS-NFT (part one of session one). This assumption is therefore also tested through an unpaired one sided student's t-test, Welch's t-test or permutation test.

H1.d. assumes that there is a correlation between INS-CCorr and trial number for the experimental group. This would mean that an increase in trials actually significantly increases INS-CCorr. This can be tested through a Pearson correlation coefficient analysis, or if the data is non-parametric, the Spearman's rank correlation.

RQ2 (Is team coordination higher after three sessions of theta and delta frequency INS-NFT?) is also examined through multiple sub-hypotheses. The hypothesis is split into five sub-hypotheses and three assumptions:

H2.a: Interpersonal Coordination Winning rate is higher after three sessions of (experimental group) INS-NFT

> *H2.a. Assumption: Interpersonal Coordination Winning rate is not higher after three sessions of control INS-NFT*

H2.b: The difference in reaction speed is smaller after three sessions of (experimental group) INS-NFT

> *H2.b. Assumption: The difference in reaction speed is not smaller after three sessions of control INS-NFT*

H2.c: The winning rate after INS-NFT is higher in the experimental group than in the control group

> *H2.c. Assumption: The winning rate before INS-NFT is not higher in the experimental group than in the control group*

H2.d: The difference in reaction speed after INS-NFT is smaller for the experimental group than for the control group.

> *H2.d. Assumption: The difference in reaction speed before INS-NFT is not smaller for the experimental group than for the control group.*

H2.e: Winning an interpersonal coordination trial is associated with significantly higher mean delta and theta CCorr beforehand

For H2.a, it will be investigated whether the interpersonal coordination winning rate increases after three sessions of (experimental group) INS-NFT. Here, we speak of winning rate as the chance of "winning" a trial instead of "losing" by responding too differently from each other with respect to time, or before the circle turns green (see section 5.2.3., Interpersonal Coordination Task). The statistical analysis to test this hypothesis should be a one-sided Exact McNemar Test, as the data is paired and of small sample size.

There exists an assumption for H2.a, namely that the increase of interpersonal coordination winning rate is not due to any other variable than INS-NFT. This assumption can be tested by testing whether the interpersonal coordination winning rate is also increased significantly in the control group. This can be tested through a one-sided Exact McNemar Test.

In order to more thoroughly investigate the effect of INS-NFT on interpersonal coordination between team members H2.b will also be examined. The Reaction Time Difference (RTD) of team members is a numerical representation of interpersonal coordination, and it is hypothesized that the RTD would be significantly smaller after three sessions of INS-NFT. This is tested through a one-sided and paired student's t-test, Welch's t-test, or a paired permutation test if the data is not normally distributed.

The assumption for H2.b is that for the control group the RTD is not significantly higher after INS-NFT than before. If this assumption would be incorrect, the effect is not from the feedback from INS-NFT. Therefore the assumption will be examined through the same tests as H2.b.

H2.c. suggests that the experimental group has a higher interpersonal coordination winning rate after INS-NFT than the control group. This hypothesis therefore suggests that INS-NFT has an effect on interpersonal coordination winning rate, while random feedback does not. This can be tested through a one-sided Chi-Square test of Independence, or a one-sided Fisher's Exact Test for Count Data if the sample number is too small for a Chi-Square test.

The underlying assumption for H2.c is that the interpersonal coordination winning rate before INS-NFT is not larger for the experimental group than for the control group. If this assumption would be incorrect, this could mean that there was no effect from INS-NFT. This can be tested through the same statistical tests as H2.c.

H2.c suggests that the reaction time difference (RTD) is smaller for the experimental group than for the control group after INS-NFT. The difference in reaction time is a more accurate representation of the interpersonal coordination between dyads, while the winning rate more accurately represents team performance. As the RTD between team members is a numerical variable, this hypothesis can be tested through an unpaired one sided student's t-test, Welch's t-test or permutation test.

The assumption behind H2.c is that the RTD is not smaller for the experimental group than for the control group before INS-NFT. If this assumption would be incorrect, this could mean that there was no effect from INS-NFT. This hypothesis can be tested with the same statistical tests as H2.c.

Finally, it is investigated whether H2.e. is true, so whether INS is significantly higher before a dyad wins a trial. This can be tested by conducting an unpaired, one-sided student's t-test, or Welch's t-test if the data is not normally distributed.

6 Results

The results are presented in the current section in the form of visualizations (section 6.1), statistical analyses (section 6.2) and the NASA Task Load Index results (section 6.3).

6.1 Visualizations

The current section presents visualizations of the results from the statistical analyses. This section is segmented into four subsections, starting with the visual results of the statistical analyses regarding research question 1 (6.1.1.) and research question 2 (6.1.2.), visualizations of trained channel couples (6.1.3.) and visualizations of topographic maps (6.1.4.).

6.1.1 Research Question 1

The results for research question 1 (RQ1: Is frontal theta and delta INS-CCorr higher during a third session of INS-NFT than during the first session?) are visualized in the current subsection.

The main results can be viewed in figure 6.1.1.1, which illustrates the CCorr values (y-axis) for each of the two trained frequencies (left graph: theta, right graph: delta) in relation to the session number (x-axis) and experimental condition (blue: control, red: experimental). The visualized data is from the last part of each individual INS-NFT session (i.e., part three), which is the part in which the teams of participants have already been training INS for at least two thirds of a given session. It can be seen from visual inspection that there seems to be some growth in delta CCorr for the experimental condition between session 1 and session 2, while little difference can be seen for theta CCorr. The results for the second sub-hypothesis (*H1b: The Mean CCorr in the theta and delta band is significantly larger for Part 3 of Session 3 than for Part 3 Session 1)* and for the third sub-hypothesis *(H1c: The Mean CCorr in the theta and delta band is significantly higher for Session 3 Part 3 of the experimental group than in Session 3 Part 3 of the control group)* can be viewed in figure 6.1.1.1.

Additionally, the results for sub-hypothesis H1a *(The Mean CCorr in the theta and delta band is significantly larger for Part 3 of Session 3 than for Part 1 Session 1)* and the assumption of sub-hypothesis H1c (*The Mean CCorr in the theta and delta band is not significantly higher for Part 1 of Session 1 of the experimental group than Part 1 of Session 1 of the control group)* can be seen in figure 6.1.1.2. These graphs show the theta (left graph) and delta (right graph) CCorr value (y-axis) for the learning phase before any INS-NFT training occurred (part 1 of session 1) and after three sessions of training (part 3 of session 3), which can be found on the x-axis, for both the control (red) and experimental (blue) condition.

From a visual inspection, there seems to be some increase in CCorr value for both the theta and delta band in the experimental group (H1a). There appears to be a slight increase in delta CCorr and a slight decline in theta CCorr within the control condition. Moreover, there appears to be a slight difference between the initial CCorr values of the control condition and the experimental condition for both theta and delta (assumption H1c).

6.1.2 Research Question 2

The results for research question 2 (H2: The mean team coordination score is higher after three sessions of theta and delta frequency INS-NFT than before training) are visualized in the current subsection.

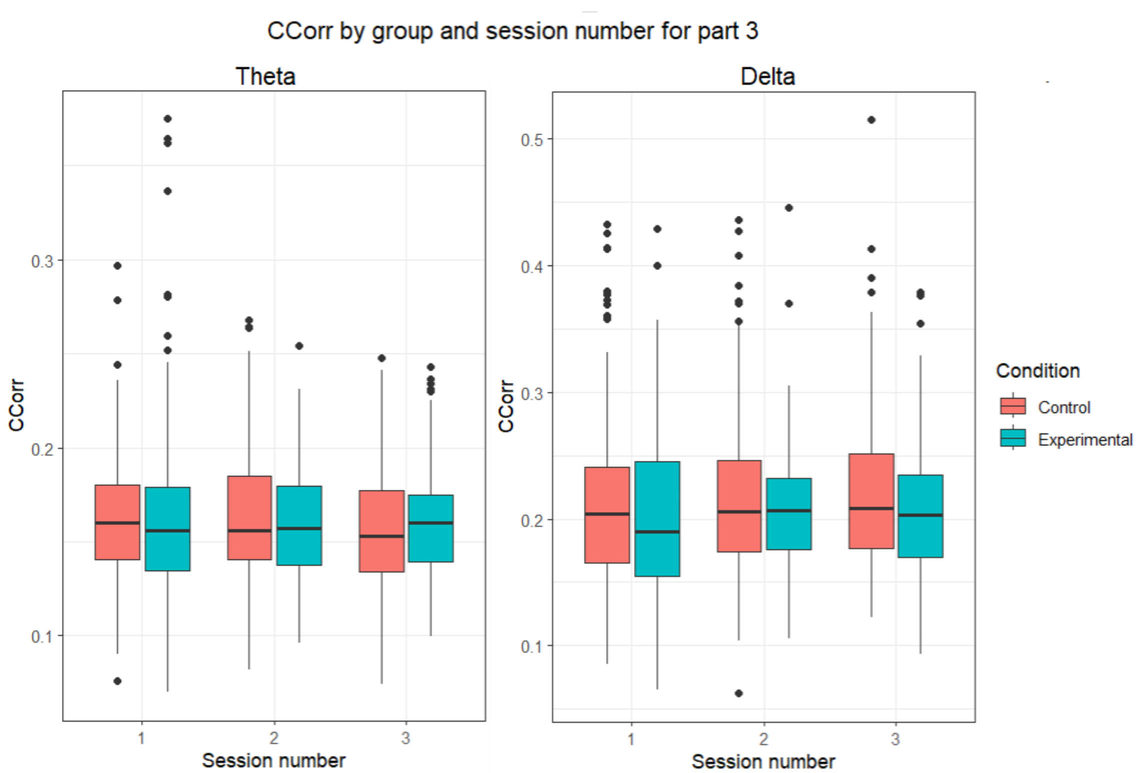


Figure 6.1.1.1: The results for part 3 of theta (left graph) and delta (right graph) INS-NFT as the CCorr value (y-axis) for each of the three given sessions (x-axis) and the two experimental conditions (red: control, blue: experimental). The session numbers are from left to right: session 1, session 2, and session 3.

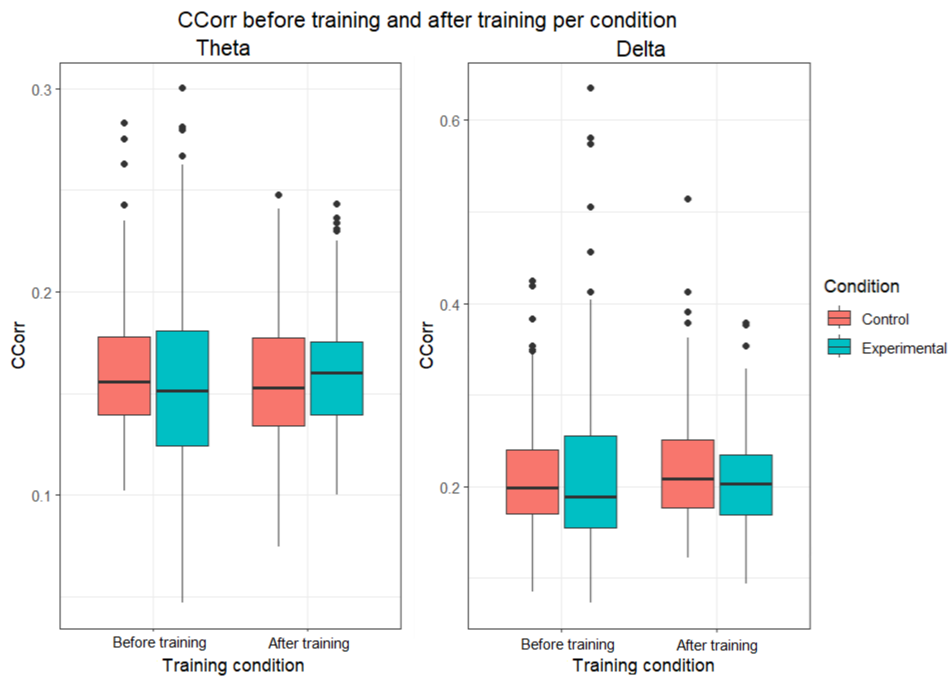


Figure 6.1.1.2: left: Theta (left graph) and delta (right graph) CCorr values (y-axis) before training (session 1 part 1) and after training (session 3 part 3) on the x-axis, per experimental condition (red: control, blue: experimental).

Firstly, the results of the team coordination task in terms of points per trial can be viewed in figure 6.1.2.1, which is relevant for H2a (H2a: *Interpersonal Coordination Winning rate is higher after three sessions of (experimental group) INS-NFT)* and the H2a assumption *(Interpersonal Coordination Winning rate is not higher after three sessions of control INS-NFT)*. Figure 6.1.2.1. shows the cumulative points on the vertical axis, the trial (each time a team may press a key simultaneously) on the horizontal axis and the results of each individual team before (blue) and after (red) training as stars (left: team A; right: team C) or circles (left: team B; right: team D). The blue and red lines represent the regression model for both conditions before and after training respectively. The regression models for the experimental group are:

$$Score\ (before\ training) = Trial * -2.74 * 10^{-1} - 1.83 \tag{8}$$

$$Score\ (after\ training) = Trial * 1.78 * 10^{-1} - 9.20 * 10^{-1} \tag{9}$$

With the R^2 before training (formula 8) being $R^2 = 2.39 * 10^{-1}$, and after training (formula 9) $R^2 = 4.25 * 10^{-2}$, which means that much of the variance cannot be explained by the regression models.

And for the control group:

$$Score\ (before\ training) = Trial * 3.82 * 10^{-1} + 3.09 \tag{10}$$

$$Score\ (after\ training) = Trial * -3.76 * 10^{-1} - 2.231 \tag{11}$$

With the R^2 before training (formula 10) being $R^2 = 8.34 * 10^{-1}$, and after training (formula 11) $R^2 = 9.37 * 10^{-1}$. This means that much of the variance can be explained by the models.

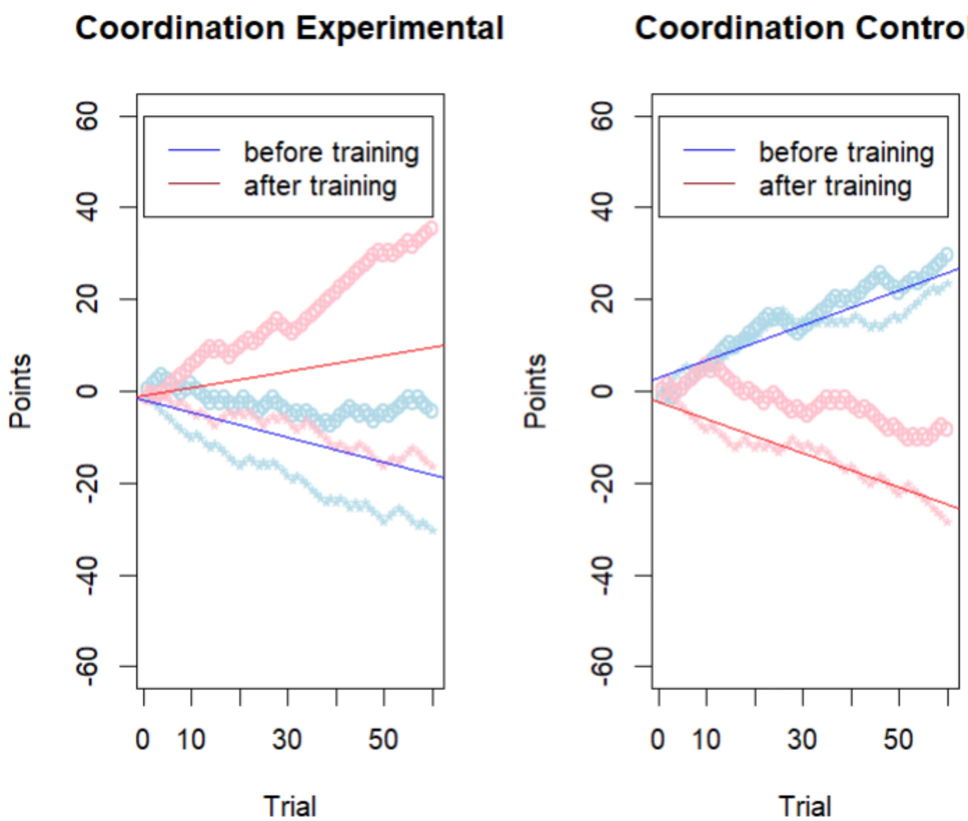


Figure 6.1.2.1: left: Team coordination scores for the experimental group, with the points gained on the y-axis and trials on the x-axis. Right: Team coordination scores for the control group, with the points gained on the y-axis and trials on the x-axis.

The results for the reaction time difference (RTD) per trial during the interpersonal coordination task can be viewed in figure 6.1.2.2., which are relevant for H2b (*The difference in reaction speed is smaller after three sessions of (experimental group) INS-NFT)* and the assumption of H2b (*The difference in reaction speed is not smaller after three sessions of (control group) INS-NFT)*. The RTD is here shown on the y-axis and the trials on the x-axis for individual teams as stars (left: team A; right: team C) or circles (left: team B; right: team D) in both conditions (left: experimental; right: control) before (coloured blue) and after (coloured red) INS-NFT. The blue and red lines represent the regression model for both conditions before and after training respectively. The regression models are as following for the experimental group:

$$RTD\ (before\ training) = Trial * -1.66 * 10^{-1} + 77.6 \tag{12}$$

$$RTD\ (after\ training) = Trial * -1.81 * 10^{-1} + 67.1 \tag{13}$$

With the R^2 before training (formula 12) being $R^2 = 2.12 * 10^{-3}$, and after training (formula 13) $R^2 = 1.66 * 10^{-3}$. The amount of variance that can be explained by the regression models is very low.

And as following for the control group:

$$RTD\ (before\ training) = Trial * 6.83e * 10^{-2} + 24.0 \tag{14}$$

$$RTD\ (after\ training) = Trial * .190 + 36.82 \tag{15}$$

With the R^2 before training (formula 14) being $R^2 = 2.32 * 10^{-3}$, and after training (formula 15) $R^2 = 8.94 * 10^{-3}$. The amount of variance that can be explained by the regression model is very low.

For H2d (H2d: *The RTD after INS-NFT is smaller in the experimental group than in the control group)* the RTD is compared between the experimental group and the control group, of which the results can be viewed in figure 6.1.2.3. Figure 6.1.2.3. illustrates the results in terms of RTD (on the y-axis) for both conditions (experimental and control) for the interpersonal coordination task before (session 1) and after (session 2) INS-NFT on the left and right respectively. The results from the experimental group are illustrated in the blue coloured box plot labelled group "1", and the results from the control group are illustrated in the red coloured box plot labelled group "0".

Finally, the delta and theta CCorr is compared for losing and winning an interpersonal coordination trial *(*for *H2e: Winning an interpersonal coordination trial is associated with significantly higher mean delta and theta CCorr beforehand)*, of which the results can be viewed in figure 6.1.2.4. Figure 6.1.2.4. shows the CCorr on the y-axis and the group on the x-axis, with the groups being the lost trials (as the red box plot) and the won trials (as the green box plot). The CCorr is in the delta frequency in the left subfigure and in the theta frequency in the right subfigure.

6.1.3 Differences between Channel Couples

The INS-CCorr was calculated by measuring the CCorr of 7x7 channels (49 couples, of which 28 are unique) and calculating its mean. The 28 couples of channels are not identical in INS-CCorr value measured during the INS-NFT sessions. This can be seen in figure 6.1.3.1. and 6.1.3.2., which illustrate the difference in INS-CCorr between two sessions (sessions one and two, two and three and one and three) of INS-NFT in the theta and delta band respectively for the experimental condition. Lighter and more yellow colours denote an increase in CCorr between sessions, while darker, red colours denote a decrease in CCorr. The significance of the difference of CCorr between session one and three is tested by performing 28 paired permutation tests with 100,000 permutations (H_a = the mean INS-CCorr during session 3 is different from the mean INS-CCorr during session 1, α = .05), of which the results can be viewed in Appendix M for delta-CCorr and Appendix N for theta-CCorr. The results can also be seen in figure 6.1.3.5. (top-left: delta, top-right: theta),

Figure 6.1.2.2: left: The reaction time difference (RTD) between team members for the experimental group, with the RTD on the y-axis and trials on the x-axis. The solid lines represent regression models (formula 12 and 13). Right: The RTD between team members for the experimental group, with the RTD on the y-axis and trials on the x-axis. The solid lines represent regression models (formula 14 and 15).

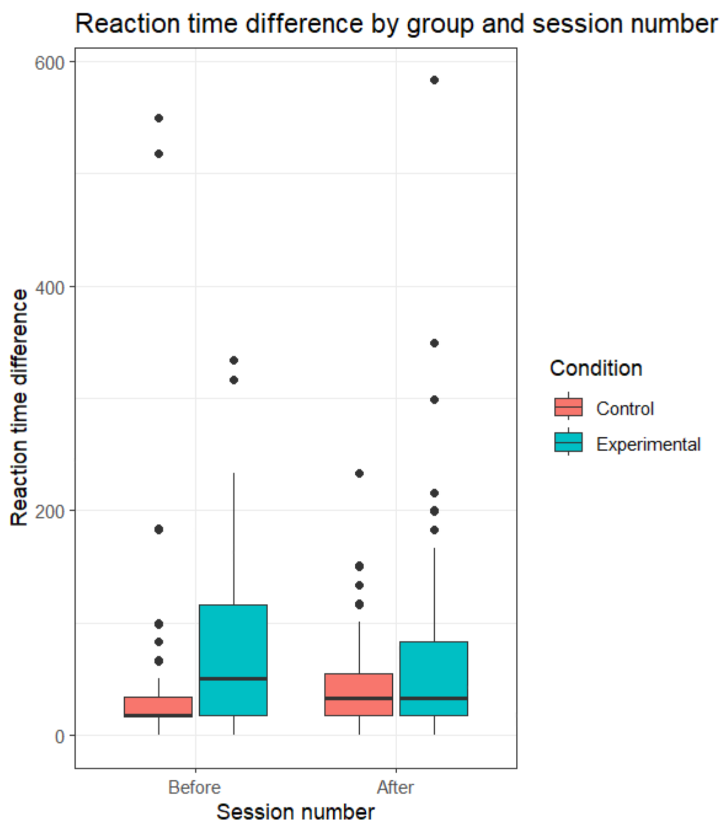


Figure 6.1.2.3: The reaction time difference (RTD) between team members (y-axis) before and after neurofeedback training (x-axis) and per condition (red: control, blue: experimental).

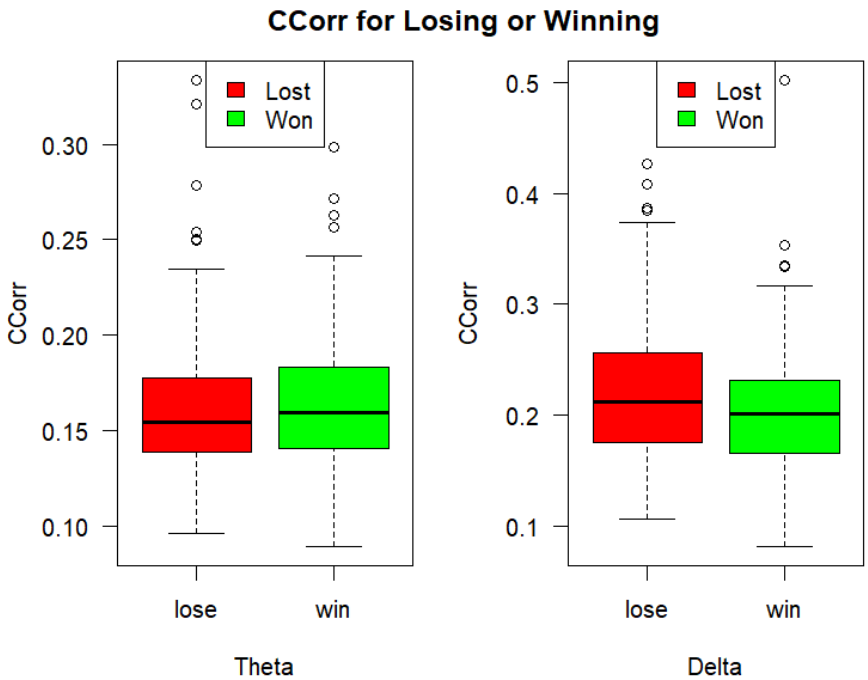


Figure 6.1.2.4: left: Box plot of the CCorr in the delta frequency before losing (red) or winning (green) a single trial. Right: Box plot of the CCorr in the theta frequency before losing (red) or winning (green) a single trial.

which illustrates all electrode pairs for which the change of CCorr between session one and three is significant. With Bonferroni Correction, the significant electrode pairs are F4xF8 and FzxF8 for delta, and no electrode pair has changed significantly for theta.

Within-session results can be seen in figures 6.1.3.3. and 6.1.3.4., which illustrate the change of INS-CCorr for the 28 measured electrode combinations between parts one and three of NFT for sessions one, two and three. This is illustrated for theta and delta respectively. Whether the difference in INS between part one and part three is significant was tested through 28 paired permutations tests with 100,000 permutations (H_a = the mean INS-CCorr during part 3 is different from the mean INS-CCorr during part 1, $\alpha = .05$), of which the results can be viewed in Appendix O for delta-CCorr and Appendix P for theta-CCorr. There are no electrode couples that changed significantly within sessions.

6.1.4 Topographic Maps

In this subsection figures are presented which illustrate the distribution of amplitude between the twenty-four measured electrodes as topographic maps. Firstly, the change in amplitude distribution between session one and session three of INS-NFT can be viewed in figure 6.1.4.1. for the experimental group, and figure 6.1.4.2. for the control group. This change has been calculated for five frequencies: delta, theta, alpha, beta and gamma. The change in delta and theta amplitide can be examined more closely in figure 6.1.4.5 (for delta) and 6.1.4.6. (for theta), where the results from the experimental group are shown next to the control group (between -4.5 μV and 4.5 μV for delta, and between -1.0 μV and 1.0 μV for theta).

Experimental delta amplitude decreased the most for both in Fp1, Fp2, F7 and P10 while increasing most in F3, Fz, F4, F8 and F10. Theta amplitude decreased most in Fp2, F7, F10 and F8, while increasing most for Fp1, F3, F4, Fz and C4. The increase or decrease in amplitude between sessions

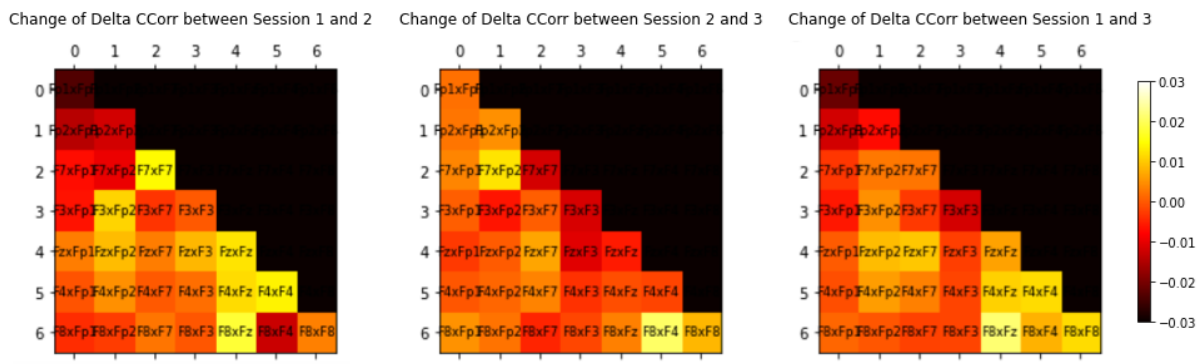


Figure 6.1.3.1: Change in delta-CCorr between part 1 and part 3 of each NFT session for the experimental condition. Left: The change in delta-CCorr between part 3 of session 1 and part 1 of session 1. Middle: The change in delta-CCorr between part 3 of session 2 and part 1 of session 2. Right: The change in delta-CCorr between part 3 of session 3 and part 1 of session 3. Lighter colors denote a higher CCorr value.

