Optimization of near-infrared spectroscopy-based neurofeedback for use in the treatment of attention-deficit hyperactivity disorder

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

M.Sc. Justin Hudak

aus Baltimore, Maryland

USA

Tübingen

2019

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	14.03.2019
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Dr. Ann-Christine Ehlis
2. Berichterstatter:	Prof. Dr. Caterina Gawrilow

Acknowledgments

I would like thank Ann-Christine for persistent, loving support during the entire course of my PhD, as well as convincing me beforehand to apply for the position. I enjoyed getting to know you first as an Advisor, and then most of all as a person. I would like to thank Caterina for being open and providing extra support in the field of school psychology. I would like to thank Thomas for being our lab stalwart, for taking care of Coffee and Statistics, and providing Saturday working relief. I would like to thank Ben-Steffen and David for technical, algorithmic and idea-ical support when caught in a tight spot. I would like to thank Ramona and Betti for all their support in organizing the miasma of measurement conflicts. I would like to thank my HIWIs for their constant support and ability to quickly learn and adapt to our difficult experimental paradigms, especially Nana and Claudia, who took on the lion's share of responsibility. I would like to thank Friederike for her collaboration on our main projects, sometimes through bouts of strong-willed personalities. I would like to thank my three parents, Tommy and Emmy for their unshakable love and support, and for inspiring me to live my life like they do. I would like to thank my grandfather Phil, who was the original inspiration. And finally, I would like to thank Bärchen, who made my dissertation last a year longer than it should have, but what a year it was...

List of Commonly Used Abbreviations

ADHD Attention-deficit hyperactivity disorder

BOLD Blood-oxygen level-dependent

CAR Common average reference

DLPFC Dorsolateral prefrontal cortex

DSM-V Diagnostic and Statistical Manual of Mental Disorders, Version 5

EEG Electroencephalograpy

EMG Electromyography

ERN/Ne Error-related negativity

ERP Event-related potential

FPCN Frontoparietal control network

HC Healthy controls

NF Neurofeedback

NIRS Near-infrared spectroscopy

O_2HB Oxygenated hemoglobin

VR Virtual reality

fMRI (Functional) magnetic resonance imaging

rtfMRI Real-time functional magnetic resonance imaging

Abstract

Attention-deficit hyperactivity disorder (ADHD) belongs to the most common disorders in both children and adults, affecting anywhere from 2.5-5% of the population worldwide. The disorder affects patients on three key axes: hyperactivity, impulsivity, and inattention; and is pervasive in several aspects of life, most clearly in the classroom, at work, and in interpersonal relations. Several treatment options exist that target the various symptomatology of ADHD; however, the most popular method, medication, has several side effects which are not advantageous. Neurofeedback (NF), while not a new method per se, is a treatment method that has garnered great contention in recent years for its questionable efficacy. However, its promise as a method that lies at the crossroads between behavioral therapy and the latest neurophysiological research means that it has exciting promise as a method that can constantly be updated to reflect the latest scientific findings. The present dissertation seeks to optimize near-infrared spectroscopy (NIRS)-based NF paradigms for the treatment of ADHD through examining the following: 1) How can NIRS-based NF paradigms be implemented in VR? 2) What are the underlying mechanisms responsible for a successful NF trial and paradigm? 3) How can we combine methods to improve NF paradigms even further?

This dissertation is comprised of four studies. In **Study 1**, we explored the efficacy of a NIRS-based NF paradigm, embedded inside a virtual classroom, in treating highly-impulsive college students. The study was designed as a pilot study to test the efficacy of integrating NF in virtual reality (VR) for its prospective use in a study with children with ADHD. Highly impulsive subjects were recruited for this study using an online version of the Barrett Impulsiveness Scale (BIS). The subjects were then randomly assigned to one of two groups. The experimental NIRS-based NF group underwent 8 sessions of prefrontal oxygenated hemoglobin (0_2HB)-based NF, in which their task was to control the lighting in the classroom by either increasing prefrontal 0_2HB (brighten the light) or decreasing 0_2HB (dim the light). In the electromyography-based (EMG) control group, the subjects used their right supraspinatus muscle (brighten the light) or left *supraspinatus muscle* (dim the light) to control the lighting. Both groups undertook the 8 session across 2 weeks and were assessed on go/no-go and stop-signal tasks both before and after the training. Subjects in the experimental group showed a significant pre-post improvement in prefrontal 0_2HB during the more difficult no-go condition, in which they had to withhold their response on no-go cues. Simultaneously, they showed pre-post reductions in impulsive commission errors, the strength of which was correlated significantly with their ability to learn the NF training. Additionally, the experimental group showed reduced reaction time variability (SDRT) on the stop-signal task. The control group, by way of comparison, showed none of these improvements. There were also no improvements on an n-back task for either group. We explained this result in terms of the specificity of the training. Subjects in the NIRS group trained the dorsolateral prefrontal cortex (DLPFC), an area widely associated with ADHD and impulsive behavior. Learning to control this region, then, seemed to have an effect on impulsive behavior that was tied to the strength of learning.