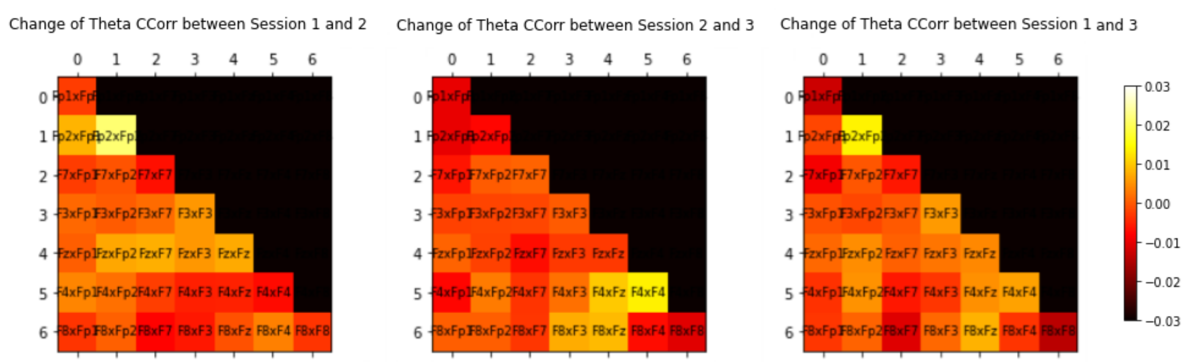


Figure 6.1.3.2: Change in theta-CCorr between part 1 and part 3 of each NFT session for the experimental condition. Left: The change in theta-CCorr between part 3 of session 1 and part 1 of session 1. Middle: The change in theta-CCorr between part 3 of session 2 and part 1 of session 2. Right: The change in theta-CCorr between part 3 of session 3 and part 1 of session 3. Lighter colors denote a higher CCorr value.

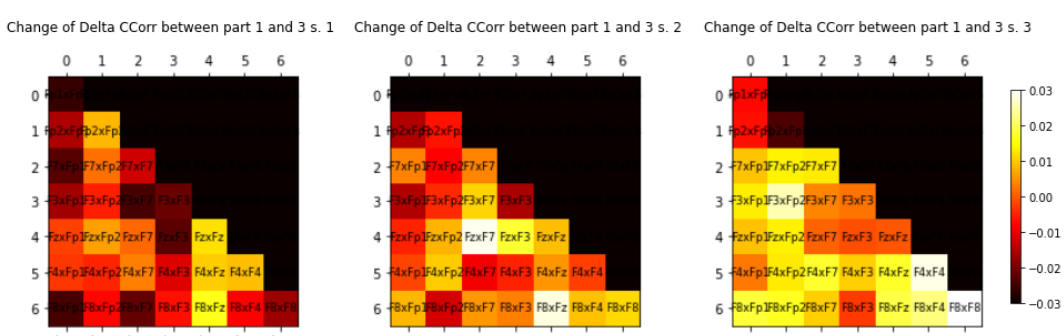


Figure 6.1.3.3: Change in delta-CCorr between part 1 and part 3 of each NFT session for the experimental condition. Left: The change in delta-CCorr between part 3 of session 1 and part 1 of session 1. Middle: The change in delta-CCorr between part 3 of session 2 and part 1 of session 2. Right: The change in delta-CCorr between part 3 of session 3 and part 1 of session 3. Lighter colors denote a higher CCorr value

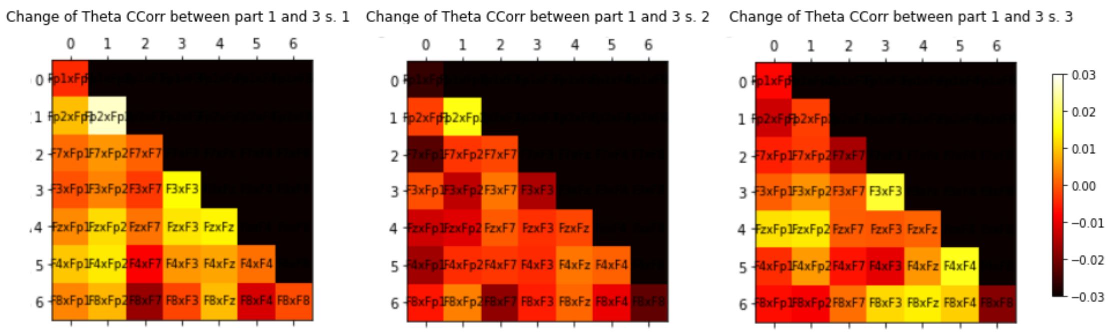


Figure 6.1.3.4: Change in theta-CCorr between part 1 and part 3 of each NFT session for the experimental condition. Left: The change in theta-CCorr between part 3 of session 1 and part 1 of session 1. Middle: The change in theta-CCorr between part 3 of session 2 and part 1 of session 2. Right: The change in theta-CCorr between part 3 of session 3 and part 1 of session 3. Lighter colors denote a higher CCorr value

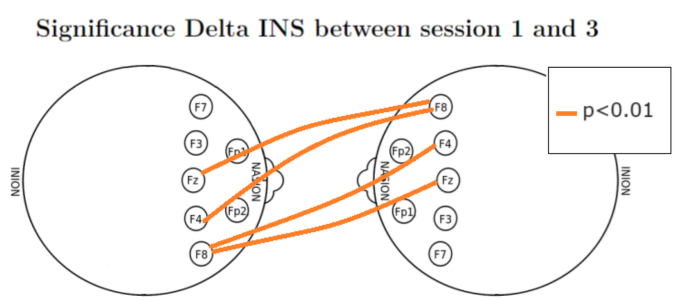


Figure 6.1.3.5: Electrode combinations that changed significantly in delta CCorr between session one and session three (F4xF8 and FzxF8). Both combinations increased for session three. All p-values are Bonferroni-corrected.

1 and 3 are smaller for the control group than for the experimental group. The significance of the change in amplitude was tested through paired permutation tests with 20000 permutations (H_a = there is a difference between the mean delta amplitude of session 1 and session 3, α = .05), of which the results can be viewed in appendix Q.

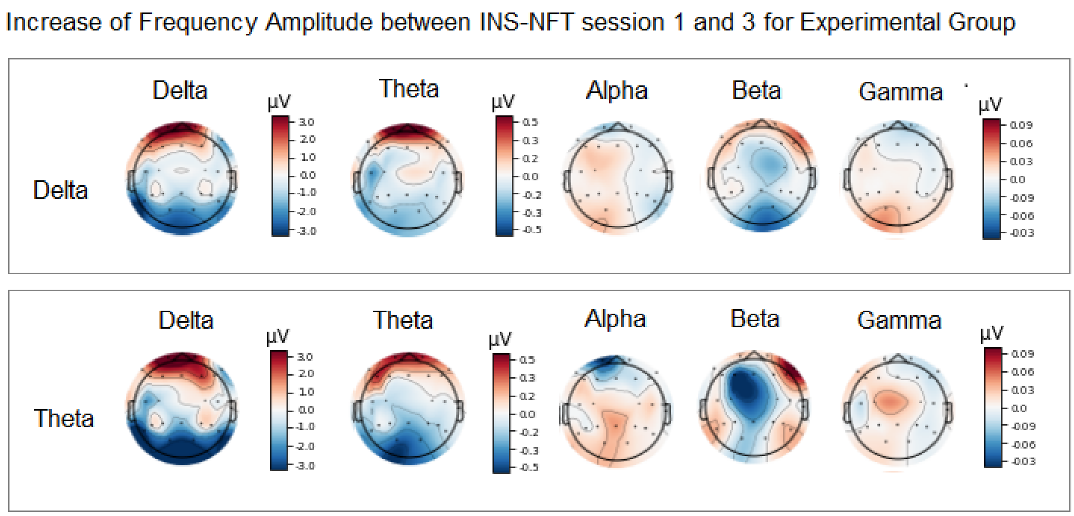


Figure 6.1.4.1: The change in frequency amplitude between INS-NFT session one (before INS-NFT) and three (after two sessions of INS-NFT) during Theta and Delta INS-NFT for the experimental condition. The difference in amplitude is plotted for values between -3.0 μV and 3.0 μV for delta, between -0.5 μV and 0.5 μV for theta and between -0.1 μV and 0.1 μV for alpha, beta and gamma.

The change in frequency amplitude is also plotted for the coordination task, which can be viewed in figure 6.1.4.4. This figure shows the topographic maps of amplitude distribution for the experimental and the control groups, for delta, theta, alpha, beta and gamma. Some differences between the groups can be seen in the delta, theta and beta bands, while the change in the alpha and gamma bands are relatively similar. For the experimental condition, delta amplitude increased most for Fp2, Fz, F4, F8, F10 and C4, and decreased for Fp1, F7, F3, T7, P10 and C3. In the control condition an increment of amplitude can be seen for Fp1 and Fp2, F9 and F7. In the theta band, Fp1 and Fp2 increased the most within the experimental condition, and this increase is greater than for the control condition. In the control condition, all but one trained electrode increased in amplitude (Fp1, Fp2, F9, F7, F3, F4, F8 and F10). For beta, the experimental group shows a greater increment of amplitude in O1, O2, P10, P8, P4, P3, Pz and T8 than in the control group. The decrease of beta is also greater in electrodes Fz, F4, F8, F10, C3, C4, T7 and F9.

6.2 Statistical Analyses

The following subsections report the results of nine statistical analyses which have been conducted to investigate Research Question 1 and Research Question 2. The results are presented in subsections 6.2.1. and 6.2.2. for Research question 1 and 2 respectively. None of the quantitative variables used in the analyses are normally distributed, which has been established through the Shapiro-Wilk Normality Test before running any statistical test with quantitative data. The results of the Shapiro-Wilk Normality Test can be viewed in Appendix E for RQ1 and Appendix F for RQ2. All of the results (except for H1d) are corrected for False Discovery Rate through the Benjamini-Hochberg method, which decreases the chance of false positives for large sets of statistical tests.

6.2.1 Research Question 1

H1.a. *The Mean CCorr in the theta and delta band is significantly larger for Part 3 of Session 3 than for Part 1 Session 1.*

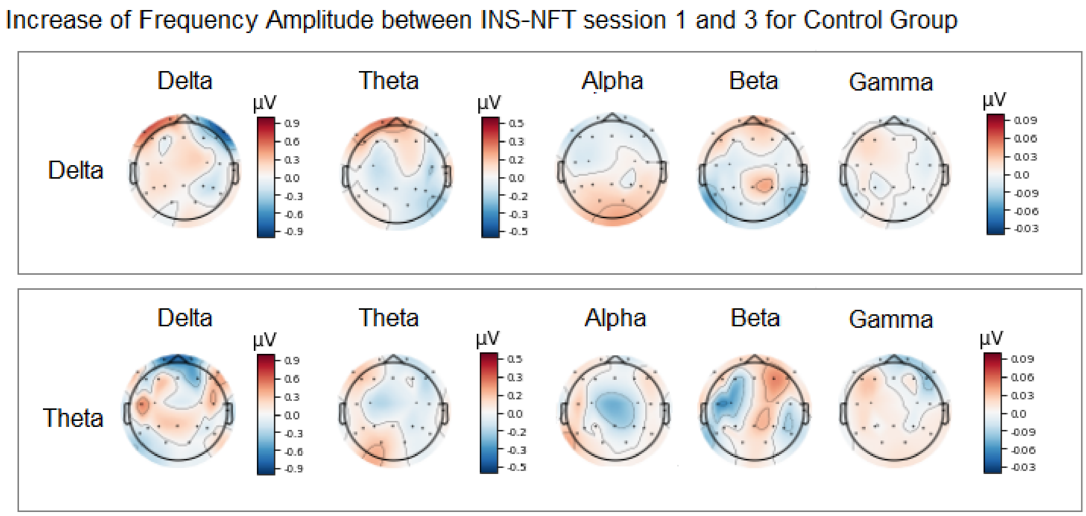


Figure 6.1.4.2: The change in frequency amplitude between INS-NFT session one (before INS-NFT) and three (after two sessions of INS-NFT) during Theta and Delta INS-NFT for the control condition. The difference in amplitude is plotted for values between -1.0 μV and 1.0 μV for delta, -0.5 μV and 0.5 μV for theta, and .1 μV and .1 μV for alpha, beta and gamma.

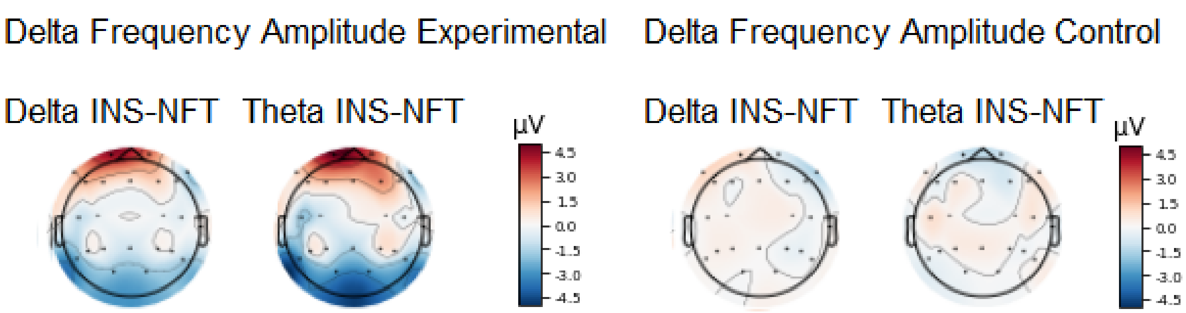


Figure 6.1.4.3: The change in delta amplitude between INS-NFT session one (before INS-NFT) and three (after two sessions of INS-NFT) during delta and theta INS-NFT. The difference in amplitude is plotted for values between -4.5 μV and 4.5 μV. Left: Experimental group. Right: Control group.

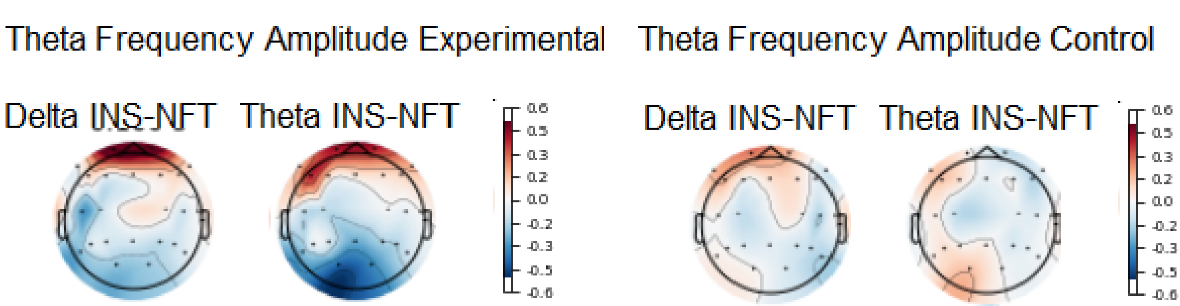


Figure 6.1.4.4: The change in theta amplitude between INS-NFT session one (before INS-NFT) and three (after two sessions of INS-NFT) during delta and theta INS-NFT. The difference in amplitude is plotted for values between -0.6 μV and 0.6 μV. Left: Experimental group. Right: Control group.

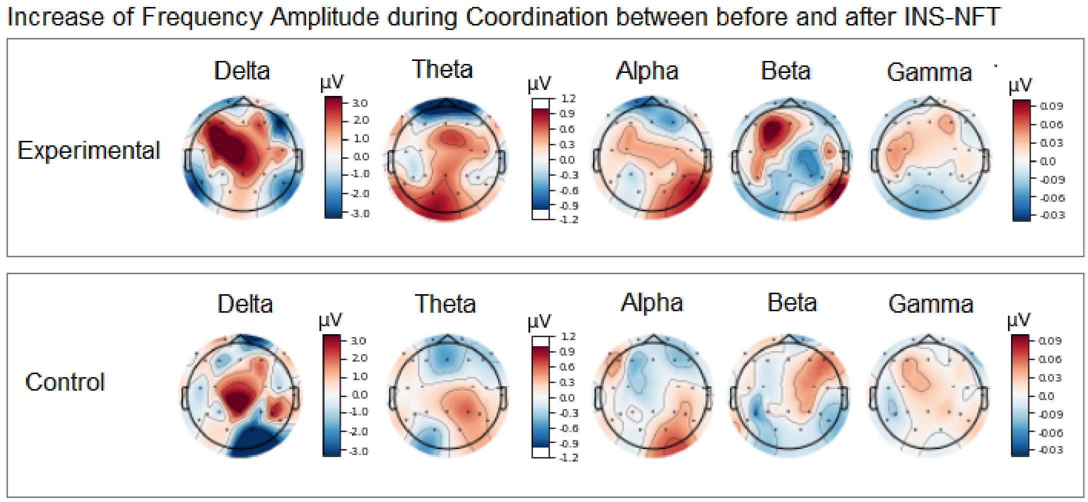


Figure 6.1.4.5: The change in frequency amplitude between team coordination task session one (before INS-NFT) and two (after INS-NFT) for Delta (0.5-4 Hz), Theta (4-8 Hz), Alpha (8-12 Hz), Beta (12-25 Hz) and Gamma (25-45 Hz). The difference in amplitude is plotted for values between -0.1 μV and 0.1 μV for alpha, beta, gamma and control theta, between -3.0 μV and 3.0 μV for delta and between -1.0 μV and 1.0 μV for theta.

Two paired permutation tests with 100,000 permutations were performed. For delta (H1: *The Mean CCorr in the delta band is significantly larger for Part 3 of Session 3 than for Part 1 Session 1*, H0: *The Mean CCorr in the delta band is not significantly larger for Part 3 of Session 3 than for Part 1 Session 1*) the results suggest that the delta CCorr is not significantly larger in Part 3 of Session 3 than in Part 1 Session 1 (p= .357 , $p > \alpha$, $\alpha = .05$). For theta (H1: *The Mean CCorr in the theta band is significantly larger for Part 3 of Session 3 than for Part 1 Session 1,* H0: *The Mean CCorr in the theta band is not significantly larger for Part 3 of Session 3 than for Part 1 Session 1*) the results suggest that theta CCorr is not significantly larger in Part 3 of Session 3 than in Part 1 Session 1 (p =.357, $p > \alpha$, $\alpha = .05$).

H1.b. The Mean CCorr in the theta and delta band is significantly larger for Part 3 of Session 3 than for Part 3 Session 1.

The paired permutation test (100,000 permutations) for delta (H1: *The Mean CCorr in delta band is significantly larger for Part 3 of Session 3 than for Part 3 Session 1*, H0: *The Mean CCorr in the delta band is not significantly larger for Part 3 of Session 3 than for Part 3 Session 1*) suggested no significant difference in CCorr for the delta band (p= .504, $p > \alpha$, $\alpha = .05$) nor for theta (H1: *The Mean CCorr in the theta band is significantly larger for Part 3 of Session 3 than for Part 3 Session 1*, H0: *The Mean CCorr in the theta band is not significantly larger for Part 3 of Session 3 than for Part 3 Session 1*) (p = .964, $p > \alpha$, $\alpha = .05$).

H1.c. Assumption: The Mean CCorr in the theta and delta band is not significantly higher for Part 1 of Session 1 of the experimental group than Part 1 of Session 1 of the control group.

In order to test the assumption a permutation test is performed for both the delta and theta frequency band. For delta (H1: *The Mean CCorr in the delta band is significantly higher for Part 1 of Session 1 of the experimental group than Part 1 of Session 1 of the control group*, H0: *Mean CCorr in the delta band is not significantly higher for Part 1 of Session 1 of the experimental group than Part 1 of Session 1 of the control group*) the results are not significant (p=.504, $p > \alpha$, $\alpha = .05$) and H0 can be assumed. For theta (H1: *The Mean CCorr in the theta band is significantly higher for Part 1 of Session 1 of the experimental group than Part 1 of Session 1 of the control group*, H0: *Mean CCorr in the theta band is not significantly higher for Part 1 of Session 1 of the* *experimental group than Part 1 of Session 1 of the control group*) the results are also not significant (p=.964, p $> \alpha$, $\alpha = .05$) and H0 can be assumed.

H1.c. The Mean CCorr in the theta and delta band is significantly higher for Session 3 Part 3 of the experimental group than in Session 3 Part 3 of the control group.

A permutation test is performed with 100,000 permutations for delta (H1: *The Mean CCorr in the delta band is significantly larger for Part 3 of Session 3 of the experimental group than for Part 3 Session 3 of the control group,* H0: T*he Mean CCorr in the delta band is not significantly larger for Part 3 of Session 3 of the experimental group than for Part 3 Session 3 of the control group*) suggests that the delta CCorr is not significantly larger in Part 3 of Session 3 of the experimental group than in Part 3 Session 3 of the control group (p= .964, p $> \alpha$, $\alpha = .05$). For theta (H1: *The Mean CCorr in the theta band is significantly larger for Part 3 of Session 3 of the experimental group than for Part 3 Session 3 of the control group,* H0: *The Mean CCorr in the theta band is not significantly larger for Part 3 of Session 3 of the experimental group than for Part 3 Session 3 of the control group*) the results suggests that the theta CCorr is not significantly larger in Part 3 of Session 3 of the experimental group than in Part 3 Session 3 of the control group (p=.357, p $> \alpha$, $\alpha = .05$).

H1.d.: There exists a correlation between CCorr and trial number for the experimental group.

A Spearman's rank correlation test is conducted in order to test this hypothesis (H1: *There exists a correlation between delta CCorr and trial number for the experimental group,* H0: *There does not exist a correlation between delta CCorr and trial number for the experimental group).* The results suggest a very weak positive correlation, r(1631)= $3.66*10^{-2}$, p= .140, p $> \alpha$, $\alpha = .05$. This is also the case for theta (H1: *There exists a correlation between theta CCorr and trial number for the experimental group,* H0: *There does not exist a correlation between theta CCorr and trial number for the experimental group):* r(1620)= $9.41*10^{-3}$., p= .705, p $> \alpha$, $\alpha = .05$.

6.2.2 Research Question 2

H2.a: Interpersonal Coordination Winning rate is higher after three sessions of (experimental group) INS-NFT

An Exact McNemar Test is performed to test the hypothesis (H1: *Interpersonal Coordination Winning rate is higher after three sessions of experimental INS-NFT, H0: Interpersonal Coordination Winning rate is not higher after three sessions of experimental INS-NFT)* which does not show a significant increase of score after INS-NFT (n = 240, m = 147, x = 77, p-value = .775, $\alpha = .05$).

The assumption of H2.a. is also tested:

H2.a assumption: Interpersonal Coordination Winning rate is not higher after three sessions of control INS-NFT

An Exact McNemar Test is performed (H1: *Interpersonal Coordination Winning rate is higher after three sessions of control INS-NFT, H0: Interpersonal Coordination Winning rate is not higher after three sessions of control INS-NFT) and is not significant* (n = 240, m = 75, x = 33, p-value = 1.00, $\alpha = .05$).

In order to investigate these results further, a similar statistical analysis is performed with the time difference between the reaction time of two team members (RTD) as a dependent variable:

H2.b: RTD is smaller after three sessions of (experimental group) INS-NFT

A paired permutation test is performed with 100,000 permutations (H1: *The RTD is smaller after three INS-NFT sessions for the experimental group , H0: The RTD is not smaller after three INS-NFT sessions for the experimental group)*, suggesting that RTD is not significantly smaller after NFT than before (p= .515, p $> \alpha$, $\alpha = .05$).

H2.b. Assumption: RTD is not smaller after three sessions of control INS-NFT

The same test is performed for the control group (H1: *The RTD is smaller after three INS-NFT sessions for the control group , H0: The RTD is not smaller after three INS-NFT sessions for the control group). T*his test did not show statistical significance (p -value = 1.00, p $> \alpha$, $\alpha = .05$).

H2.c: The winning rate after INS-NFT is higher in the experimental group than in the control group

To test this hypothesis a Fisher's Exact Test for Count Data was performed, as the number of trials was not sufficient for a Chi-square test. The analysis (H1: *The winning rate after INS-NFT is higher in the experimental group than in the control group , H0: The winning rate after INS-NFT is not higher in the experimental group than in the control group)*, suggests a significant increase in task score in the experimental group in comparison to the control group (p = $2.26{*}10^{-3}$, p $< \alpha$, $\alpha = .05$).

H2.c. Assumption: The winning rate before INS-NFT is not higher in the experimental group than in the control group

A Fisher's Exact Test for Count Data is performed (H1: *The winning rate before INS-NFT is higher in the experimental group than in the control group , H0: The winning rate before INS-NFT is not higher in the experimental group than in the control group)*, which is not significant and H1 can be rejected (p = 1.00, p $> \alpha$, $\alpha = .05$).

H2.d: The RTD after INS-NFT is smaller for the experimental group than for the control group.

An unpaired permutation test with 100,000 permutations (H1: *The RTD after INS-NFT is smaller for the experimental group than for the control group, H0: The RTD after INS-NFT is not smaller for the experimental group than for the control group)* suggests that RTD between two team members is not significantly smaller in the experimental group than in the control group after NFT (p= 1.00, p $> \alpha$, $\alpha = .05$).

H2.d. Assumption: The RTD before INS-NFT is not smaller for the experimental group than for the control group.

*A*n unpaired permutation test with 100,000 permutations is conducted (H1: *The RTD before INS-NFT is smaller for the experimental group than for the control group, H0: The RTD before INS-NFT is not smaller for the experimental group than for the control group).* The test is not significant and supports H0 (p= 1.00, p $> \alpha$, $\alpha = .05$).

H2.e: Delta and theta CCorr has a positive impact on winning an interpersonal coordination trial

A permutation test for the hypothesis for theta (H1: *Winning an interpersonal coordination trial is associated with significantly higher mean theta CCorr beforehand than losing, H0: Winning an interpersonal coordination trial is not associated with significantly higher mean theta CCorr beforehand than losing)* with 100,000 permutations suggests no significant effect (p = .962, p $> \alpha$, $\alpha = .1$). For delta *(H1: Winning an interpersonal coordination trial is associated with significantly higher mean theta CCorr beforehand than losing, H0: Winning an interpersonal coordination trial is not associated with significantly higher mean theta CCorr beforehand than losing)* the permutation test with 100,000 permutations does also not support the hypothesis (p =.653, p $> \alpha$, $\alpha = .05$).

6.3 NASA Task Load Index

The NASA Task Load Index was used to measure task load and frustration level during each NFT session (for which the results can be viewed in Appendix J.). There are no large differences between the control and experimental condition in overall task load. The score obtained as a summation for each factor (in the following order: mental demand, physical demand, temporal demand, performance, effort and frustration) and each condition (experimental and control) are: 136, 49, 48, 127, 157 and 109 (experimental); 144, 53, 74, 123, 157 and 96 (control). Each factor receives a particular weight that corresponds to its importance within the task, and the following weights have been selected: $w_{mental} = 1.0$; $w_{physical} = 0.5$; $w_{temporal} = 0.5$; $w_{performance} = 1.0$;

$w_{effort} = 0.5$; $w_{frustration} = 2.0$. The total score for the NASA Task Load Index is 608 for the experimental condition and 601 for the control condition (maximal score to be obtained: 1320).

There are two main differences between the control and the experimental conditions, which are the ratings of temporal demand and frustration. The control condition rates temporal demand higher than the experimental condition, while also rating frustration level lower. This is notable as frustration level is considered to be the most relevant factor for this study (see subsection 5.2.1, questionnaire). The highest rating of frustration was 13 for the experimental group and 19 for the control group. The frustration level is rated higher than 10 (above medium) in two measurements within the control condition, while it is rated higher than 10 for six measurements within the experimental condition. Frustration level decreased between sessions for participants B1, D1 and D2 (slightly), and increased for participants A2 and C2.

7 Discussion

The goal of the current thesis is to find answers for the following research questions:

> RQ1. Is frontal theta and delta INS-CCorr higher during a third session of INS-NFT than during the first session?
>
> RQ2. Is team coordination higher after three sessions of theta and delta frequency INS-NFT?

In order to answer these questions as thoroughly as possible with a small sample size, nine statistical analyses have been performed. The goal of these analyses is to investigate whether a trend exists towards further hypotheses.

In summary, the results for RQ1 show a trend towards the hypothesis that theta and delta INS-CCorr is not significantly higher during a third session of INS-NFT than during a first session, as the statistical tests did not show any significant increase in INS-CCorr after training (H1a and H1b) or in comparison to the control condition (H1c). INS-CCorr was also not strongly positively correlated to trial number.

The results for RQ2 show a trend towards the hypothesis that team coordination is not higher after three sessions of delta and theta INS-NFT. From the statistical analyses, only H2c showed significant results, suggesting that winning rate after INS-NFT is higher in the experimental condition than in the control condition. However, winning rate was not significantly higher after INS-NFT than before INS-NFT for the experimental condition (H2a). The reaction time difference was also not significantly smaller after INS-NFT (H2b), and it was not significantly smaller than in the control condition (H2d). The mean INS-CCorr was also not significantly higher before winning a trial than before losing (H2e).

A detailed discussion regarding the results, implications and limitations will follow in the coming subsections, which are split according to hypothesis (subsections 7.1 and 7.2 being about hypothesis 1 and hypothesis 2 respectively), followed by a methodological disussion (subsection 7.3), limitations (subsection 7.4) and future research (subsection 7.5).

7.1 Research Question 1

The results for RQ1 *(Is frontal theta and delta INS-CCorr higher during a third session of INS-NFT than during the first session?)* demonstrate a trend towards the hypothesis that theta and delta INS is not higher during the third session of INS-NFT than during the first session. This is firstly demonstrated by the statistical analysis of H1a (*The mean CCorr in the theta and delta band is significantly larger for Part 3 of Session 3 than for Part 1 Session 1)*, which did not yield statistically significant results. This suggests a trend for the null hypothesis, i.e. that the mean CCorr in the theta and delta band is not significantly larger for Part 3 of Session 3 than for Part 1 Session 1. Additionally, sub-hypothesis H1b (*The Mean CCorr in the theta and delta band is* *significantly larger for Part 3 of Session 3 than for Part 3 Session 1)* was also not supported by the results, suggesting that there is no difference between one session of INS-NFT and three sessions of INS-NFT. The results for H1a and H1b suggest that there is little effect of INS-NFT on the INS between two team members in the current setting.

It was also examined whether mean INS-CCorr would be higher after actual INS-NFT feedback than after random feedback (H1c, *(The Mean CCorr in the theta and delta band is significantly higher for Session 3 Part 3 of the experimental group than in Session 3 Part 3 of the control group)*. The hypothesis was not supported by the results. The assumption for H1c (*The Mean CCorr in the theta and delta band is not significantly higher for Part 1 of Session 1 of the experimental group than Part 1 of Session 1 of the control group)* was supported, meaning that INS-CCorr was not significantly higher in the experimental group than in the control group before training. There is also no significant difference between the control and experimental groups during the baselining period (see appendix H., illustrating a boxplot of the theta and delta INS-CCorr values per experimental group for the baseline period). These three results suggest that actual INS-NFT feedback does not increase INS significantly more than random feedback during INS-NFT.

Finally, the last sub-hypothesis tested whether there is a correlation between CCorr and trial number within the INS-NFT sessions (*H1d: There exists a correlation between CCorr and trial number for the experimental group).* The results suggest that this is not the case, as no significant correlation has been found between CCorr and trial number. All team INS-NFT training results can also be viewed in Appendix G, which shows the mean INS-CCorr value per trial for both experimental conditions.

In summary, the results do not suggest that one to three sessions of INS-NFT may have a significant effect on INS, while previous studies have reported trainability of INS within one session [77, 127]. This could be due to multiple reasons. In regards to the study by Yang et al (2020), there are many methodological differences between this study and the current study. The most relevant differences between these studies that may explain the difference in the obtained results are the difference in training schedule, trained frequency, training, motivation or difference in species. The pigeons in the study by Yang et al. trained INS (in PLV) for twenty minutes per day for three months, which is a significantly longer duration than the duration of training in the current study (72 minutes for two frequencies in three days). Additionally, the average training duration for performance enhancement in NFT studies is 3.8 to 7.7 sessions [93], which is also longer than the duration of the current study (three sessions). It is therefore possible that the learning rate for INS through INS-NFT is very low, and that more or longer sessions are necessary. Furthermore, the pigeons trained every day on a particular time of day, which is in contrast to the sometimes variable training schedule in the current study. The manner in which the pigeons have been trained is also different, as each trial started out with a resting period (two seconds) followed by the cue and synchronization measurements (one second) and a reward period (three seconds). It can be argued that the pigeons learned to increase INS for one second after the cue instead of an increase over the whole trial period. Moreover, the difference in motivation could also have had an effect on the results. The motivation for learning in the study by Yang et al. was enhanced by giving a food reward directly after winning a trial, and the pigeons were not fed before training. In contrast, the human participants in the current study received student credits or a monatary reward after the whole experiment (five sessions) and regardless of the obtained results. It has been known for some time that motivation is crucial for effective learning, as it promotes reflection, attention, interest and effort [11], and the level of motivation may predict NFT outcomes [50]).

As for the study by Müller et al. (2021), the participants trained delta and theta INS in a single session for eighty-four minutes. The study reported on significant differences in INS between baseline measurements and INS-NFT measurements in the delta and theta band but not in the alpha and beta band. In the current study, there did seem to be a small difference between baseline and task related INS in the theta band (see appendix K), but this difference does not seem to be as large as in the study of Müller et al. There are multiple possibilities that could explain the differences between these results. Firstly, one large difference between the studies is the difference in duration of training. In the current study, participants trained for twenty-four minutes per session, whilst in Müller et al. participants trained for eighty-four minutes. It is possible that a longer duration is necessary to gain significant results within a session of INS-NFT, and that multiple sessions may not have an effect. Another fairly large difference is the difference in electrode location. The current study trained Fp1, Fp2, F7, F3, F4, Fz and F8, while Müller et al. trained INS in F3, Fz, F4, C3, Cz, and C4. Other differences that may have had an effect are experimental setup, used similarity measure, slight difference in frequency (Müller only utilised 2.5 Hz for delta and 5.0 Hz for theta) and difference in sampling rate.

Regardless, it is also possible that the sample size was simply not large enough to measure any significant effect (see subsection 7.5, Limitations), as a small increase in CCorr can be seen in the graphs that visualize the results for RQ1 (see figures 6.1.1.1. and 6.1.1.2., section 6.1. Visualizations). Another reason could be that only a subsection of electrode combinations is suitable for INS-NFT, as only electrode combinations F4xF8 and FzxF8 have increased significantly within the experimental condition (delta).

The following subsection (7.1.1.) will look further into the results of the individual teams.

7.1.1 Differences between teams

Due to the small sample size (N=8, n=4, with 2 participants per team) it is important to look at the progress of individual teams, which can be viewed in Appendix G. In order to refer to the teams more easily, the four teams are labeled as team "A","B","C" and "D", with teams A and B being of the experimental group and team C and D of the control group (see figure 7.1.1.1.).

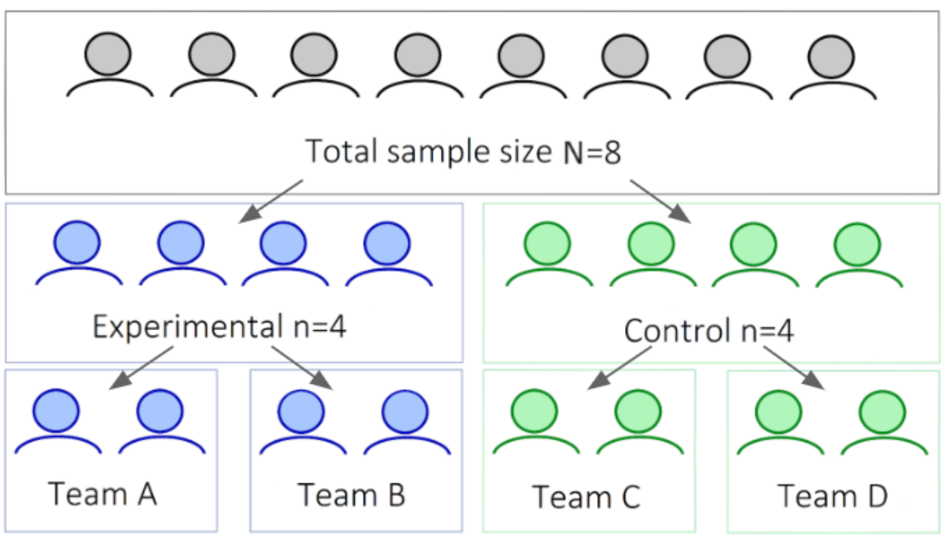


Figure 7.1.1.1: Subject teams.

Appendix G shows individual results per team and frequency as a scatter plot with linear regression lines. It can be seen that results vary per session, per frequency, per group and per condition. The experimental groups show increase in delta CCorr within-session for sessions two and three, but not for session one. Both experimental groups have a decline in CCorr within session one. Interestingly, opposite results can be seen for the theta band, with a decline for sessions two and three (except for team A in session three), and an increase during session one. Delta CCorr in team C of the control condition increased between sessions and is also seen increasing within each session. For team D, delta CCorr decreased for sessions one and two and increased during session three. Theta CCorr is similar for both teams C and D during sessions one, two and three, with the only difference being an increase in theta CCorr within session two of team D.

Investigating these results, it seems that team C from the control condition shows the most increase within and between sessions (delta), and team D also increased INS between sessions. This is interesting, as one would expect the experimental condition to increase INS more when looking at figure 6.1.1.1. in the result section for RQ2 (section 6.1.2.). Importantly, the results from section 6.1.2. compare part three of session three to part one of session one or part three of session one. The results from the scatterplots in appendix G show the regression models for the whole session. Moreover, the accuracy of the regression models may be low due to high variability of the data.

What the results from invidivual teams also show is the high variability within conditions.

7.2 Research Question 2

The results for RQ2 *(Is team coordination higher after three sessions of theta and delta frequency INS-NFT?)* supported a trend for the hypothesis that team coordination is not higher after three sessions of delta and theta INS-NFT. First, hypothesis H2a was tested *(H2a: Interpersonal Coordination Winning rate is higher after three sessions of (experimental group) INS-NFT).* The results were not statistically significant, although some increase can be seen through visual inspection of the results in figure 6.1.2.1. The control condition did not show an increase in winning rate, and the experimental group had a significantly larger winning rate than the control group (*H2c: The winning rate after INS-NFT is higher in the experimental group than in the control group*), which was also significant after a Bonferroni correction. The winning rate was not significantly higher for the experimental condition than for the control condition before INS-NFT (*H2c assumption: The winning rate before INS-NFT is not higher in the experimental group than in the control group.* These results may suggest a possible effect of INS-NFT on interpersonal coordination winning rate, but it is also possible that the winning rate improved in the experimental groups due to learning and that the winning rate decreased in the control groups due to human error (two or three errorous trials have been reported, see section 7.4) or simply a sample size that is too small. Another explanation may be that random feedback during INS-NFT decreased team coordination. Importantly, no large differences have been found between teams in terms of frustration level or task load (see section 6.3, NASA Task Load Index), and no significant problems, frustrations or mistakes have been communicated to the researcher about the tasks or the experiment (except for the previously mentioned errorous trials).

While the winning rate can be seen as a representation of team performance which can only be achieved through interpersonal coordination, the RTD is a more precise representation of interpersonal coordination. Therefore, the RTD is also analyzed. First, it was investigated whether the RTD is smaller after three sessions of (experimental group) INS-NFT (H2b). This hypothesis was tested and found not to be statistically significant. The assumption that *The RTD is not smaller after three sessions of (control group) INS-NFT* (H2b assumption) is statistically significant. The RTD is also not significantly smaller in the experimental group than in the control group after INS-NFT, which is also not the case before INS-NFT (H2d assumption). These results may be interpreted as that there are no differences in RTD between conditions and before and after INS-NFT. However, when inspecting visualizations of the results (figure 6.1.2.3) it can be seen that RTD decreased in the experimental group after INS-NFT, while it increased in the control group. These results are not significant, which could be due to the training not having a large enough effect or the sample size being too small for statistical significance. The variability for the results within groups is high, which may be due to the many variables that may affect interpersonal coordination training results (e.g. individual learning rate, tiredness [115] and perceived partner likability [110]).

The last sub-hypothesis that was tested is H2e (*Winning an interpersonal coordination trial is associated with significantly higher mean delta and theta CCorr beforehand than losing).* The results are not significant for theta and not for delta. Some studies have reported a higher INS during task success than during task failure in fMRI [131] and EEG [109]. These results do not support such an effect.

While it was investigated whether INS-NFT would have an effect on interpersonal coordination, the main purpose of the investigation into RQ2 is to examine the relationship between INS and interpersonal coordination within a team. Multiple studies have reported an increase of INS during cooperative interpersonal coordination, but this relationship had not been researched thoroughly enough to explain whether the possible effect may be due to causation or correlation. It was theorized within this study that INS could have a positive effect on interpersonal coordination (see section 4, Research Question), but this theory remains unresearched as INS was not increased significantly after INS-NFT.

7.3 Methodology

The current study is the first study to examine the effect of multiple neurofeedback training sessions on INS and interpersonal coordination. The methodology for this type of study has therefore not yet been estabilished, and one goal of the current study is to lay a foundation for future studies to build upon. For this reason this subsection will provide an overview of the downsides of the current methodology, and which alterations may be applied in order to obtain higher quality results in future research.

7.3.1 Participant Selection

Participant selection is an important part of an experiment in INS-NFT, as NFT results depend on a number of psychological variables such as attention, motivation, mood and various personality traits [50], while there are results to support that INS may be dependent on the amount of affinity for the other team member [77, 65]. . When selecting participants for the purpose of training INS-NFT it is therefore recommended to select participants who are likely to learn (preferably as quickly as possible) and have a similar amount of affinity towards their team members. As partner likability and mood are difficult to control for within INS-NFT research, it is recommended to use a questionnaire in order to assess these variables and to determine their possible effect.

In order to increase instrinsic motivation for NFT, it may be helpful to select participants with an affinity with the research and the research topic. It has been seen throughout the study that participants who were enthousiastic about the research were more interested in finding the best methodology to synchronize with their team member. This was apparent as these participants mentioned verbally to have tried a number of methods to synchronize, until they found a methodology that seemed to work (whether that would be in the experimental condition or control). Participants that were less enthousiastic for the research did not try as many methods in order to synchronize. One participant even mentioned that they did not try any particular methodology for the last of the three sessions, and they chose to meditate through the whole session.

Additionally, it is recommended to test for Neurofeedback training-responsiveness for a longer period of time than what was performed in the current study. While there is no clear scientifically defined manner in which NFT-responsiveness screening should be performed, it is recommended to use a simple and quite short task, such as the task used by Escalano et al. (2012) [30]. For the current thesis, the task was shortened and the calculation for upper-alpha was simplified. However, it would be more reliable to perform a longer session of upper-alpha upregulation, more similar to Escalano et al. [30].

7.3.2 Participant Treatment

This study did not have any strict rules regarding conversing with team members before or after the experimental tasks. This is however not an ideal environment, as the team members may increase or decrease affinity with each other through conversation. The methodology for RQ1 and RQ2 may improve from a prohibition in conversation between participants who are in the same team.

7.3.3 Neurofeedback Training Task

The participants have been positive about the design of the NFT-task. Overall it was found to be easily understandable and relatively fun. All participants have reported at least once that they became excited at the moment that the heads started to move towards each other. It is unknown whether this excitement could have had an influence on INS during training. If the excitement after feedback enhances or decreases INS, this could have influenced the training to some extent. It may be interesting to investigate whether results would be different with a resting period between feedback and a new trial, such as in the research by Yang et al. (2020). The downside to this methodology is that INS after a stimulus cue may have different effects from constant INS.

The duration of the complete NFT-task (26 minutes) was found to be do-able by all participants, while the duration of each block (12 minutes) was found to be somewhat long and a little difficult to sit through. Overall, the task design may improve from shorther blocks of NFT trials, with resting time between the blocks. The ideal duration for each block would probably be somewhere between four and eight minutes. The ideal duration for the complete INS-NFT task would be around 20 to 30 minutes, as this would be around the maximum time to maintain the participant's attention span in the current setting.

It may be helpful to add a different looking NFT task to the training schedule as this may prevent participants from getting bored and losing attention. This is especially important if the number of sessions would be increased, or if the INS-NFT task would take longer than 26 minutes. An alternative INS-NFT task would ideally work similarly to the current task, but would have more variable and engaging aesthetics. It would be preferable to give the participants rewards if the results are positive. One may think of rewarding participants with praise (on-screen) after reaching the threshold.

It is also recommended to work with the highest possible sampling frequency, as this would increase the accuracy of the measurement.

7.3.4 Interpersonal Coordination Task

The coordination task was found to be fun, easy to understand and do-able by the participants. There are some downsides to the task, however. It can be argued that the score that participants could gain from this task does not measure the interpersonal coordination directly. A more accurate measure would be the RTD, which measures the difference in reaction time between the two team members. It may be more useful to use the RTD in future studies as a measure of interpersonal coordination, and proceed using the task as it has been used in the current study. One alteration should be made, however, and that is to make it clear to the participants which player number they have at all times. During one of the sessions one participant forgot their player number and did not know for a couple of trials whether they had to increase or decrease their reaction speed in order to coordinatate their actions with their team member. In order to prevent this from happening one can place a reminder of the player number on each computer screen.

7.4 Limitations

There have been a number of limitations to this experiment. The primary limitation is the low number of participants per experimental condition, which is only four per condition. Additionally, all participants have been grouped together in teams of two, and all data points used for the statistical analyses have been acquired from teams and not individuals. This means that the data acquired is from two individual teams per condition, which is an objectively low number for a sample. It can be argued that this sample size is not representative of the overall population, especially as there are multiple variables that are difficult to control for that may affect INS, such as the liking of the other team member or the mental state of an individual at a given moment [65].

The methods deployed by participants for the training of INS may also have some influence on the results. One participant from the experimental condition (team B) explained that they tried meditation in order to increase INS during session 3. This method has not been used by any other participant, and as this participant did not change meditation style according to the feedback, it can be argued that it did not support the learning of INS. Due to the small sample size, even single participants may have influenced the results by deploying relatively ineffective training techniques.

Furthermore, this research is primarily about learning. It has been known for some time that learning rate depends on a variety of factors (e.g. motivation [32], personality, mood and attention [50]) which are different for each individual, which means that the learning rate may vary strongly between teams. The sample size that has been used for this study may have been insufficient to cover for all of the possible differences in learning rate. It is also plausible that, while all participants have been screened for NFT-responsiveness, some participants were not able to be trained through NFT or were not able to train INS.

Moreover, the similarity measure which was used to measure INS (adjusted-CCorr) has previously not been used in any hyperscanning studies. While results for adjusted-CCorr show lower variability and accumulated error than for CCorr [133], there is no literature to confirm these results.

One practical limitation of this research is the amount of time in between the given sessions. Unfortunately, much may happen in between training or task sessions which may slow the training down, such as illness, strikes of public transport, traffic queues or important personal matters. As these events have taken place at some point in time, some of the training sessions were not scheduled with one or two days in between the sessions. In practice, the minimum amount of time in between sessions was approximately 24 hours, while the maximum amount of time has been approximately seven days.

Another practical limitation is that two different labs have been used for the experiment. All teams in the control condition have been trained in the same laboratory, and all teams in the experimental condition have been trained in a different laboratory. The experimental setup was exactly the same for both labs, but the laboratory used in the control condition was slightly more spacious.

Finally, two mistakes have been made during the experimental phase. Firstly, during session one of team A, some of the electrodes showed many artifacts which led to these electrodes being eliminated from the calculation of CCorr in offline analysis. These electrodes had been included in the calculation of CCorr during online INS-NFT, which made the training of session 1 slower (as participants did not receive feedback if the amplitude was too high) and less effective. Additionally, a mistake was made during the team coordination task of team C after INS-NFT. One of the team members mentioned that they became confused about which participant number they had for approximately two or three rounds, but eventually understood their participant number through the feedback that was provided.

7.5 Future Research

In the current section the recommendations for future research are discussed for RQ1 and RQ2.

7.5.1 Research question 1

The findings from this exploratory study did not show a trend towards INS being trainable through NFT. The question arises whether the results are interesting enough to continue examining this research question on a larger scale. There are arguments that can be made in support of the continuation of the research as well as arguments against continuation.

First, the arguments in support of the continuation of research. As is mentioned in the discussion section (section 7.1.), previous work has showed some support for the hypothesis that INS can be trained. The training in these experiments was of a longer duration in terms of sessions (Yang et al., 2020) or session length (Müller et al., 2021) than the training in the current experiment. While the results from the statistical analyses have not been significant, visual inspection shows some increment in INS (figures 6.1.1.1. and 6.1.1.2.), and it is possible that this increase proceeds with more sessions. The average number of NFT training sessions for performance enhancing is 3.8 to 7.7 [93], and due to the complexity of the training (as the result is dependent on two individuals instead of one) it is not unlikely that INS-NFT would require more training sessions than were conducted in the current study.

Another argument that can be made for continuation is that there exists much variability in the learning process between teams and between sessions, which makes it likely that there exist variables that intervene with the learning progress which could be made less relevant if the sample number would be larger.

There are also arguments against continuation. The main argument is that the results from this experiment do not seem to show a trend towards the hypothesis that INS is trainable through NFT. The statistical analyses did not yield significant results, the graphs from individual teams did not show more growth of INS in the experimental condition (Appendix G.) and electrode matrices did not show a stronger increase of INS for the experimental condition (Appendix I and figures 6.1.3.1. - 6.1.3.4.). Additionally, the research may become very time consuming due to the recommended number of sessions (3.8 to 7.7), larger sample size and longer NFT-responsiveness screening duration.

If further research is conducted into RQ1, it is advised to include electrode couples F8xF4 and F8xFz in INS calculation due to the significant change in delta INS-NFT within these electrode pairs. Without a Bonferroni correction, the increase in electrode pair F4xFz would also have been significant between session one and three, and the increase in FzxF7 and Fp1xF7 would have been significant between part one and three of the sessions. Therefore, it may be interesting to investigate training effectiveness in F4xFz, FzxF7 and Fp1xF7.

Training a single frequency instead of two may be more effective for future research, as more time could be invested into the training of a single frequency and less variables could interfere with the results. As delta INS increased the most from visual inspection and was increased significantly for some electrode pairs, it is recommended to train INS in the delta frequency.

Additionally, it may be valuable to investigate which variables may account for the learning variability between sessions and teams, and examine whether the effect of learning would increase with other motivating factors.

7.5.2 Research question 2

The majority of the findings do not support a trend towards the hypothesis that INS-NFT has an effect on interpersonal coordination between team members. There are some arguments that can be made in favour of continuation of the research into RQ2, and some arguments that can be made against.

Firstly, the arguments in favour of the continuation of research. The experimental results showed some growth in interpersonal coordination for the experimental group in both RTD and interpersonal coordination score, which was not large enough to be statistically significant. It can be argued that the sample size for the experimental condition (n=2) was not large enough to achieve statistically significant growth in studies like this, as too many variables may interfere with the results (e.g. individual learning rate, tiredness [115], perceived partner likability [110]) which may also be observed from the high variability in experimental results. There are also results that support a trend towards the hypothesis, as the winning rate after INS-NFT was significantly higher for the experimental condition than for the control condition.

To the contrary, there are also arguments that can be made against continuation of the research. The results of the majority of the statistical analyses have not been significant. This indicates some trend towards the hypothesis that INS-NFT does not increase interpersonal coordination. Additionally, the decrease in winning rate within the control condition may in part be explained by the mistake that had been made by one of the participants. Another reason for discontinuation is that it is possible that INS is not trainable through INS-NFT, which would take away most of the reason for investigating RQ2. Indeed, the main purpose of the investigation into RQ2 was to examine whether an increase in INS could increase interpersonal coordination. As the results for RQ1 show no trend towards the hypothesis that INS is trainable through INS-NFT, it is unknown whether future studies will be able to increase INS and investigate this relationship further.

Overall, it is generally not recommended to continue research into RQ2 unless the experiment is combined with further research into RQ1. The reason for this is that there is no trend in empirical data to support the hypothesis that INS-NFT is effective in increasing INS. If research into RQ2 would be continued, it would be advised to include a condition without INS-NFT to investigate the effect of INS-NFT on interpersonal coordination. It is recommended to train a single frequency instead of two, as this would leave out the possiblity of interfering effects, with the additional benefit of being able to prolong training time.

8 Conclusion

The current study is one of the first studies to investigate whether interpersonal neural synchrony can be increased through multi-session neurofeedback training, and the first to examine the effect of this training on interpersonal coordination between team members. Two research questions have been investigated:

> RQ1. Is frontal theta and delta INS-CCorr higher during a third session of INS-NFT than during the first session?
>
> RQ2. Is team coordination higher after three sessions of theta and delta frequency INS-NFT?

These research questions have been investigated by performing an experiment with nine statistical analyses. The results showed a trend towards the hypothesis that delta and theta INS in frontal regions is not larger during the third session of INS-NFT than during the first session. It also showed a trend towards the hypothesis that team coordination does not increase after three sessions of delta and theta INS-NFT.

The results from the statistical analyses suggest that mean INS-CCorr is not significantly larger after three sessions of INS-NFT, and that giving actual neurofeedback did not yield a significantly different result in INS-CCorr value from giving random feedback. There is also no significant correlation between INS-NFT and trial number, which suggests that INS-NFT does not significantly increase with each trial during a training session. Future research should focus on researching the effectiveness of training electrode pairs F8xFz and F8xF4, and it is recommended to train INS in the delta frequency.

The majority of the results for RQ2 did not yield statistically significant results, indicating no significant increase of interpersonal coordination after INS-NFT. The hypothesis that interpersonal coordination winning rate is higher after INS-NFT for the experimental condition than for the control condition was statistically significant. There was also no significant difference found in INS-CCorr value between losing or winning an interpersonal coordination trial. Future research could investigate the effect of control-type INS-NFT, and additionally the effect of INS-NFT when INS is significantly increased.

This study has looked thoroughly at the effect of multi-session INS-NFT on INS and on interpersonal coordination. It can be concluded that at this moment there is no trend to suggest that INS-NFT can be used to increase INS between two team members effectively, and that there is no trend to suggest that interpersonal coordination can be increased by INS-NFT.

References

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

[2] D. Araújo and J. Bourbousson. Theoretical perspectives on interpersonal coordination for team behaviour. In *Interpersonal coordination and performance in social systems*, pages 144–157. Routledge, 2016.

[3] H. Arrow, J. E. McGrath, and J. L. Berdahl. *Small groups as complex systems: Formation, coordination, development, and adaptation*. Sage Publications, 2000.

[4] F. G. Ashby. An introduction to fmri. *An introduction to model-based cognitive neuroscience*, pages 91–112, 2015.

[5] L. Astolfi, J. Toppi, F. De Vico Fallani, G. Vecchiato, S. Salinari, D. Mattia, F. Cincotti, and F. Babiloni. Neuroelectrical hyperscanning measures simultaneous brain activity in humans. *Brain topography*, 23:243–256, 2010.

[6] M. Balconi, E. Brambilla, and L. Falbo. Appetitive vs. defensive responses to emotional cues. autonomic measures and brain oscillation modulation. *Brain Research*, 1296:72–84, 2009.

[7] M. Balconi, M. E. Vanutelli, et al. Eeg hyperscanning and behavioral synchronization during a joint actions. *Neuropsychological Trends*, 2018(24):23–47, 2018.

[8] M. Balconi, I. Venturella, G. Fronda, and M. E. Vanutelli. Leader-employee emotional "interpersonal tuning". an eeg coherence study. *Social Neuroscience*, 15(2):234–243, 2020.

[9] D. Bevilacqua, I. Davidesco, L. Wan, K. Chaloner, J. Rowland, M. Ding, D. Poeppel, and S. Dikker. Brain-to-brain synchrony and learning outcomes vary by student–teacher dynamics: Evidence from a real-world classroom electroencephalography study. *Journal of cognitive neuroscience*, 31(3):401–411, 2019.

[10] P. Billeke, F. Zamorano, D. Cosmelli, and F. Aboitiz. Oscillatory brain activity correlates with risk perception and predicts social decisions. *Cerebral Cortex*, 23(12):2872–2883, 2013.

[11] M. Borah. Motivation in learning. *Journal of Critical Reviews*, 8(2):550–552, 2021.

[12] R. Borbás, L. V. Fehlbaum, U. Rudin, C. Stadler, and N. M. Raschle. Neural correlates of theory of mind in children and adults using catoon: Introducing an open-source child-friendly neuroimaging task. *Developmental Cognitive Neuroscience*, 49:100959, 2021.

[13] A. P. Burgess. On the interpretation of synchronization in eeg hyperscanning studies: a cautionary note. *Frontiers in human neuroscience*, 7:881, 2013.

[14] V. K. Campos da Paz, A. Garcia, A. Campos da Paz Neto, and C. Tomaz. Smr neurofeedback training facilitates working memory performance in healthy older adults: A behavioral and eeg study. *Frontiers in behavioral neuroscience*, 12:321, 2018.

[15] P. Chen, S. Hendrikse, K. Sargent, M. Romani, M. Oostrik, T. F. Wilderjans, S. Koole, G. Dumas, D. Medine, and S. Dikker. Hybrid harmony: a multi-person neurofeedback application for interpersonal synchrony. *Frontiers in Neuroergonomics*, 2:687108, 2021.

[16] Q. Chen. Eeg hyperscanning study of team neurodynamics analysis during cooperative and competitive interaction. Master's thesis, University of Twente, 2019.

[17] X. Cheng, X. Li, and Y. Hu. Synchronous brain activity during cooperative exchange depends on gender of partner: A fnirs-based hyperscanning study. *Human brain mapping*, 36(6):2039–2048, 2015.

[18] M. X. Cohen. Where does eeg come from and what does it mean? *Trends in neurosciences*, 40(4):208–218, 2017.

[19] S. S. Cohen and L. C. Parra. Memorable audiovisual narratives synchronize sensory and supramodal neural responses. *ENeuro*, 3(6), 2016.

[20] X. Cui, D. M. Bryant, and A. L. Reiss. Nirs-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *Neuroimage*, 59(3):2430–2437, 2012.

[21] A. Czeszumski, S. Eustergerling, A. Lang, D. Menrath, M. Gerstenberger, S. Schuberth, F. Schreiber, Z. Z. Rendon, and P. König. Hyperscanning: a valid method to study neural inter-brain underpinnings of social interaction. *Frontiers in Human Neuroscience*, 14:39, 2020.

[22] I. Davidesco, E. Laurent, H. Valk, T. West, S. Dikker, C. Milne, and D. Poeppel. Brain-to-brain synchrony predicts long-term memory retention more accurately than individual brain measures. *BioRxiv*, page 644047, 2019.

[23] M. Demir, A. D. Likens, N. J. Cooke, P. G. Amazeen, and N. J. McNeese. Team coordination and effectiveness in human-autonomy teaming. *IEEE Transactions on Human-Machine Systems*, 49(2):150–159, 2018.

[24] S. Dikker, L. Wan, I. Davidesco, L. Kaggen, M. Oostrik, J. McClintock, J. Rowland, G. Michalareas, J. J. Van Bavel, M. Ding, et al. Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current biology*, 27(9):1375–1380, 2017.

[25] M. L. Dixon, R. Thiruchselvam, R. Todd, and K. Christoff. Emotion and the prefrontal cortex: An integrative review. *Psychological bulletin*, 143(10):1033, 2017.

[26] M. Doppelmayr and E. Weber. Effects of smr and theta/beta neurofeedback on reaction times, spatial abilities, and creativity. *Journal of Neurotherapy*, 15(2):115–129, 2011.

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

[28] M. Ertl, M. Hildebrandt, K. Ourina, G. Leicht, and C. Mulert. Emotion regulation by cognitive reappraisal—the role of frontal theta oscillations. *NeuroImage*, 81:412–421, 2013.

[29] C. Escolano, M. Aguilar, and J. Minguez. Eeg-based upper alpha neurofeedback training improves working memory performance. In *2011 annual international conference of the IEEE engineering in medicine and biology society*, pages 2327–2330. IEEE, 2011.

[30] C. Escolano, B. Oliván, Y. Lopez-del Hoyo, J. Garcia-Campayo, and J. Minguez. Double-blind single-session neurofeedback training in upper-alpha for cognitive enhancement of healthy subjects. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4643–4647. IEEE, 2012.

[31] X. Feng, B. Sun, C. Chen, W. Li, Y. Wang, W. Zhang, W. Xiao, and Y. Shao. Self–other overlap and interpersonal neural synchronization serially mediate the effect of behavioral synchronization on prosociality. *Social Cognitive and Affective Neuroscience*, 15(2):203–214, 2020.

[32] J. Filgona, J. Sakiyo, D. Gwany, and A. Okoronka. Motivation in learning. *Asian Journal of Education and social studies*, 10(4):16–37, 2020.

[33] F. A. Fishburn, V. P. Murty, C. O. Hlutkowsky, C. E. MacGillivray, L. M. Bemis, M. E. Murphy, T. J. Huppert, and S. B. Perlman. Putting our heads together: interpersonal neural synchronization as a biological mechanism for shared intentionality. *Social cognitive and affective neuroscience*, 13(8):841–849, 2018.

[34] K. Gouweleeuw. Using neurophysiological signals to measure social exclusion induced by a language barrier. Master's thesis, University of Twente, 2021.

[35] A. Grinsted, J. C. Moore, and S. Jevrejeva. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear processes in geophysics*, 11(5/6):561–566, 2004.

[36] M. Grosse-Wentrup, B. Schölkopf, and J. Hill. Causal influence of gamma oscillations on the sensorimotor rhythm. *NeuroImage*, 56(2):837–842, 2011.

[37] I. Gumilar. Inter-brain synchronization during collaboration in virtual reality. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 491–494. IEEE, 2021.

[38] H. Z. Gvirts and R. Perlmutter. What guides us to neurally and behaviorally align with anyone specific? a neurobiological model based on fnirs hyperscanning studies. *The Neuroscientist*, 26(2):108–116, 2020.

[39] S. G. Hart. Nasa task load index (tlx). 1986.

[40] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

[41] C. S. Herrmann, D. Strüber, R. F. Helfrich, and A. K. Engel. Eeg oscillations: from correlation to causality. *International Journal of Psychophysiology*, 103:12–21, 2016.

[42] J. Hogeveen, C. Salvi, and J. Grafman. 'emotional intelligence': lessons from lesions. *Trends in neurosciences*, 39(10):694–705, 2016.

[43] B. Hommel, L. S. Colzato, and W. P. Van Den Wildenberg. How social are task representations? *Psychological Science*, 20(7):794–798, 2009.

[44] Y. Hu, Y. Hu, X. Li, Y. Pan, and X. Cheng. Brain-to-brain synchronization across two persons predicts mutual prosociality. *Social cognitive and affective neuroscience*, 12(12):1835–1844, 2017.

[45] Y. Hu, Y. Pan, X. Shi, Q. Cai, X. Li, and X. Cheng. Inter-brain synchrony and cooperation context in interactive decision making. *Biological psychology*, 133:54–62, 2018.

[46] M. Iacoboni and M. Dapretto. The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience*, 7(12):942–951, 2006.

[47] P. Israsena, S. Jirayucharoensak, S. Hemrungrojn, S. Pan-Ngum, et al. Brain exercising games with consumer-grade single-channel electroencephalogram neurofeedback: pre-post intervention study. *JMIR Serious Games*, 9(2):e26872, 2021.

[48] S. Järvelä, M. Salminen, A. Ruonala, J. Timonen, K. Mannermaa, N. Ravaja, and G. Jacucci. Dynecom: Augmenting empathy in vr with dyadic synchrony neurofeedback. 2019.

[49] M. Jensen, E. Hüttenrauch, J. Schmidt, G. Andersson, M.-L. Chavanon, and C. Weise. Neurofeedback for tinnitus: study protocol for a randomised controlled trial assessing the specificity of an alpha/delta neurofeedback training protocol in alleviating both sound perception and psychological distress in a cohort of chronic tinnitus sufferers. *Trials*, 21(1):1–13, 2020.

[50] K. C. Kadosh and G. Staunton. A systematic review of the psychological factors that influence neurofeedback learning outcomes. *Neuroimage*, 185:545–555, 2019.

[51] E. Kalbe, M. Schlegel, A. T. Sack, D. A. Nowak, M. Dafotakis, C. Bangard, M. Brand, S. Shamay-Tsoory, O. A. Onur, and J. Kessler. Dissociating cognitive from affective theory of mind: a tms study. *cortex*, 46(6):769–780, 2010.

[52] D. Kemmerer. What modulates the mirror neuron system during action observation?: Multiple factors involving the action, the actor, the observer, the relationship between actor and observer, and the context. *Progress in Neurobiology*, 205:102128, 2021.

[53] W. Klimesch. Eeg-alpha rhythms and memory processes. *International Journal of psychophysiology*, 26(1-3):319–340, 1997.

[54] W. Klimesch. Eeg alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, 29(2-3):169–195, 1999.

[55] W. Klimesch. Alpha-band oscillations, attention, and controlled access to stored information. *Trends in cognitive sciences*, 16(12):606–617, 2012.

[56] W. Klimesch. The frequency architecture of brain and brain body oscillations: an analysis. *European Journal of Neuroscience*, 48(7):2431–2453, 2018.

[57] G. Knyazev, J. Y. Slobodskoj-Plusnin, and A. Bocharov. Event-related delta and theta synchronization during explicit and implicit emotion processing. *Neuroscience*, 164(4):1588–1600, 2009.

[58] G. G. Knyazev. Motivation, emotion, and their inhibitory control mirrored in brain oscillations. *Neuroscience & Biobehavioral Reviews*, 31(3):377–395, 2007.

[59] L. Koban, A. Ramamoorthy, and I. Konvalinka. Why do we fall into sync with others? interpersonal synchronization and the brain's optimization principle. *Social Neuroscience*, 14(1):1–9, 2019.

[60] T. Koike, H. C. Tanabe, S. Okazaki, E. Nakagawa, A. T. Sasaki, K. Shimada, S. K. Sugawara, H. K. Takahashi, K. Yoshihara, J. Bosch-Bayard, et al. Neural substrates of shared attention as social memory: a hyperscanning functional magnetic resonance imaging study. *NeuroImage*, 125:401–412, 2016.

[61] J.-P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela. Measuring phase synchrony in brain signals. *Human brain mapping*, 8(4):194–208, 1999.

[62] A. M. Leslie. Pretense and representation: The origins of" theory of mind.". *Psychological review*, 94(4):412, 1987.

[63] L. Li, H. Wang, H. Luo, X. Zhang, R. Zhang, and X. Li. Interpersonal neural synchronization during cooperative behavior of basketball players: a fnirs-based hyperscanning study. *Frontiers in human neuroscience*, 14:169, 2020.

[64] Y. Li, M. Chen, R. Zhang, and X. Li. Experiencing happiness together facilitates dyadic coordination through the enhanced interpersonal neural synchronization. *Social Cognitive and Affective Neuroscience*, 17(5):447–460, 2022.

[65] Y. Li, R. Chen, O. Turel, T. Feng, C.-Z. Zhu, and Q. He. Dyad sex composition effect on inter-brain synchronization in face-to-face cooperation. *Brain Imaging and Behavior*, 15:1667–1675, 2021.

[66] Y. Long, L. Zheng, H. Zhao, S. Zhou, Y. Zhai, and C. Lu. Interpersonal neural synchronization during interpersonal touch underlies affiliative pair bonding between romantic couples. *Cerebral Cortex*, 31(3):1647–1659, 2021.

[67] A. W. MacDonald, J. D. Cohen, V. A. Stenger, and C. S. Carter. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 288(5472):1835–1838, 2000.

[68] C. N. Macrae, O. K. Duffy, L. K. Miles, and J. Lawrence. A case of hand waving: Action synchrony and person perception. *Cognition*, 109(1):152–156, 2008.

[69] M. Maranesi, A. Livi, L. Fogassi, G. Rizzolatti, and L. Bonini. Mirror neuron activation prior to action observation in a predictable context. *Journal of Neuroscience*, 34(45):14827–14832, 2014.

[70] O. Mayo and I. Gordon. In and out of synchrony—behavioral and physiological dynamics of dyadic interpersonal coordination. *Psychophysiology*, 57(6):e13574, 2020.

[71] J. E. McGrath. *Groups: Interaction and performance*, volume 14. Prentice-Hall Englewood Cliffs, NJ, 1984.

[72] C. M. Michel and B. He. Eeg source localization. *Handbook of clinical neurology*, 160:85–101, 2019.

[73] J. G. Miller, P. Vrtička, X. Cui, S. Shrestha, S. H. Hosseini, J. M. Baker, and A. L. Reiss. Inter-brain synchrony in mother-child dyads during cooperation: an fnirs hyperscanning study. *Neuropsychologia*, 124:117–124, 2019.

[74] J. P. Mitchell, M. F. Mason, C. N. Macrae, and M. R. Banaji. Thinking about others: The neural substrates of social cognition. *Social neuroscience: People thinking about people*, pages 63–82, 2006.

[75] P. R. Montague, G. S. Berns, J. D. Cohen, S. M. McClure, G. Pagnoni, M. Dhamala, M. C. Wiest, I. Karpov, R. D. King, N. Apple, et al. Hyperscanning: simultaneous fmri during linked social interactions, 2002.

[76] H. S. Moon, H. Jiang, T. T. Vo, W. B. Jung, A. L. Vazquez, and S.-G. Kim. Contribution of excitatory and inhibitory neuronal activity to bold fmri. *Cerebral cortex*, 31(9):4053–4067, 2021.

[77] V. Müller, D. Perdikis, M. A. Mende, and U. Lindenberger. Interacting brains coming in sync through their minds: an interbrain neurofeedback study. *Annals of the New York Academy of Sciences*, 1500(1):48–68, 2021.

[78] S. Nam and S. Choi. Effect of threshold setting on neurofeedback training. *NeuroRegulation*, 7(3):107–107, 2020.

[79] T. Nguyen, S. Hoehl, and P. Vrtička. A guide to parent-child fnirs hyperscanning data processing and analysis. *Sensors*, 21(12):4075, 2021.

[80] T. Nguyen, H. Schleihauf, M. Kungl, E. Kayhan, S. Hoehl, and P. Vrtička. Interpersonal neural synchrony during father–child problem solving: an fnirs hyperscanning study. *Child development*, 92(4):e565–e580, 2021.

[81] T. Nozawa, Y. Sasaki, K. Sakaki, R. Yokoyama, and R. Kawashima. Interpersonal frontopolar neural synchronization in group communication: an exploration toward fnirs hyperscanning of natural interactions. *Neuroimage*, 133:484–497, 2016.

[82] A. Ossadtchi, T. Shamaeva, E. Okorokova, V. Moiseeva, and M. A. Lebedev. Neurofeedback learning modifies the incidence rate of alpha spindles, but not their duration and amplitude. *Scientific reports*, 7(1):3772, 2017.

[83] Y. Pan, X. Cheng, Z. Zhang, X. Li, and Y. Hu. Cooperation in lovers: an f nirs-based hyperscanning study. *Human brain mapping*, 38(2):831–841, 2017.

[84] J. Park, J. Shin, and J. Jeong. Inter-brain synchrony levels according to task execution modes and difficulty levels: an fnirs/gsr study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:194–204, 2022.

[85] R. Paul, J. R. Drake, and H. Liang. Global virtual team performance: The effect of coordination effectiveness, trust, and team cohesion. *IEEE Transactions on Professional Communication*, 59(3):186–202, 2016.

[86] P. Pinti, I. Tachtsidis, A. Hamilton, J. Hirsch, C. Aichelburg, S. Gilbert, and P. W. Burgess. The present and future use of functional near-infrared spectroscopy (fnirs) for cognitive neuroscience. *Annals of the New York Academy of Sciences*, 1464(1):5–29, 2020.

[87] K. Preckel, P. Kanske, and T. Singer. On the interaction of social affect and cognition: empathy, compassion and theory of mind. *Current Opinion in Behavioral Sciences*, 19:1–6, 2018.

[88] M. Proudfoot, M. W. Woolrich, A. C. Nobre, and M. R. Turner. Magnetoencephalography. *Practical neurology*, 14(5):336–343, 2014.

[89] V. Reindl, C. Gerloff, W. Scharke, and K. Konrad. Brain-to-brain synchrony in parent-child dyads and the relationship with emotion regulation revealed by fnirs-based hyperscanning. *NeuroImage*, 178:493–502, 2018.

[90] D. A. Reinero, S. Dikker, and J. J. Van Bavel. Inter-brain synchrony in teams predicts collective performance. *Social cognitive and affective neuroscience*, 16(1-2):43–57, 2021.

[91] N. L. REMPEL-CLOWER. Role of orbitofrontal cortex connections in emotion. *Annals of the New York Academy of Sciences*, 1121(1):72–86, 2007.

[92] R. Rico, M. Sánchez-Manzanares, F. Gil, and C. Gibson. Team implicit coordination processes: A team knowledge–based approach. *Academy of management review*, 33(1):163–184, 2008.

[93] J. Rogala, K. Jurewicz, K. Paluch, E. Kublik, R. Cetnarski, and A. Wróbel. The do's and don'ts of neurofeedback training: a review of the controlled studies using healthy adults. *Frontiers in human neuroscience*, 10:301, 2016.

[94] E. rol Başar, M. artin Schürmann, T. amer Demiralp, C. anan Başar-Eroglu, and A. hmet Ademoglu. Event-related oscillations are 'real brain responses'—wavelet analysis and new strategies. *International Journal of Psychophysiology*, 39(2-3):91–127, 2001.

[95] M. Saber. Removing powerline interference from eeg signal using optimized fir filters. *J. Artif. Intell. Metaheuristics*, 1(1):8–19, 2022.

[96] E. Salas, M. A. Rosen, C. S. Burke, and G. F. Goodwin. The wisdom of collectives in organizations: An update of the teamwork competencies. In *Team effectiveness in complex organizations*, pages 73–114. Routledge, 2008.

[97] E. Salas, M. L. Shuffler, A. L. Thayer, W. L. Bedwell, and E. H. Lazzara. Understanding and improving teamwork in organizations: A scientifically based practical guide. *Human resource management*, 54(4):599–622, 2015.

[98] E. Salas, S. I. Tannenbaum, K. Kraiger, and K. A. Smith-Jentsch. The science of training and development in organizations: What matters in practice. *Psychological science in the public interest*, 13(2):74–101, 2012.

[99] L. Schlaffke, S. Lissek, M. Lenz, G. Juckel, T. Schultz, M. Tegenthoff, T. Schmidt-Wilcke, and M. Brüne. Shared and nonshared neural networks of cognitive and affective theory-of-mind: A neuroimaging study using cartoon picture stories. *Human Brain Mapping*, 36(1):29–39, 2015.

[100] N. Sebanz, H. Bekkering, and G. Knoblich. Joint action: bodies and minds moving together. *Trends in cognitive sciences*, 10(2):70–76, 2006.

[101] M. Siems and M. Siegel. Dissociated neuronal phase-and amplitude-coupling patterns in the human brain. *NeuroImage*, 209:116538, 2020.

[102] R. Sitaram, T. Ros, L. Stoeckel, S. Haller, F. Scharnowski, J. Lewis-Peacock, N. Weiskopf, M. L. Blefari, M. Rana, E. Oblak, et al. Closed-loop brain training: the science of neurofeedback. *Nature Reviews Neuroscience*, 18(2):86–100, 2017.

[103] S. Siuly, Y. Li, and Y. Zhang. Eeg signal analysis and classification. *IEEE Trans Neural Syst Rehabilit Eng*, 11:141–144, 2016.

[104] J. L. Soler-Dominguez and C. Gonzalez. Using eeg and gamified neurofeedback environments to improve esports performance: project neuroprotrainer,". In *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2021)*, volume 1, pages 278–283, 2021.

[105] B. Sorger, F. Scharnowski, D. E. Linden, M. Hampson, and K. D. Young. Control freaks: Towards optimal selection of control conditions for fmri neurofeedback studies. *Neuroimage*, 186:256–265, 2019.

[106] A. Soutschek, C. C. Ruff, T. Strombach, T. Kalenscher, and P. N. Tobler. Brain stimulation reveals crucial role of overcoming self-centeredness in self-control. *Science advances*, 2(10):e1600992, 2016.

[107] C. J. Stam, G. Nolte, and A. Daffertshofer. Phase lag index: assessment of functional connectivity from multi channel eeg and meg with diminished bias from common sources. *Human brain mapping*, 28(11):1178–1193, 2007.

[108] M. B. Sterman and T. Egner. Foundation and practice of neurofeedback for the treatment of epilepsy. *Applied psychophysiology and biofeedback*, 31:21–35, 2006.

[109] C. Szymanski, A. Pesquita, A. A. Brennan, D. Perdikis, J. T. Enns, T. R. Brick, V. Müller, and U. Lindenberger. Teams on the same wavelength perform better: Inter-brain phase synchronization constitutes a neural substrate for social facilitation. *Neuroimage*, 152:425–436, 2017.

[110] J. S. Thomas, A. C. Loignon, D. J. Woehr, M. L. Loughry, and M. W. Ohland. Dyadic viability in project teams: The impact of liking, competence, and task interdependence. *Journal of Business and Psychology*, 35:573–591, 2020.

[111] M. Tomasello and M. Carpenter. Shared intentionality. *Developmental science*, 10(1):121–125, 2007.

[112] J. Toppi, G. Borghini, M. Petti, E. J. He, V. De Giusti, B. He, L. Astolfi, and F. Babiloni. Investigating cooperative behavior in ecological settings: an eeg hyperscanning study. *PloS one*, 11(4):e0154236, 2016.

[113] W. Tschacher, G. M. Rees, and F. Ramseyer. Nonverbal synchrony and affect in dyadic interactions. *Frontiers in psychology*, 5:1323, 2014.

[114] K. Tzambazis and C. Stough. Alcohol impairs speed of information processing and simple and choice reaction time and differentially impairs higher-order cognitive abilities. *Alcohol and Alcoholism*, 35(2):197–201, 2000.

[115] J. Van Den Berg and G. Neely. Performance on a simple reaction time task while sleep deprived. *Perceptual and Motor Skills*, 102(2):589–599, 2006.

[116] D. van Son, W. van der Does, G. P. Band, and P. Putman. Eeg theta/beta ratio neurofeedback training in healthy females. *Applied Psychophysiology and Biofeedback*, 45:195–210, 2020.

[117] C. Vesper, E. Abramova, J. Bütepage, F. Ciardo, B. Crossey, A. Effenberg, D. Hristova, A. Karlinsky, L. McEllin, S. R. Nijssen, et al. Joint action: Mental representations, shared information and general mechanisms for coordinating with others. *Frontiers in psychology*, 7:2039, 2017.

[118] C. Vesper and M. J. Richardson. Strategic communication and behavioral coupling in asymmetric joint action. *Experimental brain research*, 232:2945–2956, 2014.

[119] C. Wang, H. Li, L. Jia, F. Li, and J. Wang. Theta band behavioral fluctuations synchronized interpersonally during cooperation. *Psychonomic Bulletin & Review*, 27:563–570, 2020.

[120] C. Wang, T. Zhang, Z. Shan, J. Liu, D. Yuan, and X. Li. Dynamic interpersonal neural synchronization underlying pain-induced cooperation in females. *Human Brain Mapping*, 40(11):3222–3232, 2019.

[121] M.-Y. Wang, P. Luan, J. Zhang, Y.-T. Xiang, H. Niu, and Z. Yuan. Concurrent mapping of brain activation from multiple subjects during social interaction by hyperscanning: a mini-review. *Quantitative imaging in medicine and surgery*, 8(8):819, 2018.

[122] Q. Wang, Z. Han, X. Hu, S. Feng, H. Wang, T. Liu, and L. Yi. Autism symptoms modulate interpersonal neural synchronization in children with autism spectrum disorder in cooperative interactions. *Brain Topography*, 33:112–122, 2020.

[123] L. A. Weber, T. Ethofer, and A.-C. Ehlis. Predictors of neurofeedback training outcome: A systematic review. *NeuroImage: Clinical*, 27:102301, 2020.

[124] V. Wikström, K. Saarikivi, M. Falcon, T. Makkonen, S. Martikainen, V. Putkinen, B. U. Cowley, and M. Tervaniemi. Inter-brain synchronization occurs without physical co-presence during cooperative online gaming. *Neuropsychologia*, 174:108316, 2022.

[125] M. K. Wittmann, N. Kolling, N. S. Faber, J. Scholl, N. Nelissen, and M. F. Rushworth. Self-other mergence in the frontal cortex during cooperation and competition. *Neuron*, 91(2):482–493, 2016.

[126] R. C. Wolf, C. L. Philippi, J. C. Motzkin, M. K. Baskaya, and M. Koenigs. Ventromedial prefrontal cortex mediates visual attention during facial emotion recognition. *Brain*, 137(6):1772–1780, 2014.

[127] L. Yang, M. Li, L. Yang, H. Wang, H. Wan, and Z. Shang. Functional connectivity changes in the intra-and inter-brain during the construction of the multi-brain network of pigeons. *Brain Research Bulletin*, 161:147–157, 2020.

[128] K. Yun. On the same wavelength: Face-to-face communication increases interpersonal neural synchronization. *Journal of Neuroscience*, 33(12):5081–5082, 2013.

[129] K. Yun, K. Watanabe, and S. Shimojo. Interpersonal body and neural synchronization as a marker of implicit social interaction. *Scientific reports*, 2(1):959, 2012.

[130] C. Zhou, X. Cheng, C. Liu, and P. Li. Interpersonal coordination enhances brain-to-brain synchronization and influences responsibility attribution and reward allocation in social cooperation. *NeuroImage*, 252:119028, 2022.

[131] S. Zhou, Y. Zhang, Y. Fu, L. Wu, X. Li, N. Zhu, D. Li, and M. Zhang. The effect of task performance and partnership on interpersonal brain synchrony during cooperation. *Brain Sciences*, 12(5):635, 2022.

[132] X. Zhou, Y. Pan, R. Zhang, L. Bei, and X. Li. Mortality threat mitigates interpersonal competition: an eeg-based hyperscanning study. *Social cognitive and affective neuroscience*, 16(6):621–631, 2021.

[133] M. Zimmermann, K. Schultz-Nielsen, G. Dumas, and I. Konvalinka. Arbitrary methodological decisions skew inter-brain synchronization estimates in hyperscanning-eeg studies. 2023.

[134] V. Zotev, A. Mayeli, M. Misaki, and J. Bodurka. Emotion self-regulation training in major depressive disorder using simultaneous real-time fmri and eeg neurofeedback. *NeuroImage: Clinical*, 27:102331, 2020.

Appendix A. Questionnaire

1. Are you diagnosed and/or currently suffering from any mental disorder from this list:

- Panic disorder, (severe) obsessive-compulsive disorder
- Depression, bipolar disorder, dysthymia
- Borderline Personality Disorder, Antisocial Personality Disorder
- Psychotic disorder (e.g. schizophrenia)
- Post-Traumatic Stress Disorder

[] Yes

[] No

2. Are you diagnosed and/or currently suffering from any neurological disorder that could influence learning, social interactions or performance in computer tasks? (e.g. Alzheimer's, epilepsy, Parkinson, Dimentia)

[] Yes

[] No

3. Do you take any medication that is known to impair reaction time?

[] Yes

[] No

4. Do you have any visual impairments (e.g. blind in one eye) or uncorrected near- or farsightedness?

[] Yes

[] No

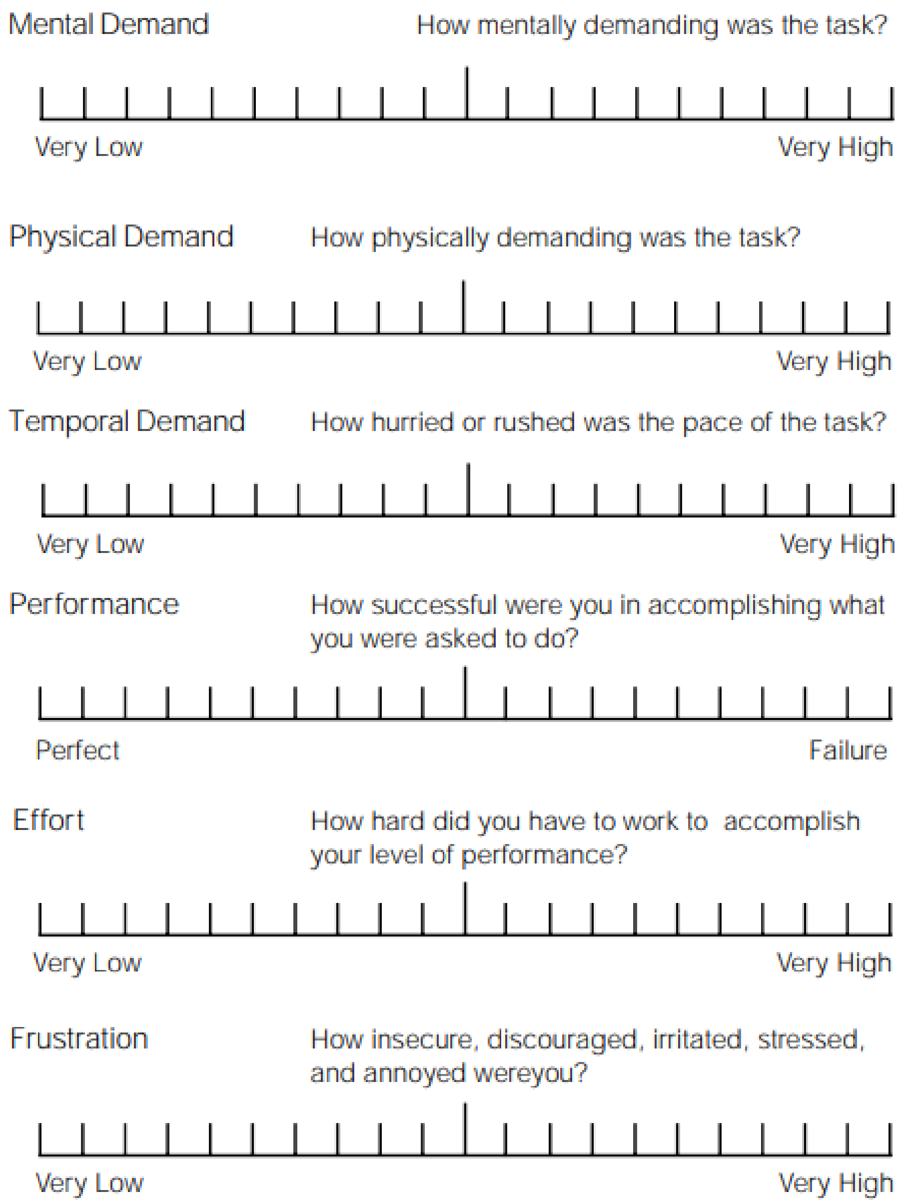
5. Do you have some type of condition that could impair your ability to quickly press a key on a keyboard?

[] Yes

[] No

Figure A.1: Questionnaire which was given to the participants before participation to access whether participants met any exclusion criteria.

Appendix B. NASA Task Load Index



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

Figure B.1: The NASA Task Load Index.

Appendix C. Change in Threshold Values Alpha-NFT

Mean Baseline amplitude	Initial Threshold (Percentages of amplitude)
0.005 > amplitude > 0	100%
Amplitude > 0.005	90%
Amplitude > 0.01	85%
Amplitude > 0.05	80%

Figure C.1: How the initial threshold is calculated during alpha neurofeedback training. Left: the measured mean amplitude. Right: the calculated initial threshold value as a percentage of the measured mean amplitude.
.

Appendix D. Change in Threshold Values INS-NFT

Mean Baseline CCorr	**Initial Threshold (Percentages of CCorr)**
0.01 > CCorr > 0	100%
CCorr > 0.01	95%
CCorr > 0.1	92%

Figure D.1: How the initial threshold is calculated during INS neurofeedback training. Left: the measured mean CCorr. Right: the calculated initial threshold value as a percentage of the measured mean CCorr.
.

Appendix E. Shapiro-Wilk Normality Test Results H1

	W	p	Significance
H1a / Delta	.936	$2.21*10^{-16}$	p < a
H1a / Theta	.951	$2.20*10^{-16}$	p < a
H1b / Delta	.982	$1.43*10^{-4}$	p < a
H1b / Theta	.861	$2.20*10^{-16}$	p < a
H1c / Delta	.957	$1.23*10^{-8}$	p < a
H1c / Theta	.990	$2.05*10^{-2}$	p < a
H1c Assumption / Delta	.893	$6.29*10^{-15}$	p < a
H1c Assumption / Theta	.975	$9.96*10^{-6}$	p < a
H1d / Delta	.936	$2.20*10^{-16}$	p < a
H1d / Theta	.953	$2.20*10^{-16}$	p < a

Figure E.1: Shapiro-Wilk Normality Test results for Hypothesis 1.

.

Appendix F. Shapiro-Wilk Normality Test Results H2

	W	p	Significance
H2b	.861	$9.50{*}10^{-14}$	p < a
H2b Assumption	.867	$2.12{*}10^{-13}$	p < a
H2d	.880	$1.03{*}10^{-12}$	p < a
H2d Assumption	.853	$3.98{*}10^{-14}$	p < a
H2e / Delta	.956	$3.00{*}10^{-9}$	p < a
H2e / Theta	.940	$3.06{*}10^{-11}$	p < a

Figure F.1: Shapiro-Wilk Normality Test results for Hypothesis 2.

.

Appendix G. Individual Neurofeedback Training Results

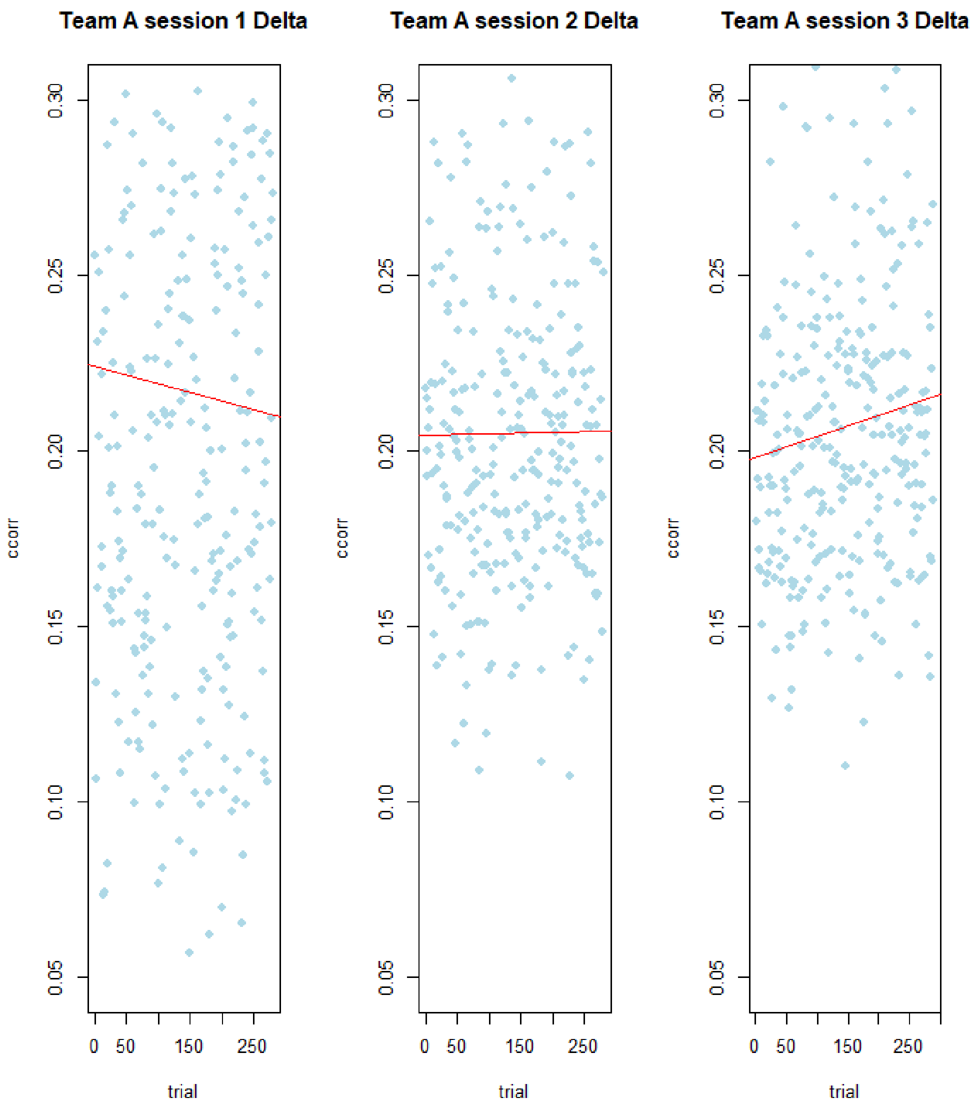


Figure G.1: Delta Neurofeedback Training Results for team A of the experimental group (y-axis: Delta-CCorr, x-axis: trial number). From left to right: results for session 1, session 2 and session 3.
.

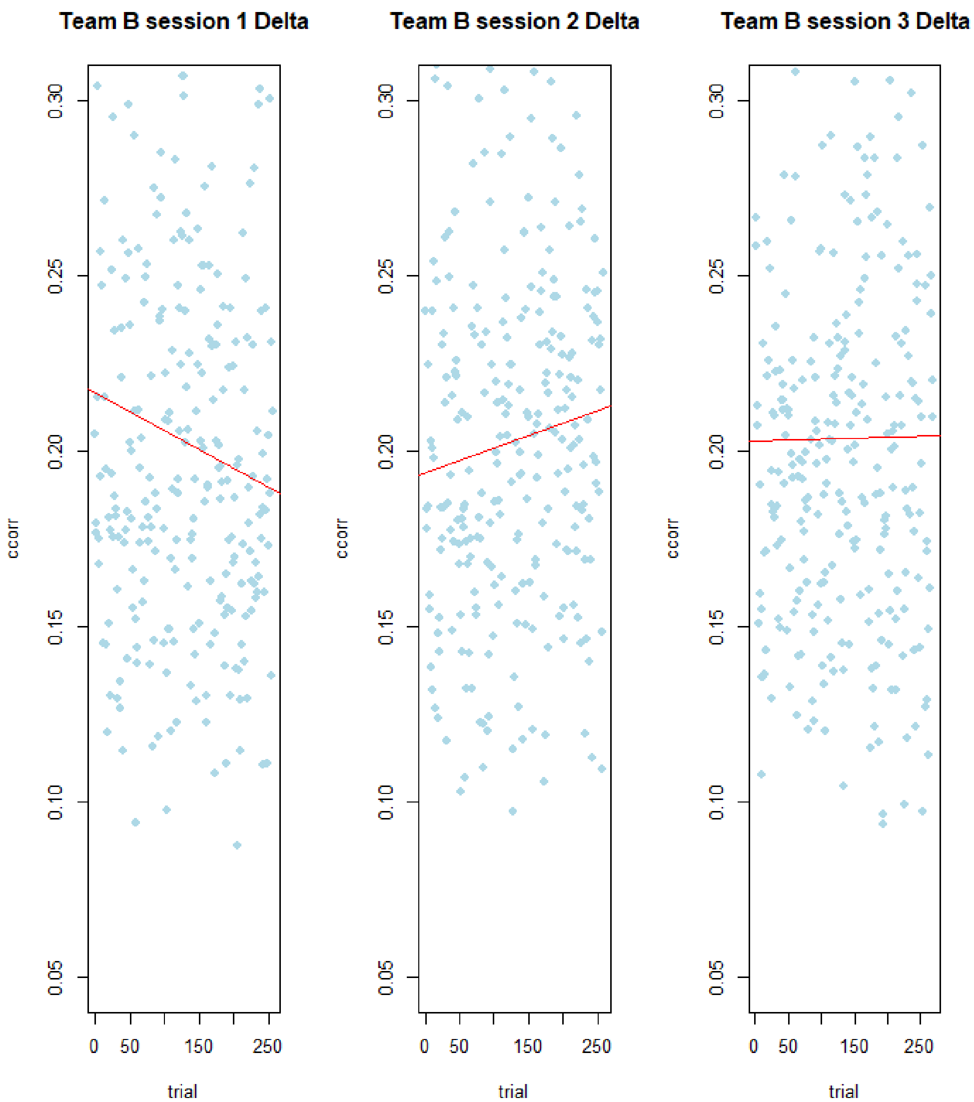


Figure G.2: Delta Neurofeedback Training Results for team B of the experimental group (y-axis: Delta-CCorr, x-axis: trial number). From left to right: results for session 1, session 2 and session 3.

.

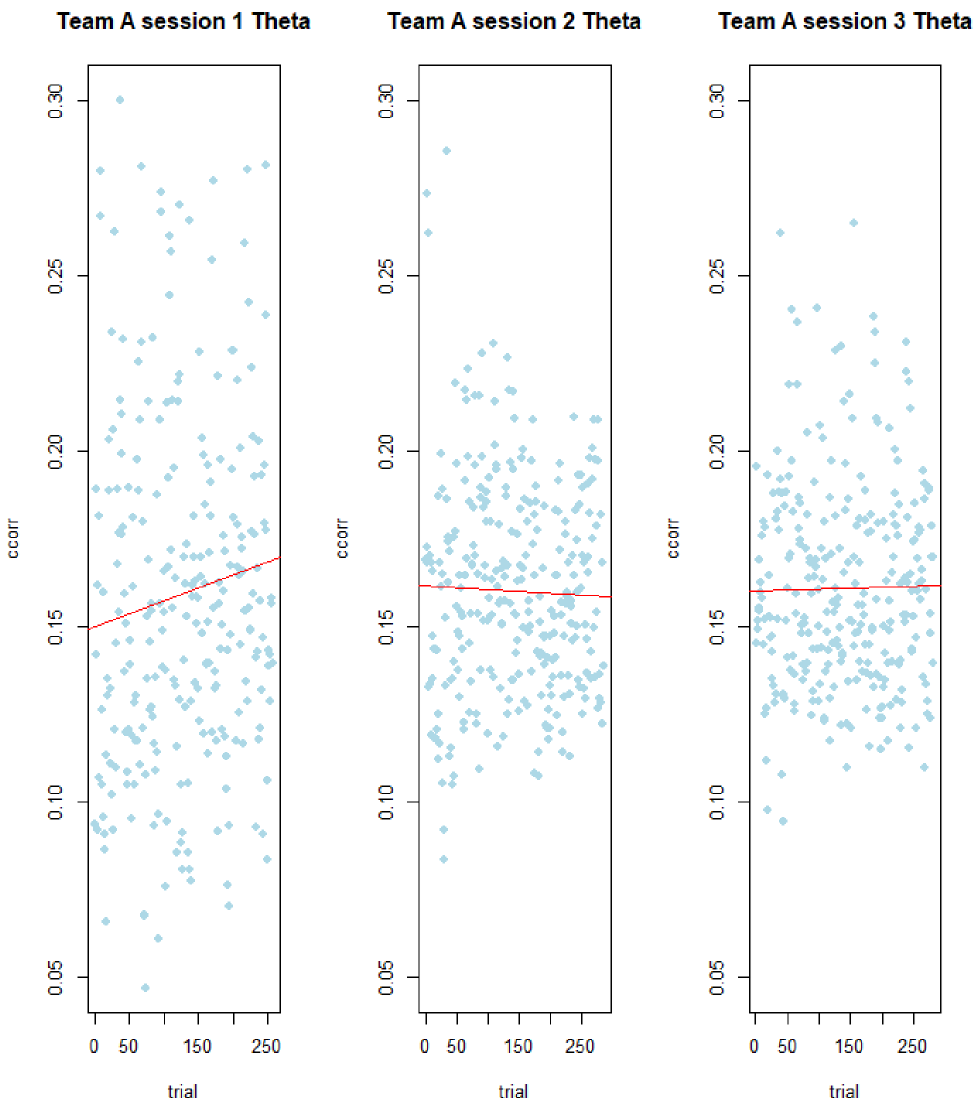


Figure G.3: Theta Neurofeedback Training Results for team A of the experimental group (y-axis: Theta-CCorr, x-axis: trial number). From left to right: results for session 1, session 2 and session 3.

.

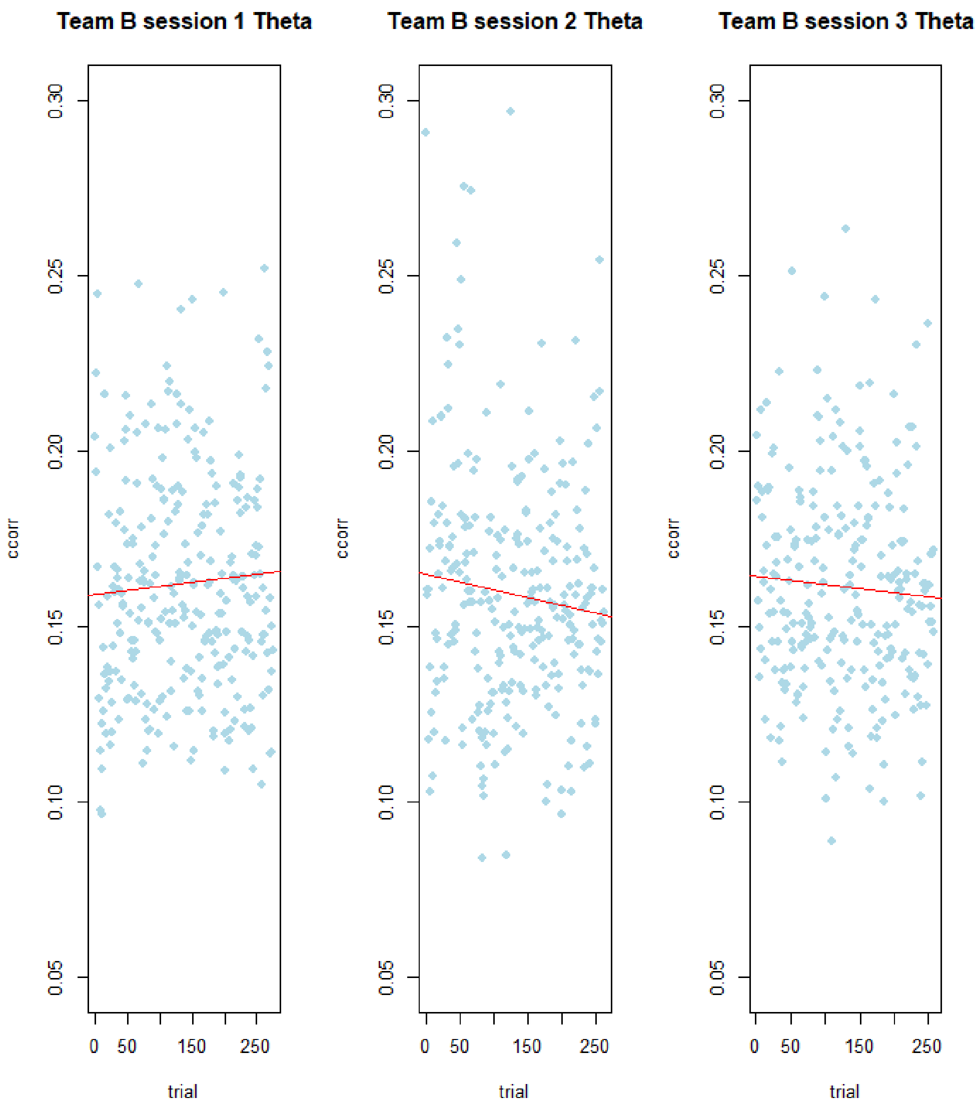


Figure G.4: Theta Neurofeedback Training Results for team B of the experimental group (y-axis: Theta-CCorr, x-axis: trial number). From left to right: results for session 1, session 2 and session 3.

.

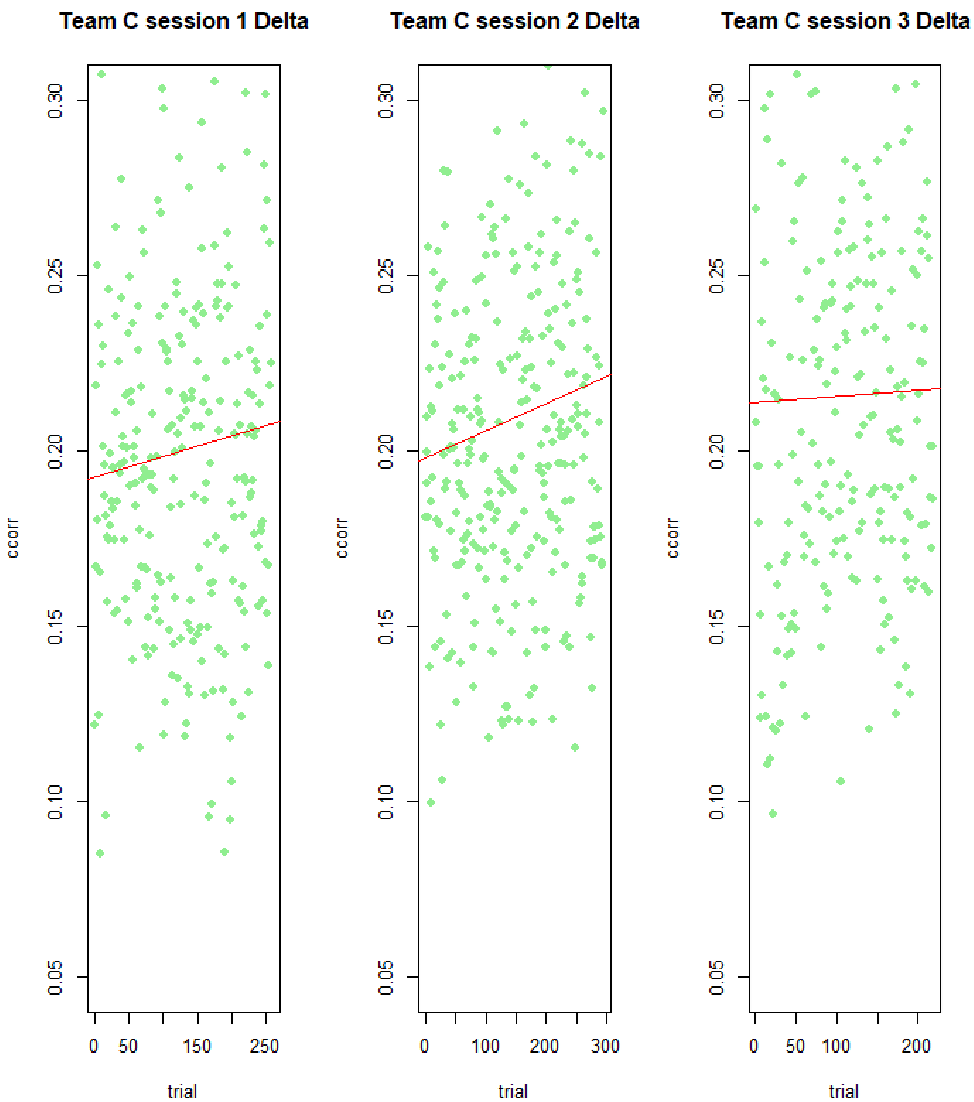


Figure G.5: Delta Neurofeedback Training Results for team C of the control group (y-axis: Delta-CCorr, x-axis: trial number). From left to right: results for session 1, session 2 and session 3.

.

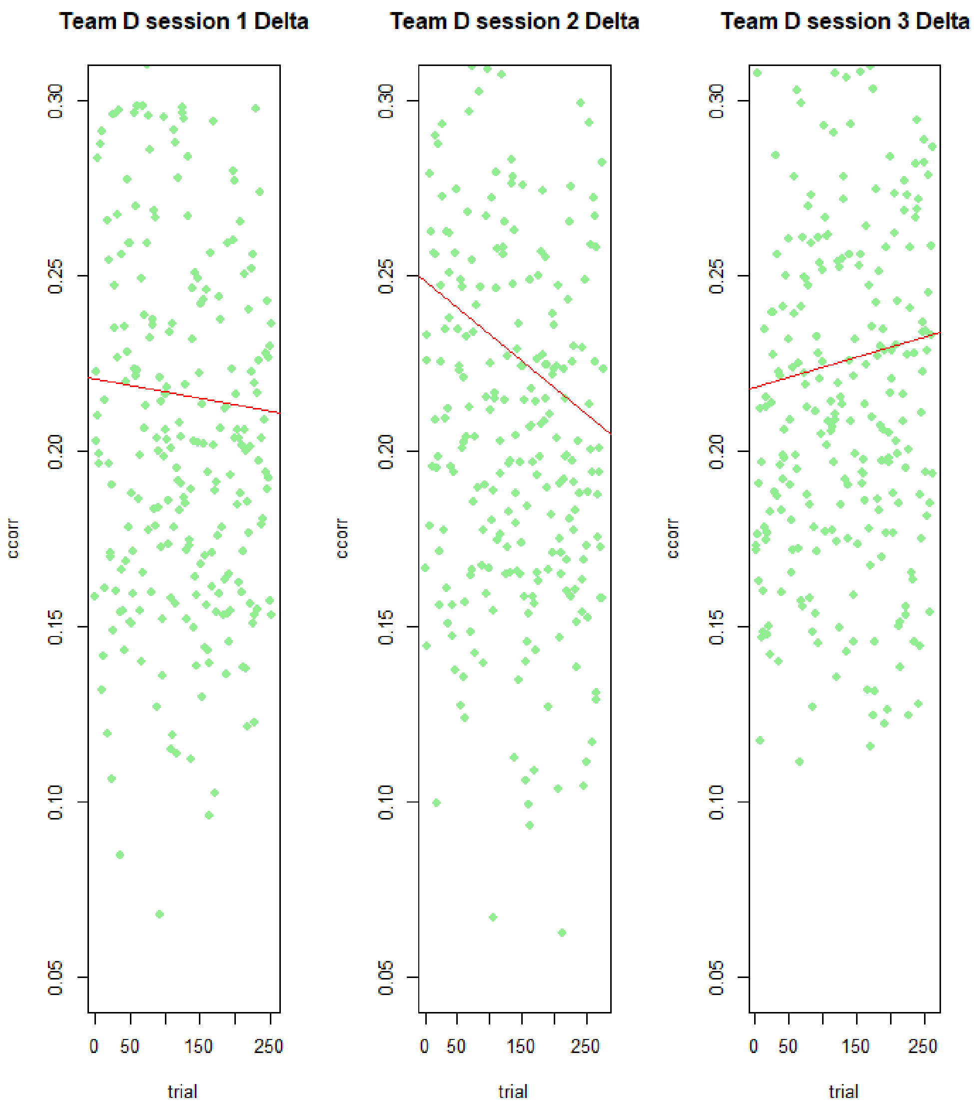


Figure G.6: Delta Neurofeedback Training Results for team D of the control group (y-axis: Delta-CCorr, x-axis: trial number). From left to right: results for session 1, session 2 and session 3.

.

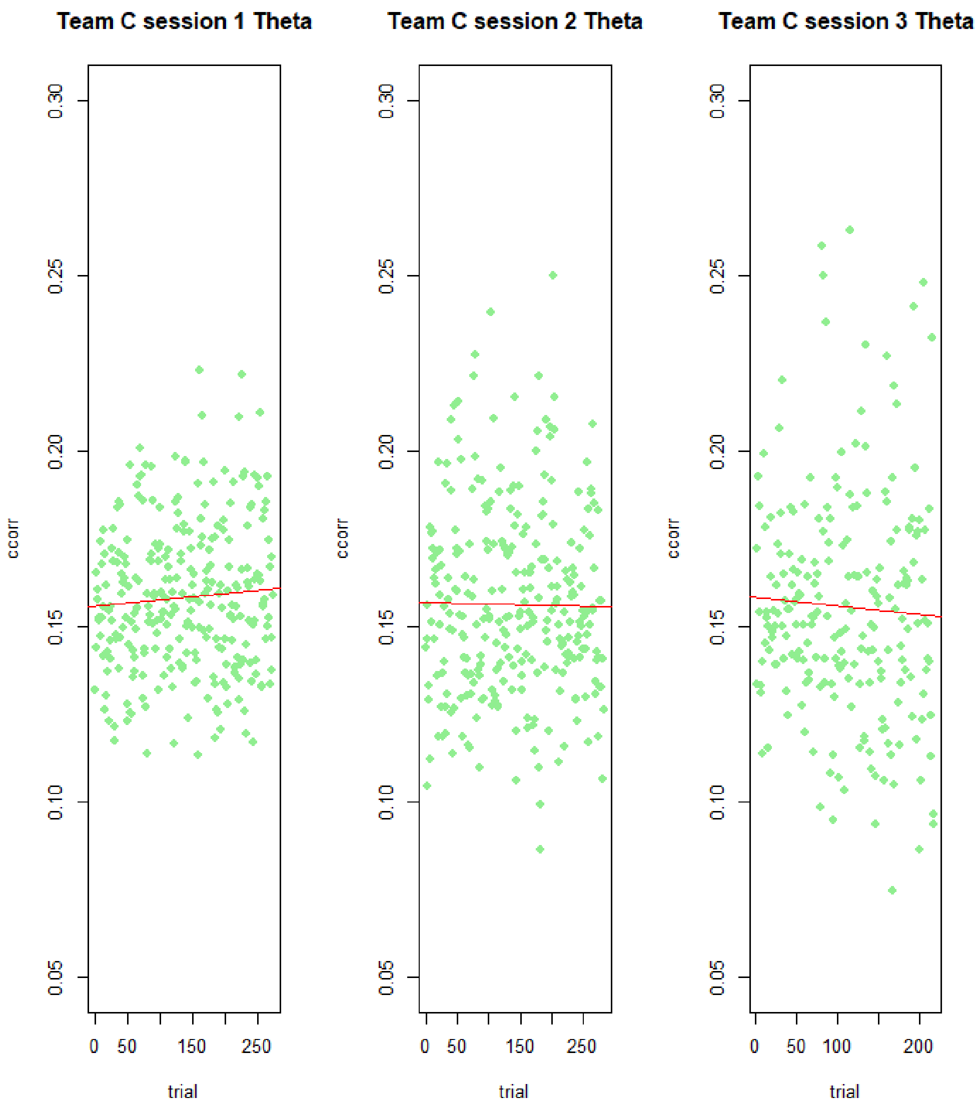


Figure G.7: Theta Neurofeedback Training Results for team C of the control group (y-axis: Theta-CCorr, x-axis: trial number). From left to right: results for session 1, session 2 and session 3.

.

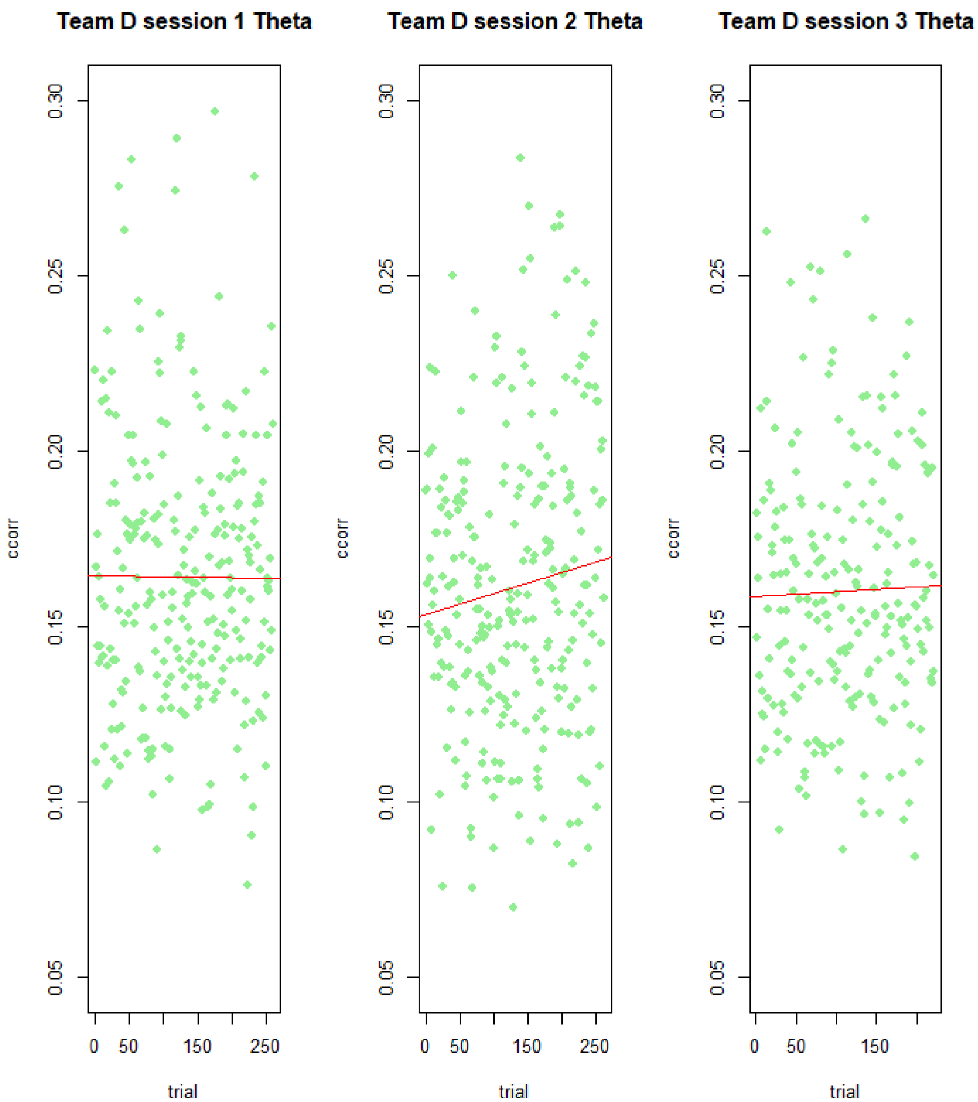


Figure G.8: Theta Neurofeedback Training Results for team D of the control group (y-axis: Theta-CCorr, x-axis: trial number). From left to right: results for session 1, session 2 and session 3.

.

Appendix H. Baseline results (Session 1 INS-NFT)

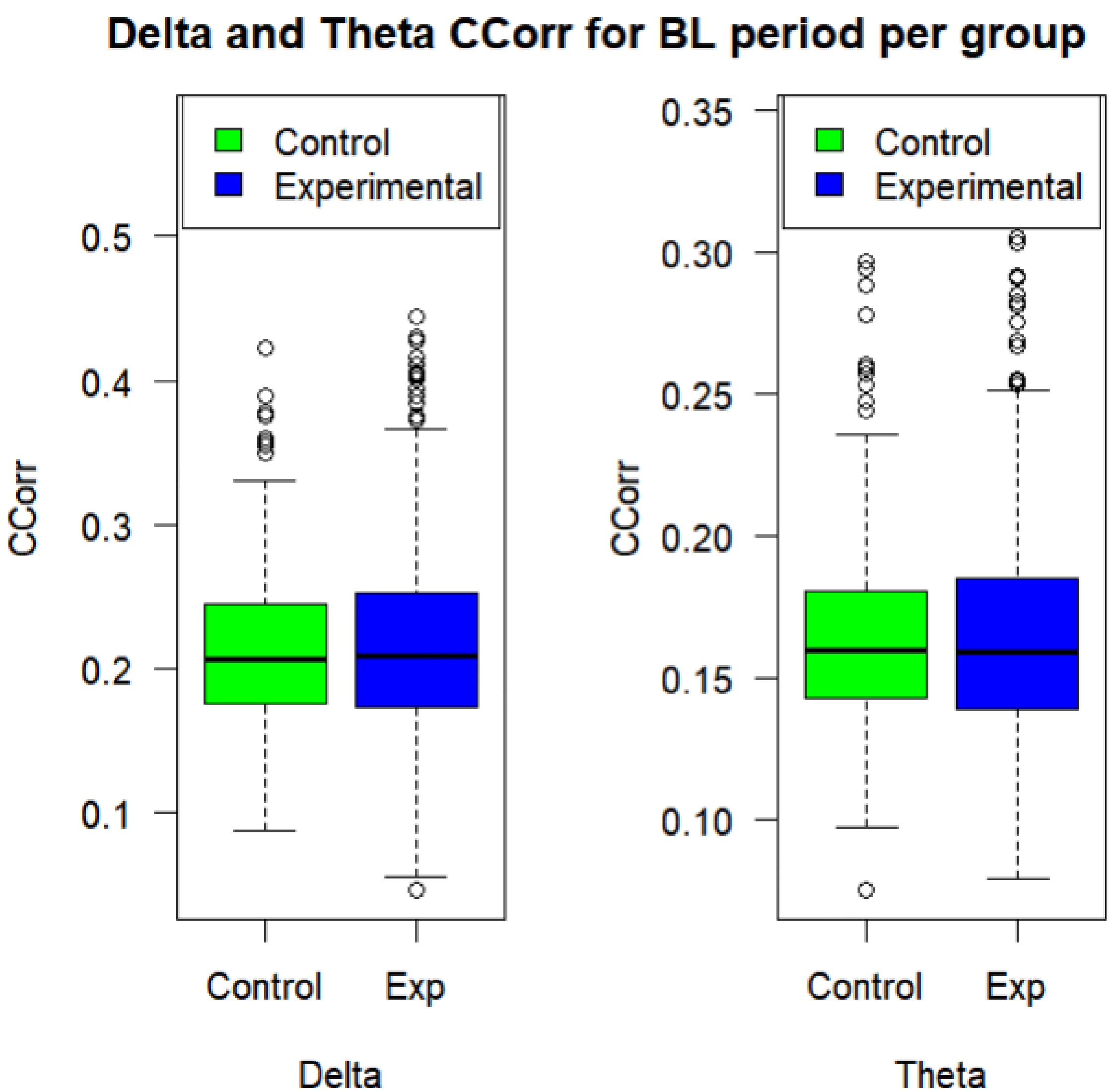


Figure H.1: Results in for baseline period of Neurofeedback Training session 1 (y-axis: CCorr, x-axis: condition, i.e. control and experimental, respectively). Left: Delta-CCorr on the y-axis, Right: Theta-CCorr on the y-axis.

.

Appendix I. Matrices Control Condition

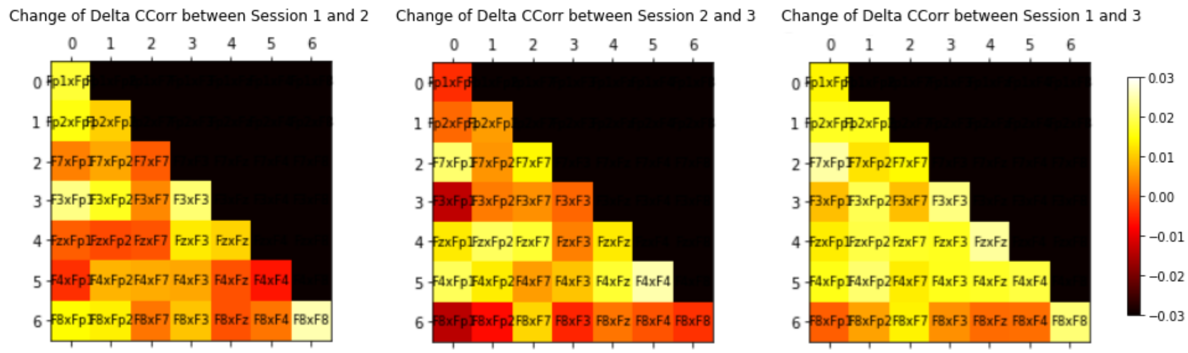


Figure I.1: Change in delta-CCorr between two sessions for the control condition. left: The change in delta-CCorr of session 2 compared to session 1. Middle: The change in delta-CCorr of session 3 compared to session 2. Right: The change in delta-CCorr of session 3 compared to session 1. Lighter colors denote a higher CCorr value.

.

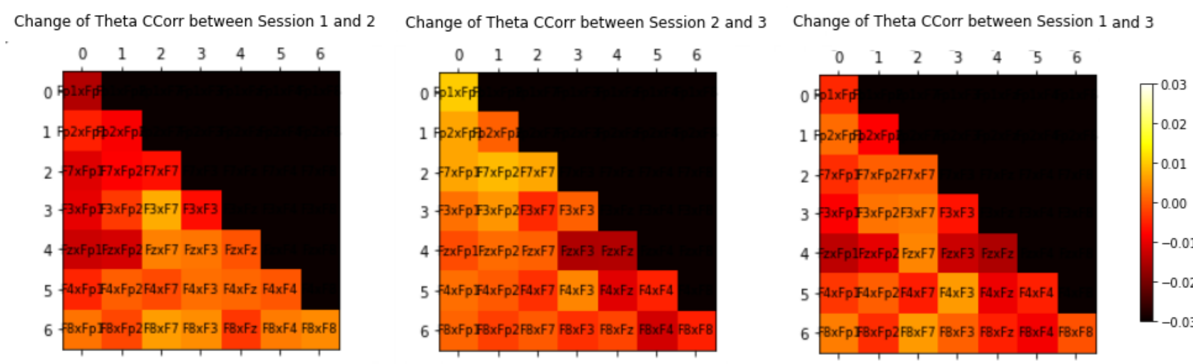


Figure I.2: Change in theta-CCorr between two sessions for the control condition. left: The change in theta-CCorr of session 2 compared to session 1. Middle: The change in theta-CCorr of session 3 compared to session 2. Right: The change in theta-CCorr of session 3 compared to session 1. Lighter colors denote a higher CCorr value.

.

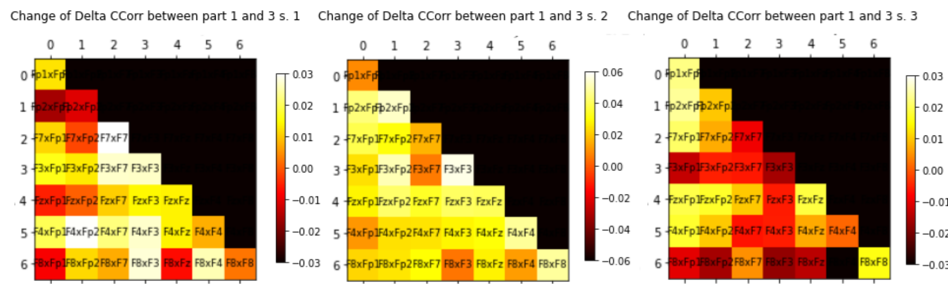


Figure I.3: Change in delta-CCorr between part 1 and part 3 of each NFT session for the control condition. Left: The change in delta-CCorr between part 3 of session 1 and part 1 of session 1. Middle: The change in delta-CCorr between part 3 of session 2 and part 1 of session 2. Right: The change in delta-CCorr between part 3 of session 3 and part 1 of session 3. Lighter colors denote a higher CCorr value.

.

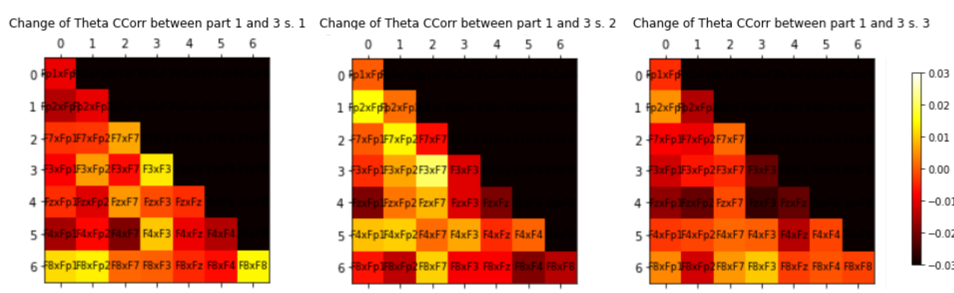


Figure I.4: Change in theta-CCorr between part 1 and part 3 of each NFT session for the control condition. Left: The change in theta-CCorr between part 3 of session 1 and part 1 of session 1. Middle: The change in theta-CCorr between part 3 of session 2 and part 1 of session 2. Right: The change in theta-CCorr between part 3 of session 3 and part 1 of session 3. Lighter colors denote a higher CCorr value.

.

Appendix J. Result NASA Task Load Index

	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration
Team A S.1 P.1	14	6	6	14	16	10
Team A S.1 P.2	13	2	4	8	17	9
Team B S.1 P.1	6	14	5	14	16	11
Team B S.1 P.2	8	0	5	7	13	13
Team A S.2 P.1	6	3	6	8	11	8
Team A S.2 P.2	17	5	3	8	17	2
Team B S.2 P.1	13	2	3	12	14	7
Team B S.2 P.2	13	4	2	13	15	13
Team A S.3 P.1	8	4	4	9	6	9
Team A S.3 P.2	17	5	4	7	5	13
Team B S.3 P.1	10	1	2	15	13	2
Team B S.3 P.2	11	3	4	12	14	12
Sum Experimental	**136**	**49**	**48**	**127**	**157**	**109**
Team C S.1 P.1	5	5	3	7	12	8
Team C S.1 P.2	19	1	8	18	16	9
Team D S.1 P.1	14	3	3	8	12	11
Team D S.1 P.2	11	2	9	14	14	5
Team C S.2 P.1	7	5	7	7	12	7
Team C S.2 P.2	18	1	6	16	16	8
Team D S.2 P.1	15	1	2	9	12	8
Team D S.2 P.2	13	10	4	5	14	6
Team C S.3 P.1	18	17	11	10	18	8
Team C S.3 P.2	18	3	17	17	17	19
Team D S.3 P.1	3	2	2	5	8	4
Team D S.3 P.2	3	3	2	7	6	3
Sum Control	**144**	**53**	**74**	**123**	**157**	**96**

Figure J.1: Results of the NASA Task Load Index for the experimental group (upper 13 rows) and the control group (bottom 13 rows).

Appendix K. Comparison Baseline and Experimental results

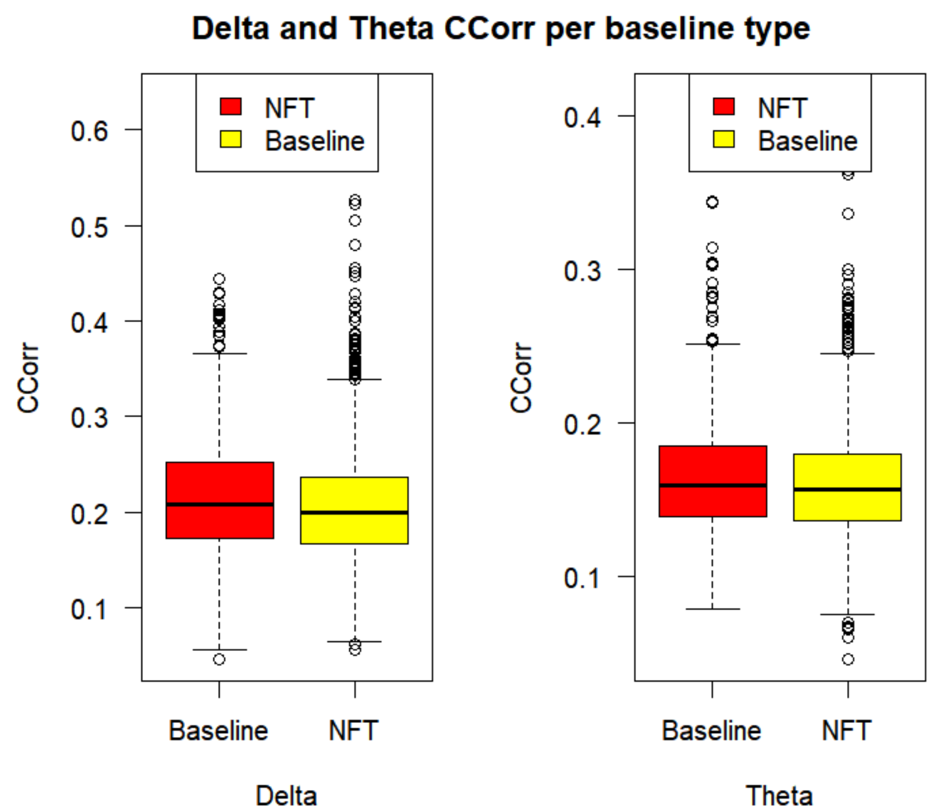


Figure L.1: left: Box plot of delta CCorr for the baseline period of session 1 (yellow) and during INS-NFT of session 3 part 3 (red). Right: Box plot of theta CCorr for the baseline period of session 1 (yellow) and during INS-NFT of session 3 part 3 (red).

Appendix L. Neurofeedback Training Results for Individual Teams

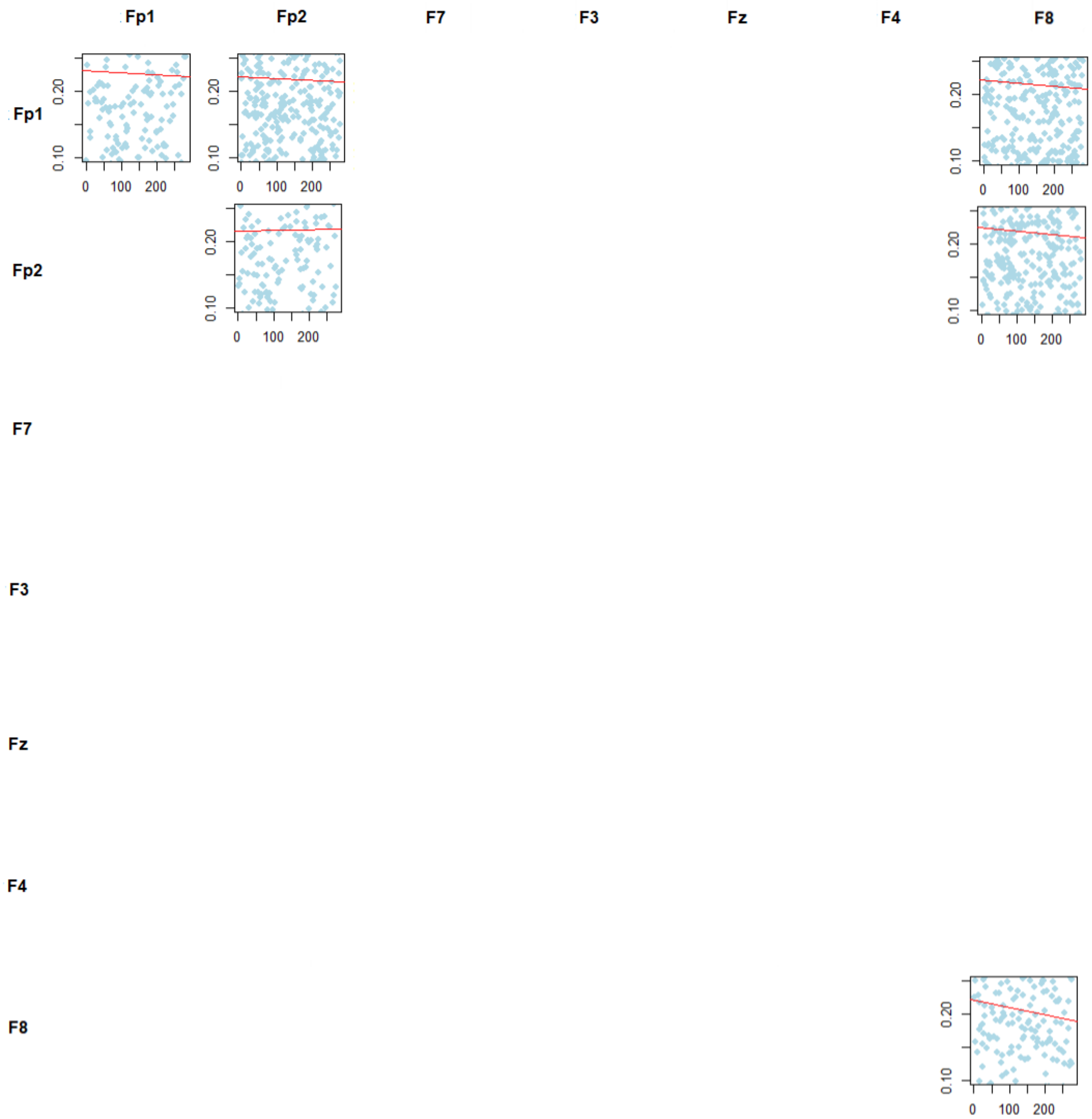


Figure L.1: Delta-CCorr per electrode combination during neurofeedback training session 1 for team A. X-axis: trial number. Y-axis: Delta-CCorr.

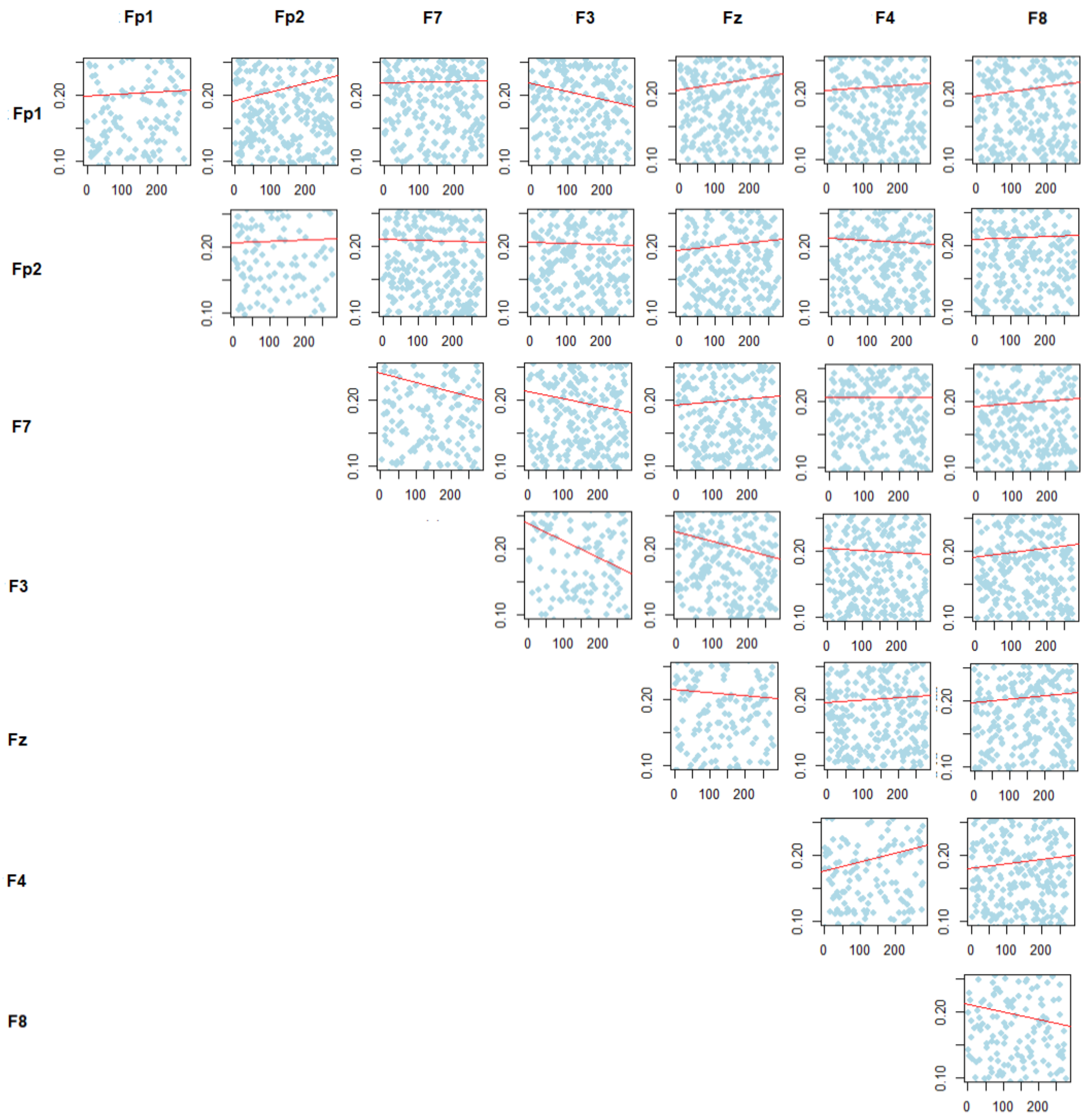


Figure L.2: Delta-CCorr per electrode combination during neurofeedback training session 2 for team A. X-axis: trial number. Y-axis: Delta-CCorr.

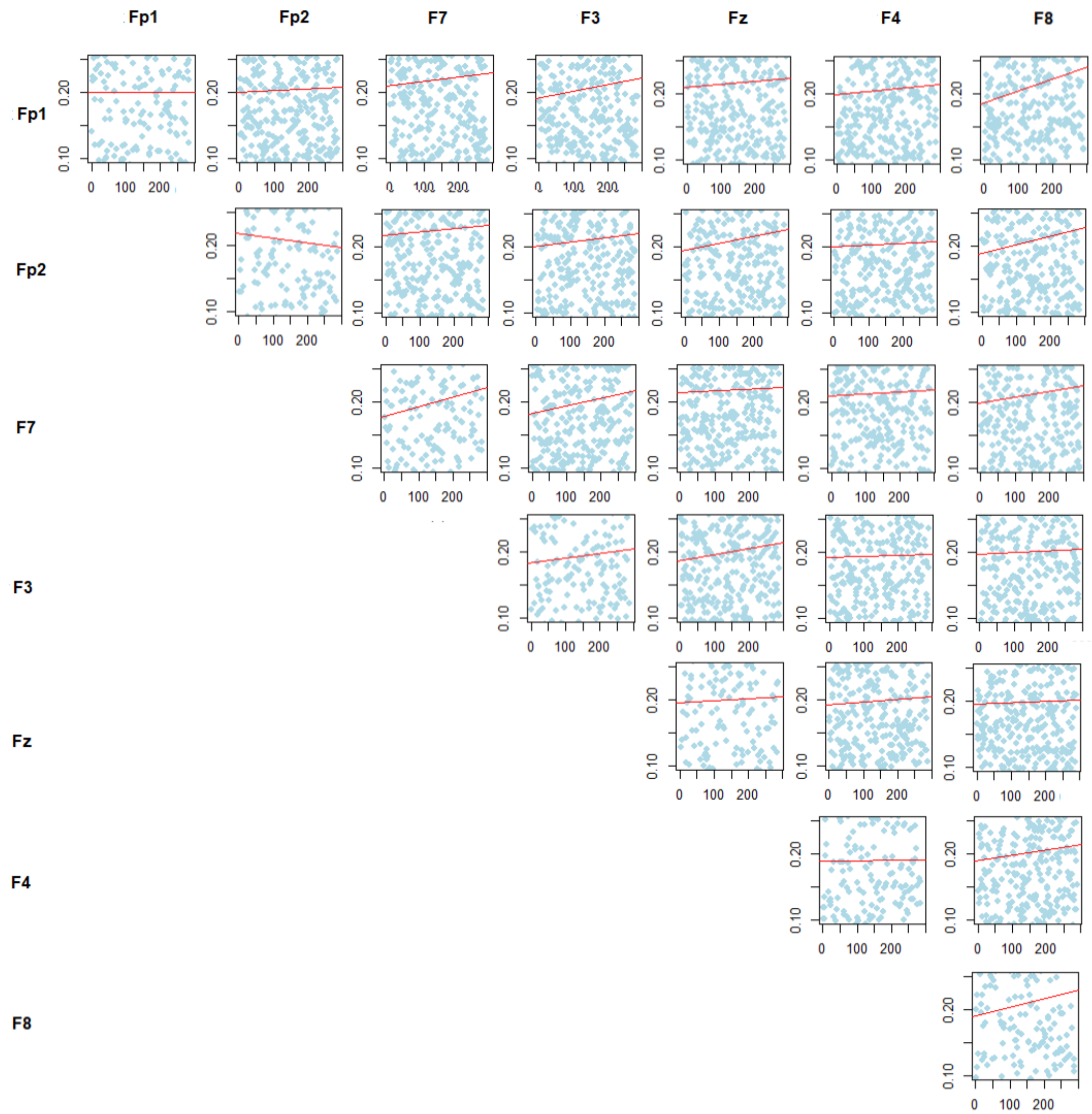


Figure L.3: Delta-CCorr per electrode combination during neurofeedback training session 3 for team A. X-axis: trial number. Y-axis: Delta-CCorr.

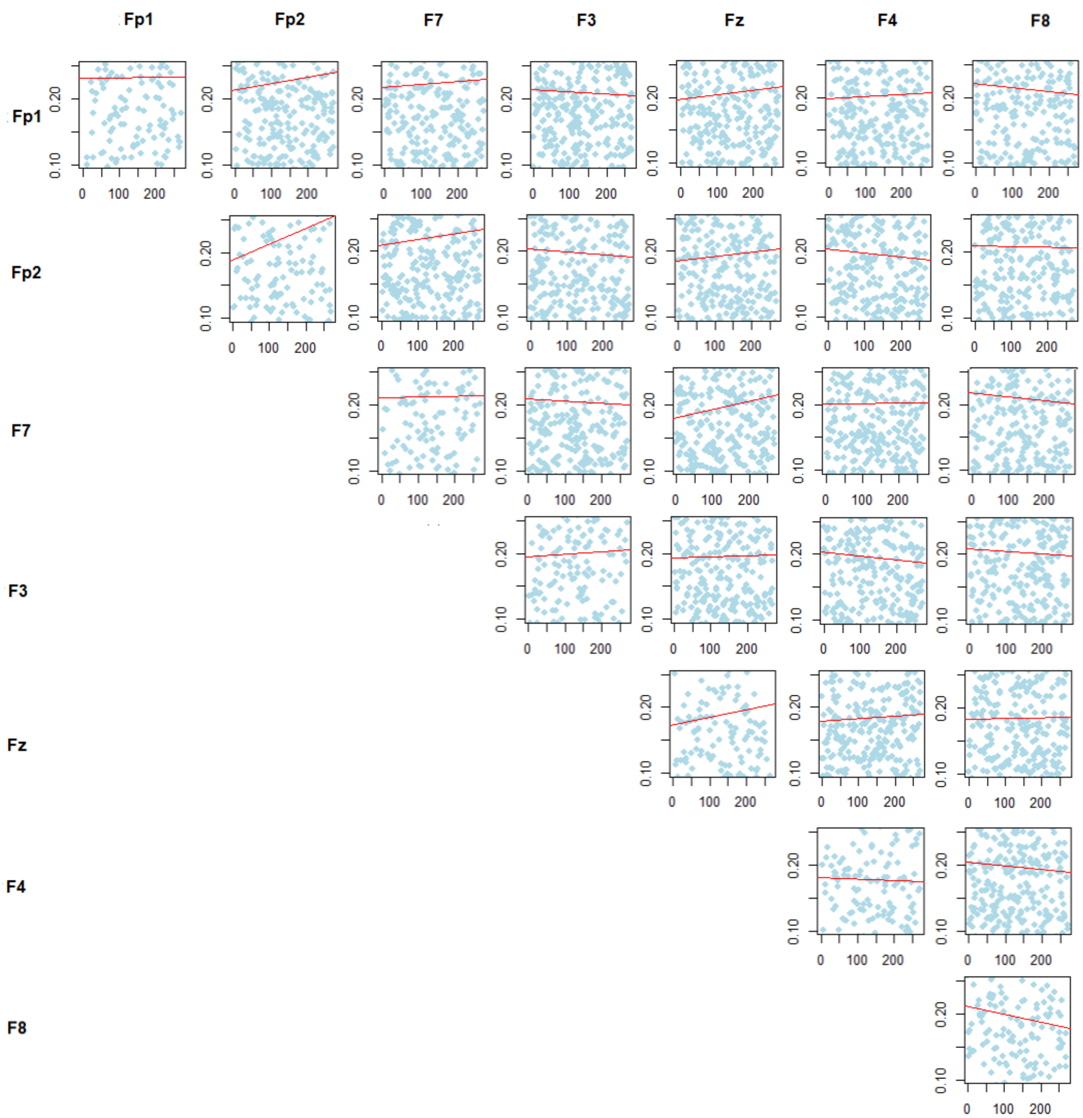


Figure L.4: Delta-CCorr per electrode combination during neurofeedback training session 1 for team B. X-axis: trial number. Y-axis: Delta-CCorr.

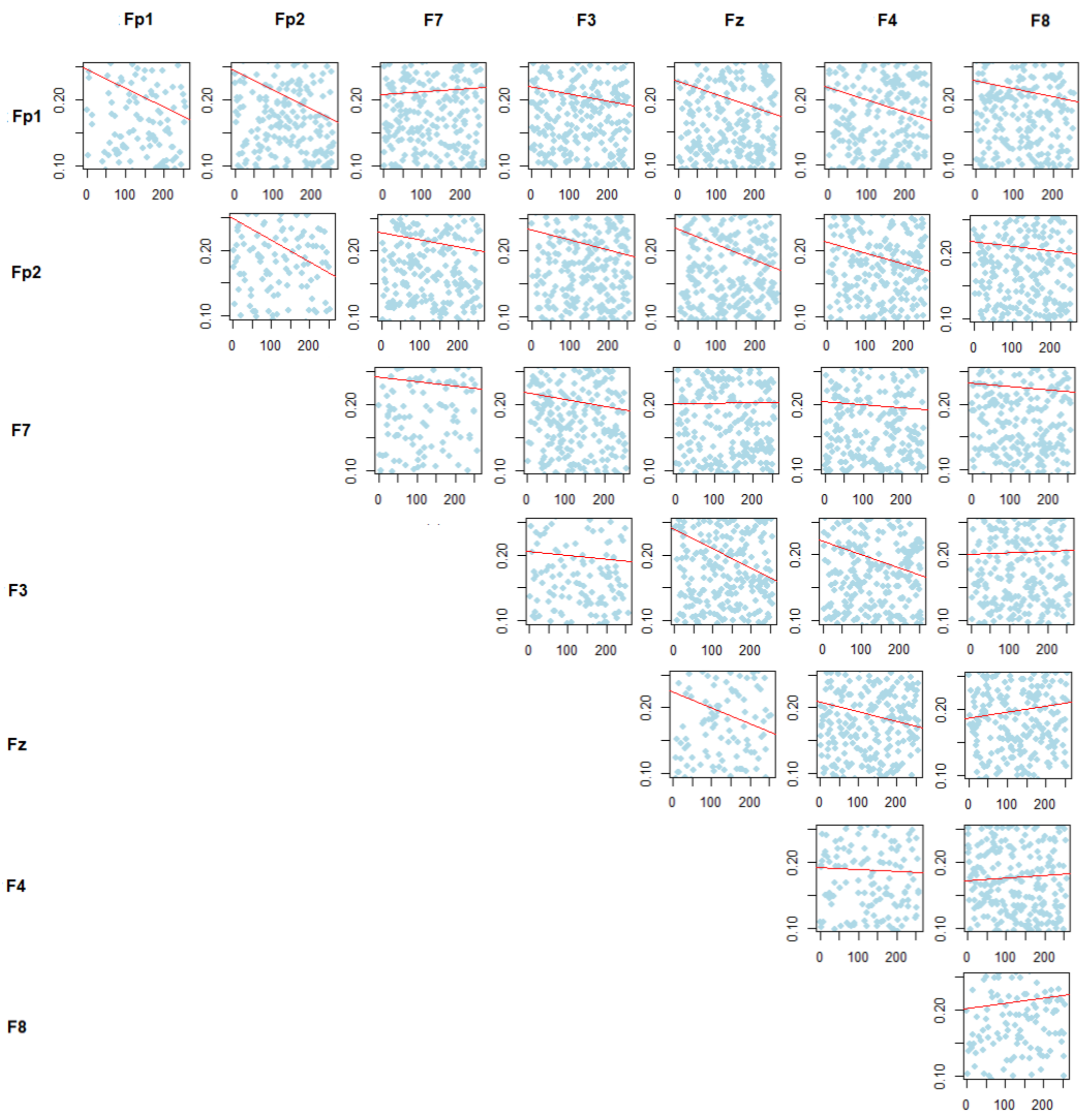


Figure L.5: Delta-CCorr per electrode combination during neurofeedback training session 2 for team B. X-axis: trial number. Y-axis: Delta-CCorr.

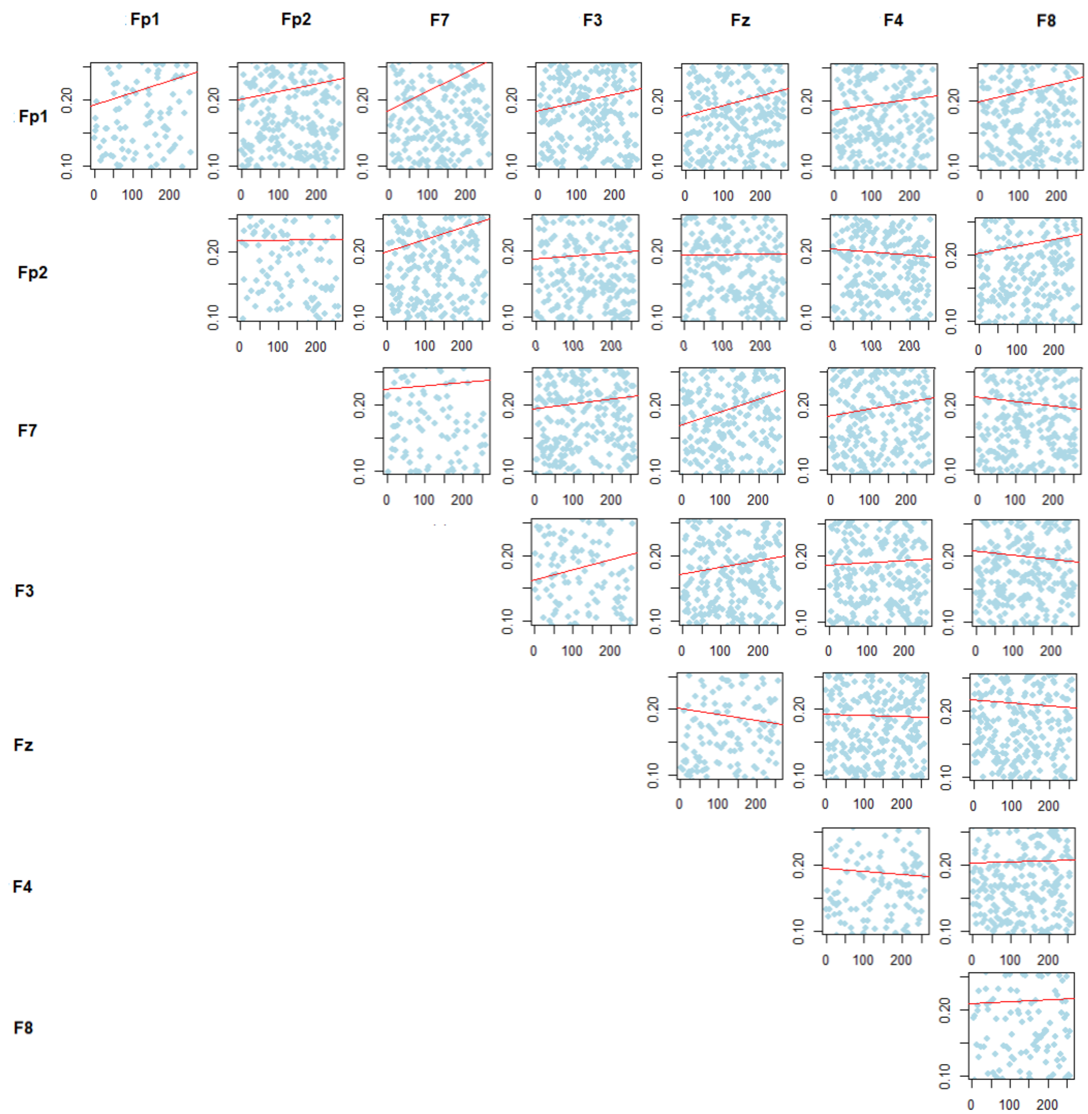


Figure L.6: Delta-CCorr per electrode combination during neurofeedback training session 3 for team B. X-axis: trial number. Y-axis: Delta-CCorr.

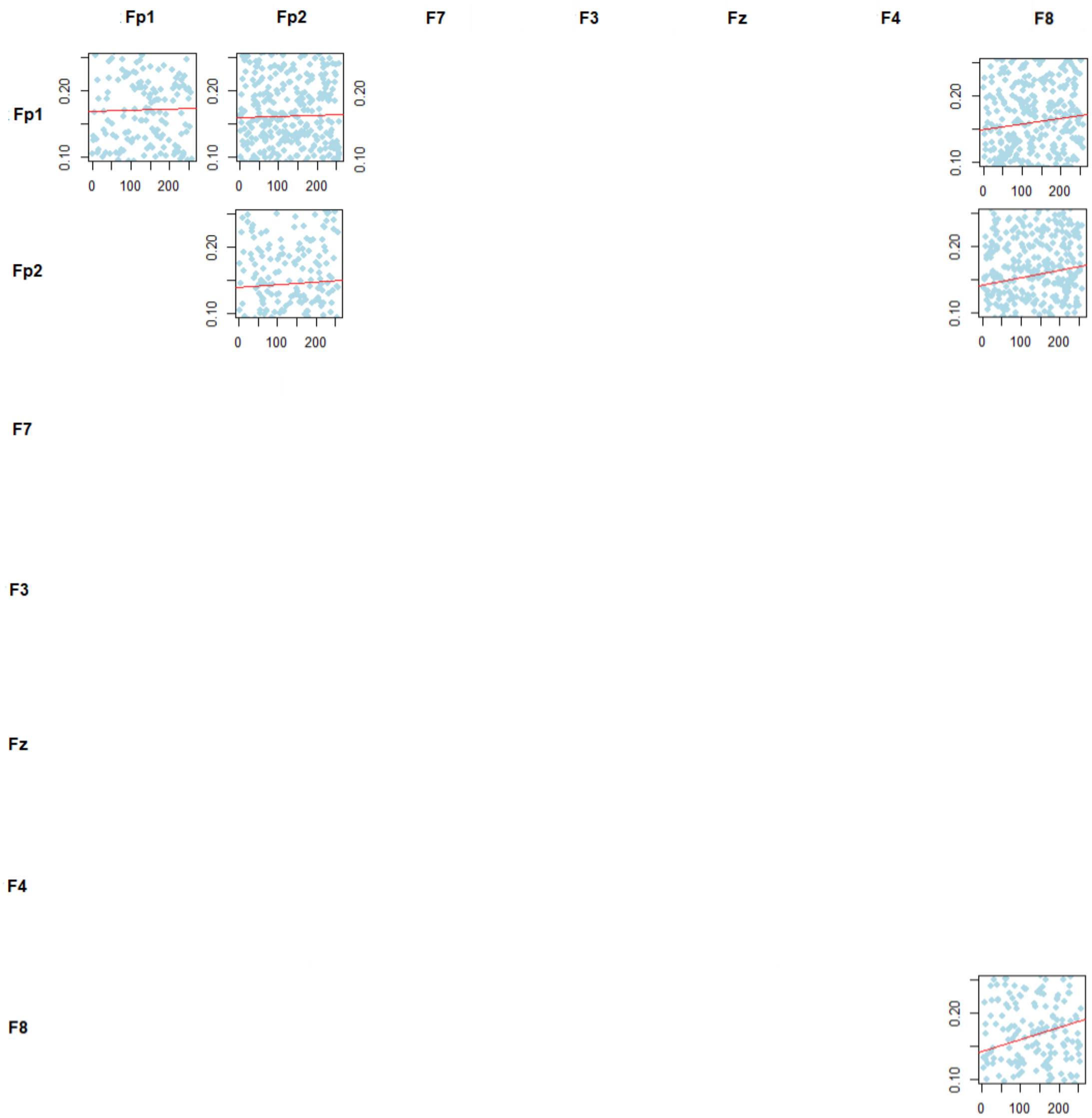


Figure L.7: Theta-CCorr per electrode combination during neurofeedback training session 1 for team A. X-axis: trial number. Y-axis: Theta-CCorr.

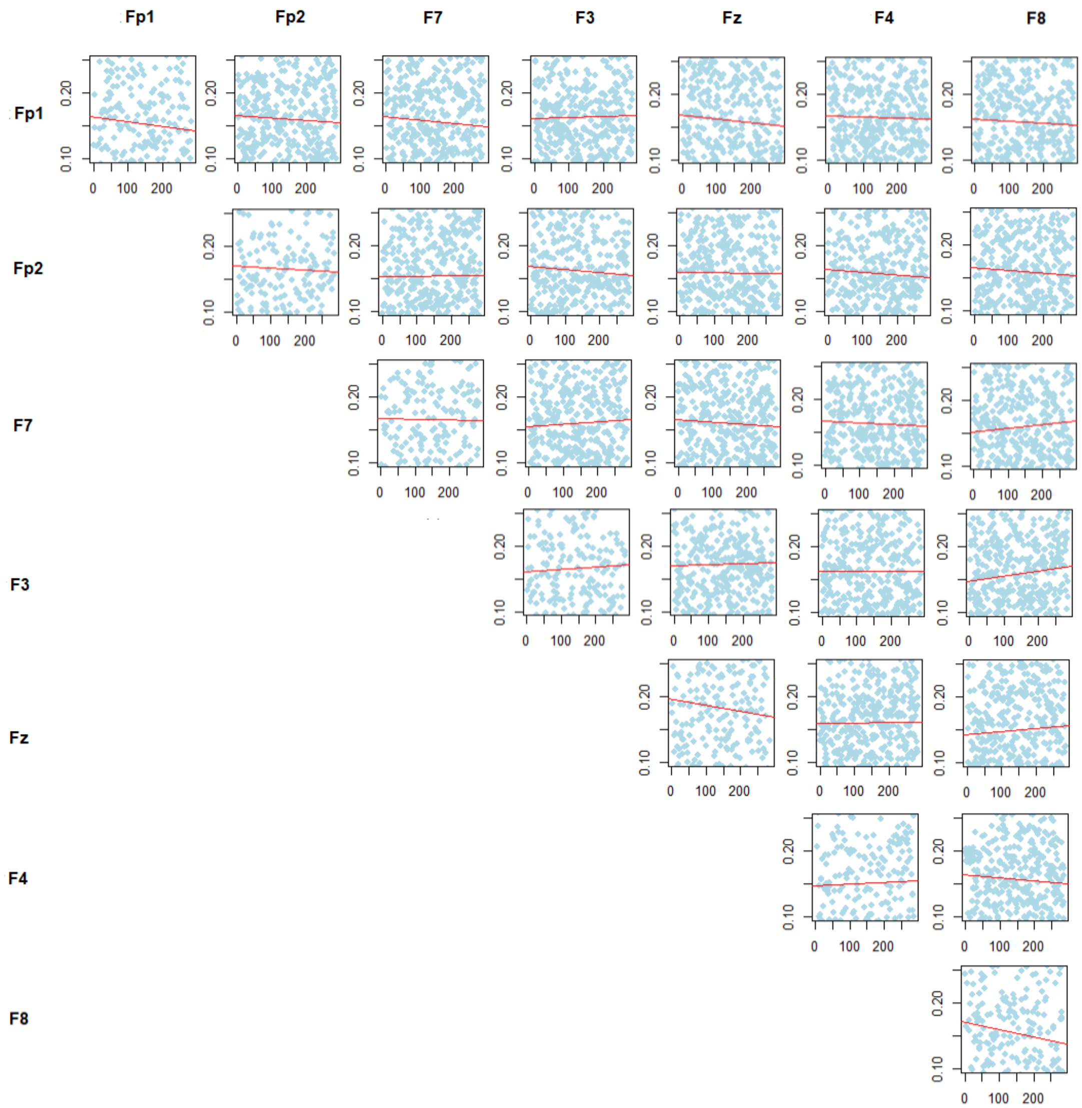


Figure L.8: Theta-CCorr per electrode combination during neurofeedback training session 2 for team A. X-axis: trial number. Y-axis: Theta-CCorr.

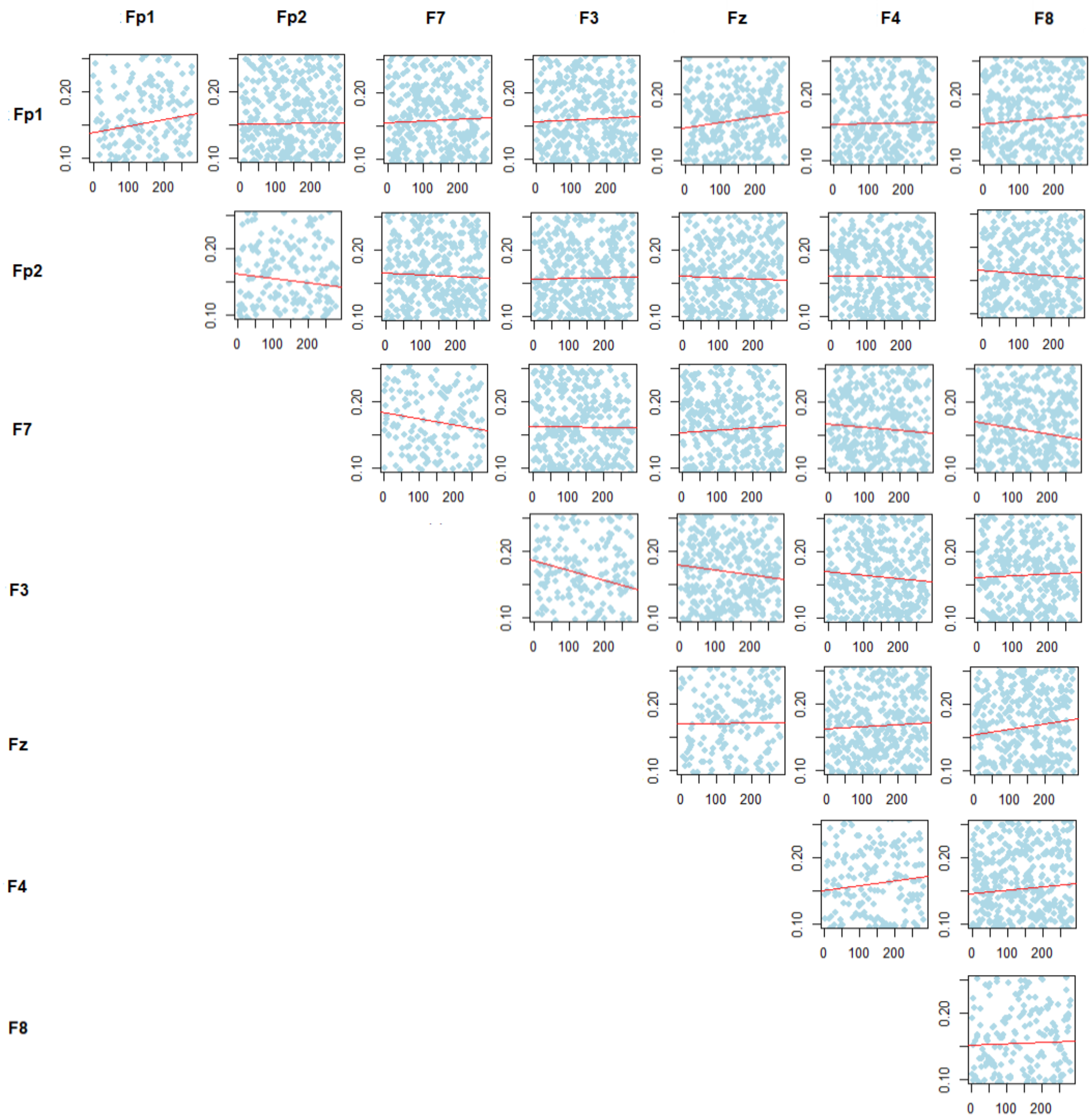


Figure L.9: Theta-CCorr per electrode combination during neurofeedback training session 3 for team A. X-axis: trial number. Y-axis: Theta-CCorr.

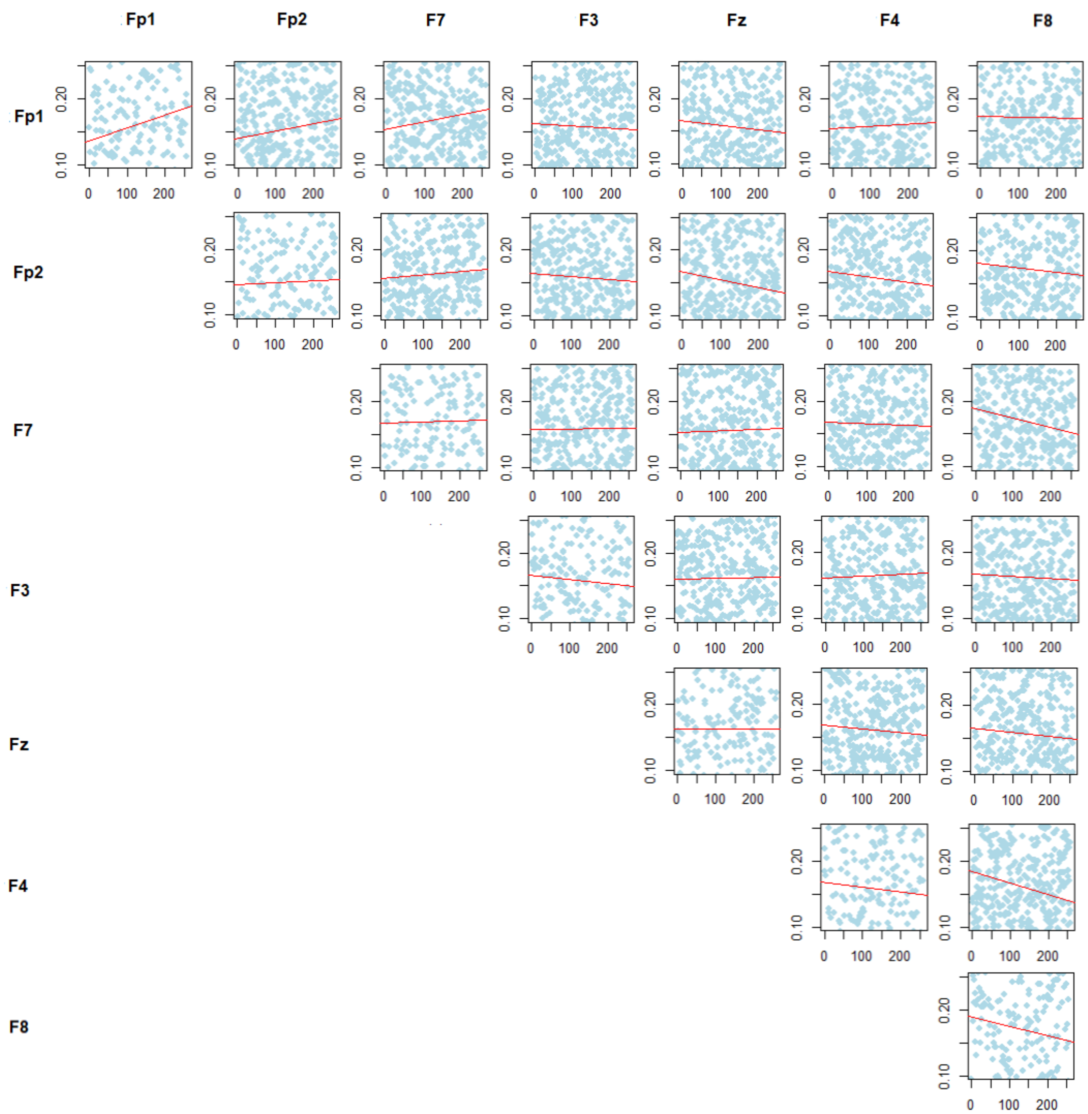


Figure L.10: Theta-CCorr per electrode combination during neurofeedback training session 1 for team B. X-axis: trial number. Y-axis: Theta-CCorr.

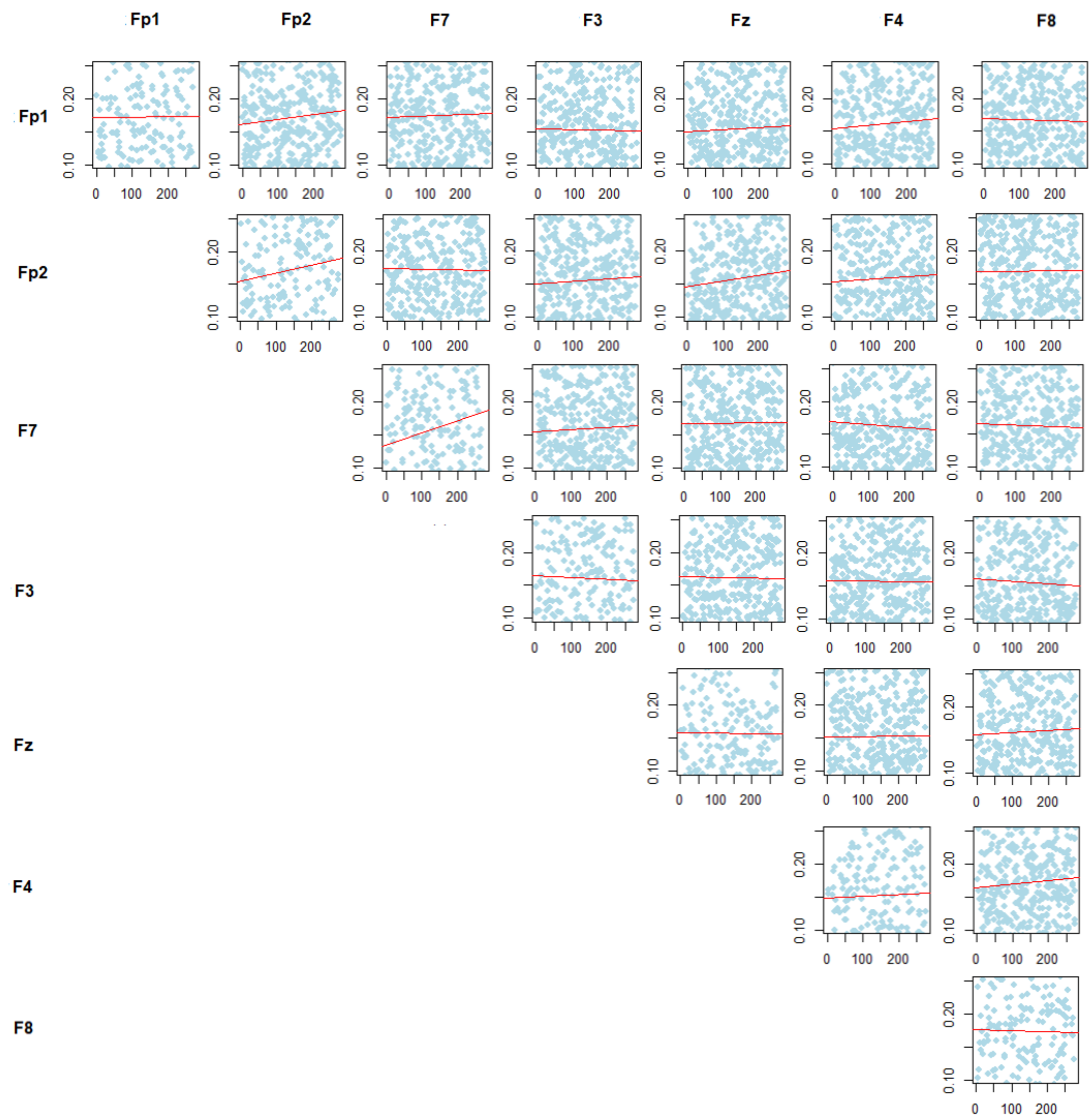


Figure L.11: Theta-CCorr per electrode combination during neurofeedback training session 2 for team B. X-axis: trial number. Y-axis: Theta-CCorr.

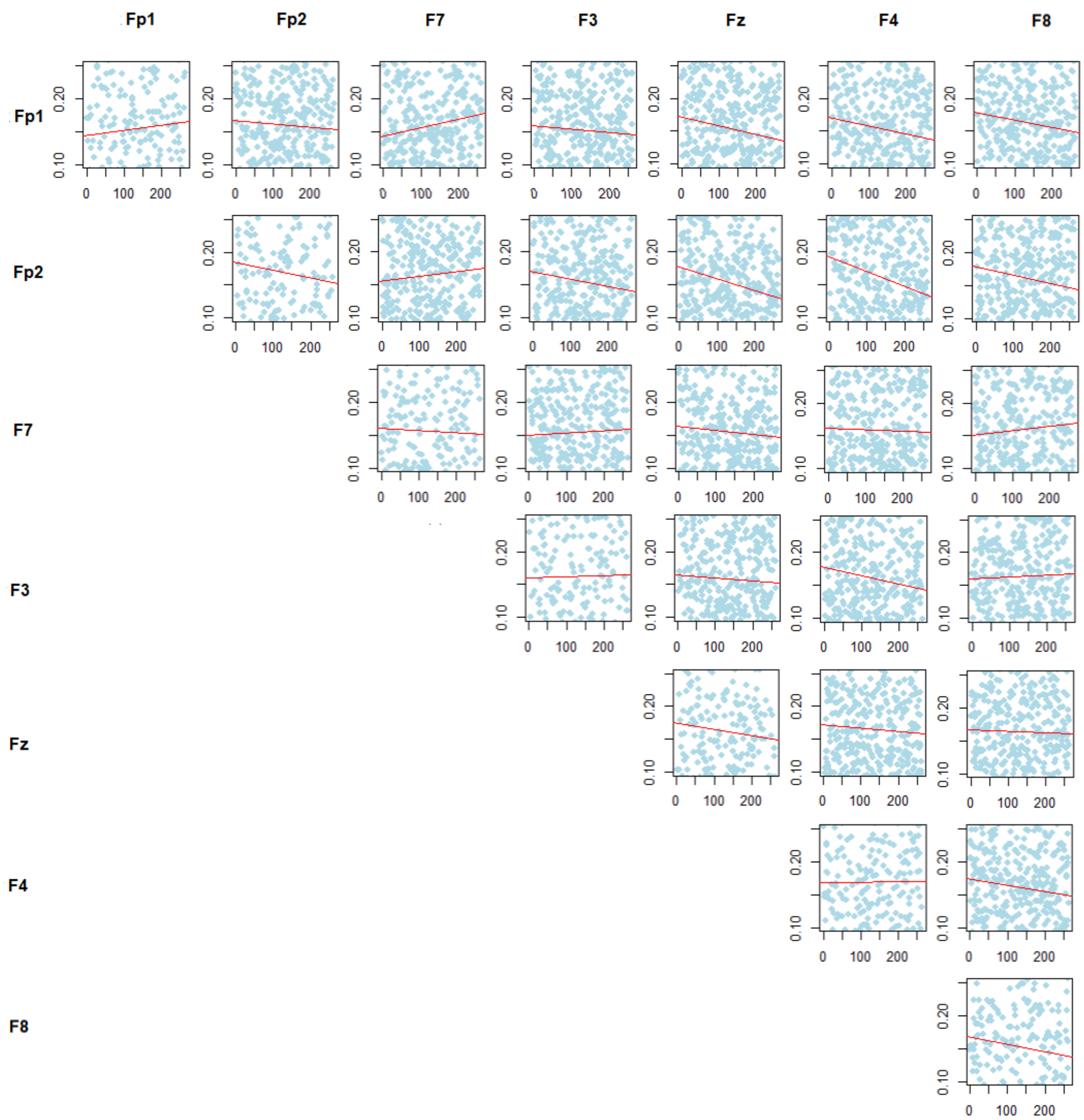


Figure L.12: Theta-CCorr per electrode combination during neurofeedback training session 3 for team B. X-axis: trial number. Y-axis: Theta-CCorr.

Appendix M. Significance Delta INS-CCorr between session 1 and 3

	Fp1	Fp2	F7	F3	Fz	F4	F8
Fp1	1.00	.613	1.00	1.00	1.00	1.00	1.00
Fp2	/	1.00	1.00	1.00	1.00	1.00	1.00
F7	/	/	1.00	1.00	1.00	1.00	1.00
F3	/	/	/	1.00	1.00	1.00	1.00
Fz	/	/	/	/	1.00	1.00	$7.84{*}10^{-3}{*}$
F4	/	/	/	/	/	1.00	$1.93{*}10^{-3}{*}$
F8	/	/	/	/	/	/	1.00

Figure M.1: Paired permutation tests with 100,000 permutations between the delta-CCorr of session 3 (during delta NFT) and the delta-CCorr of session 1 (during delta NFT) for each electrode pair. $\alpha = 5.00 *10^{-2}$, $H_{alternative}$ = two-sided, Bonferroni corrected.
.

Appendix N. Significance Theta INS-CCorr between session 1 and 3

	Fp1	Fp2	F7	F3	Fz	F4	F8
Fp1	1.00	1.00	.899	1.00	1.00	1.00	1.00
Fp2	/	.515	1.00	1.00	1.00	1.00	1.00
F7	/	/	1.00	1.00	1.00	1.00	1.00
F3	/	/	/	1.00	1.00	1.00	1.00
Fz	/	/	/	/	1.00	1.00	1.00
F4	/	/	/	/	/	1.00	1.00
F8	/	/	/	/	/	/	.865

Figure N.1: Paired permutation tests with 100,000 permutations between the theta-CCorr of session 3 (during theta NFT) and the theta-CCorr of session 1 (during theta NFT) for each electrode pair. $\alpha = 5.00 * 10^{-2}$, $H_{alternative}$ = two-sided, Bonferroni corrected.
.

Appendix O. Significance Delta INS-CCorr between part 1 and 3

	Fp1	Fp2	F7	F3	Fz	F4	F8
Fp1	1.0	1.0	.406	1.0	1.0	1.0	1.0
Fp2	/	1.0	1.0	1.0	1.0	1.0	1.0
F7	/	/	1.0	1.0	.199	1.0	1.0
F3	/	/	/	1.0	1.0	1.0	1.0
Fz	/	/	/	/	1.0	1.0	1.0
F4	/	/	/	/	/	1.0	1.0
F8	/	/	/	/	/	/	1.0

Figure O.1: Paired permutation tests with 100,000 permutations between the delta-CCorr of part 3 (during delta NFT) and the delta-CCorr of part 1 (during delta NFT) for each electrode pair. $\alpha = 5.00 * 10^{-2}$, $H_{alternative}$ = two-sided, Bonferroni corrected.
.

Appendix P. Significance Theta INS-CCorr between part 1 and 3

	Fp1	Fp2	F7	F3	Fz	F4	F8
Fp1	1.00	1.00	.899	1.00	1.00	1.00	1.00
Fp2	/	.515	1.00	1.00	1.00	1.00	1.00
F7	/	/	1.00	1.00	1.00	1.00	1.00
F3	/	/	/	1.00	1.00	1.00	1.00
Fz	/	/	/	/	1.00	1.00	1.00
F4	/	/	/	/	/	1.00	1.00
F8	/	/	/	/	/	/	.865

Figure P.1: Paired permutation tests with 100,000 permutations between the theta-CCorr of part 3 (during theta NFT) and the telta-CCorr of part 1 (during theta NFT) for each electrode pair. $\alpha = 5.00 * 10^{-2}$, $H_{alternative}$ = two-sided, Bonferroni corrected.
.

Appendix Q. Difference between the amplitude during INS-NFT session 1 and 3

	p-value theta during theta	p-value theta during delta	p-value delta during theta	p-value delta during delta
Fp1	$2.10*10^{-3}*$	.152	1.00	.116
Fp2	.265	.939	1.00	1.00
F9	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$	$9.45*10^{-3}**$	$7.67*10^{-2}$
F7	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$	1.00	1.00
F3	$p<1.00*10^{-4}***$	.128	.158	1.00
Fz	1.00	.738	1.00	1.00
F4	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$	1.00	1.00
F8	$3.78*10^{-2}*$	.128	$5.25*10^{-3}**$	1.00
F10	1.00	1.00	$4.20*10^{-2}*$	.221
T9	-	-	-	-
T7	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$	1.00	1.00
C3	.447	$2.10*10^{-2}*$	.347	$1.36*10^{-2}*$
C4	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$	.107	$p<1.00*10^{-4}***$
T8	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$
T10	-	-	-	-
P9	$3.26*10^{-2}*$	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$
P7	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$	.289	$1.68*10^{-2}*$
P3	-	-	-	-
Pz	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$	$2.00*10^{-2}*$	$p<1.00*10^{-4}***$
P4	$p<1.00*10^{-4}***$	1.0	$1.05*10^{-3}**$	1.00
P8	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$	$1.05*10^{-3}**$	1.00
P10	$8.19*10^{-2}$	$4.20*10^{-3}**$	$p<1.00*10^{-4}***$	$1.05*10^{-3}**$
O1	$p<1.00*10^{-4}***$	$p<1.00*10^{-4}***$	$1.89*10^{-2}*$	$6.30*10^{-3}**$
O2	$2.73*10^{-2}*$	$p<1.00*10^{-4}***$	1.00	$p<1.00*10^{-4}***$
O2				

Figure Q.1: Paired permutation tests with 20,000 permutations between the amplitude of session 3 and the amplitude of session 1 for each electrode, for the theta and delta frequency and for each block (theta and delta). A horizontal line denotes that not enough data is gathered of the particular electrode. $\alpha = 5.00 *10^{-2}$, $H_{alternative}$ = two-sided, Bonferroni corrected.
.

	p-value alpha during theta	p-value alpha during delta	p-value beta during theta	p-value beta during delta
Fp1	.755	1.00	1.00	$2.21{*}10^{-2}*$
Fp2	$7.35{*}10^{-3}**$	1.00	.627	$2.10{*}10^{-3}**$
F9	$p<1.00{*}10^{-4}***$	$1.05{*}10^{-3}**$	1.00	.285
F7	$6.30{*}10^{-3}**$	$p<1.00{*}10^{-4}***$	1.00	.714
F3	1.00	1.00	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$
Fz	.356	1.00	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$
F4	1.00	1.00	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$
F8	$1.58{*}10^{-2}*$	1.00	$p<1.00{*}10^{-4}***$	$1.78{*}10^{-2}*$
F10	1.00	$2.63{*}10^{-2}*$	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$
T9	-	-	-	-
T7	$2.10{*}10^{-3}**$	$p<1.00{*}10^{-4}***$	1.00	.714
C3	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$
C4	$1.05{*}10^{-2}*$	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$
T8	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$	$4.20{*}10^{-3}**$	$p<1.00{*}10^{-4}***$
T10	-	-	-	-
P9	.424	1.00	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$
P7	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$
P3	-	-	-	-
Pz	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$	$4.10{*}10^{-2}*$	.645
P4	.129	$p<1.00{*}10^{-4}***$	.606	$3.57{*}10^{-2}*$
P8	.188	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$
P10	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$
O1	1.00	$3.04{*}10^{-2}*$	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$
O2	$p<1.00{*}10^{-4}***$	1.00	$p<1.00{*}10^{-4}***$	$p<1.00{*}10^{-4}***$

Figure Q.2: Paired permutation tests with 20,000 permutations between the amplitude of session 3 and the amplitude of session 1 for each electrode, the alpha and beta frequency and for each block (theta and delta). A horizontal line denotes that not enough data is gathered of the particular electrode. $\alpha = 5.00 *10^{-2}$, $H_{alternative}$ = two-sided, Bonferroni corrected.

.

	p-value gamma during theta	p-value gamma during delta
Fp1	p<1.00*10^{-4} * **	1.00
Fp2	p<1.00*10^{-4} * **	.469
F9	.371	1.00
F7	1.00	4.31*10^{-2}*
F3	2.10*10^{-3} * *	1.00
Fz	p<1.00*10^{-4} * **	p<1.00*10^{-4} * **
F4	p<1.00*10^{-4} * **	1.05*10^{-3} * *
F8	p<1.00*10^{-4} * **	2.10*10^{-3} * **
F10	p<1.00*10^{-4} * **	p<1.00*10^{-4} * **
T9	-	-
T7	1.00	7.67*10^{-2}
C3	.375	.242
C4	p<1.00*10^{-4} * **	p<1.00*10^{-4} * **
T8	1.00	6.30*10^{-4}
T10	-	-
P9	p<1.00*10^{-4} * **	p<1.00*10^{-4} * **
P7	p<1.00*10^{-4} * **	p<1.00*10^{-4} * **
P3	-	-
Pz	p<1.00*10^{-4} * **	p<1.00*10^{-4} * **
P4	1.00	1.00
P8	p<1.00*10^{-4} * **	1.00
P10	p<1.00*10^{-4} * **	p<1.00*10^{-4} * **
O1	p<1.00*10^{-4} * **	p<1.00*10^{-4} * **
O2	p<1.00*10^{-4} * **	p<1.00*10^{-4} * **

Figure Q.3: Paired permutation tests with 20,000 permutations between the amplitude of session 3 and the amplitude of session 1 for each electrode, for the gamma frequency and for each block (theta and delta). A horizontal line denotes that not enough data is gathered of the particular electrode. $\alpha = 5.00 *10^{-2}$, $H_{alternative}$ = two-sided, Bonferroni corrected.
.

List of Abbreviations

ACC Anterior Cingulate Cortex

ACI Absolute Couping Index

ADHD Attention-Deficit Hyperactivity Disorder

AI Anterior Insula

AP Action Potential

ASD Autistic Spectrum Disorder

BOLD Blood-oxygen-level-dependent

CAR Common Average Reference

CCorr Circular Correlation Coefficient

EEG Electroencephalography / Electroencephalogram

ERO Event Related Oscillation

ERP Event Related Potential

FIR Finite Impulse Response

fMRI Functional Magnetic Resonance Imaging

fNIRS Functional Near-infrared Spectroscopy

INS Interpersonal Neural Synchrony

LED Light-emitting diode

MCN Modified combinatorial nomenclature

MEG Magnetoencephalography

MNS Mirror Neuron System

MR Magnetic Resonance

MS Mentalizing System

NFT Neurofeedback Training

OFC Orbitofrontal Cortex

PFC Prefrontal Cortex

PLI Phase-Locking Index

PLV Phase-Locking Value

PSP Postsynaptic Potential

REST Reference electrode standardization technique

RT Reaction Time

RTD Reaction Time Difference

STS Superior Temporal Sulcus

ToM Theory-of-Mind

TPJ Temporoparietal Junction