In **Study 2**, we originally wanted to examine the underlying brain processes in NF trials and rests via examining the functional connectivity (FC) differences between successful and failed NF trials from a complete NF data set from a study in our lab with adult subjects with ADHD. However, upon initial analysis, we found several problems in the dataset. For one, there was a baseline bias, causing a sharp downward spike in pre-trial feedback channel activity before activation trials and a sharp upward spike in pre-trial feedback channel activity

before deactivation trials, in both cases making the resulting feedback trial easier, but also contaminating the short analysis of FC. We found that this error was due to an improper randomization of trial order. Therefore, we grouped the trials together (de- and activation) and found that some helpful activation within the frontoparietal control network (FPCN) was actually punished due to the selection of the feedback channels and most importantly, the common average reference (CAR), included in the feedback algorithm. Because some of the vital channels of the FCPN, in particular the bilateral DLPFC, were not part of the feedback region, enhanced connectivity between these regions led to statistically greater failure than success. This study showed us that the design of NF experiments has critical small details that need to be carefully controlled before conducting the experiments.

Study 3 explored multimodal brain measurement with NIRS and electroencephalography (EEG) and as such served as a basis for a NIRS-based multimodal NF paradigm. Using a complex modified Eriksen flanker task designed to force errors, we sought to look for post-error behavioral and processing differences between adults with ADHD and healthy controls (HC). To accomplish this we employed a novel analysis using wavelet-denoised single-trial P300 event-related potentials from EEG recordings to predict the hemodynamic response of NIRS in correct trials following errors and correct trials following correct trials. Results from a traditional general linear model based regression were compared with this novel EEG-informed regression for both trial types and for ADHD versus HC. Results showed a markedly stronger activation made even more strong by the EEG-informed analysis in prefrontal, motor, and sensorimotor areas for HC following errors. These results demonstrated that the addition of the enhanced temporal information from EEG can improve the specificity of the NIRS analysis. Such techniques might also be used to design a much more specific NF study for ADHD.

Study 4 is the study design from our project comparing 2D- and 3D- classroom-based NF in children with ADHD. The manuscript details the myriad dependent variables that we are currently collecting to try and isolate the important factors in NF design for children with ADHD (including motivation, movement, combined EEG-NIRS measurements, executive functioning battery, as well as scholastic achievement). While **Study 4** passes chronologically better as **Study 2**, it is not officially part of the required manuscripts for this dissertation as it was already used in the dissertation of my colleague, Friederike Blume.

In a general discussion, I propose recommendations for the design of NF experiments for optimal treatment of ADHD, implementing the latest analytical techniques, as well as scouring the minute details which are often overlooked in commercial as well as lab software.

Zusammenfassung

Die Aufmerksamkeitsdefizit-Hyperaktivitätsstörung (ADHS) zählt zu den am häufigsten vorkommenden psychischen Störungen im Kindes- und Jugend-, aber auch im Erwachsenenalter, ca. 2,5-5% der Weltbevölkerung sind davon betroffen. Die Erkrankung wirkt sich in drei Kernbereichen aus: Hyperaktivität, Impulsivität und Unaufmerksamkeit, mit tiefgreifenden Beeinträchtigungen in verschiedenen Bereichen des Lebens, am deutlichsten in der Schule, bei der Arbeit und in interpersonellen Beziehungen. Es existieren verschiedene Behandlungsansätze, welche die vielseitige Symptomatologie der ADHS anzugehen versuchen. Die Medikation stellt dabei die gängigste Behandlungsmethode dar, ist jedoch nicht selten mit Nebenwirkungen verbunden. Die Methode des Neurofeedbacks (NF) ist zwar unlängst kein neuartiges Verfahren mehr, ist aber eine Behandlungsmethode, welche, aufgrund der fraglichen Wirksamkeit, in den letzten Jahren viele Diskussionen anregt hat. Da die Methode jedoch an der Schnittstelle zur Verhaltenstherapie und der neusten neurophysiologischen Forschung liegt, bietet sie die Möglichkeit fortwährend um die aktuellsten wissenschaftlichen Erkenntnisse erweitert zu werden. Die vorliegende Dissertation strebt die Optimierung des NIRS-basierten NF Paradigmas zur Behandlung von ADHS durch die Untersuchung folgender Punkte an: 1) Wie können NIRS-basierte Paradigmen in VR implementiert werden? 2) Welche sind die zugrundeliegenden Mechanismen, die für einen erfolgreichen NF Trial und Paradigma verantwortlich sind? 3) Wie können Methoden kombiniert werden um NF Paradigmen weiter zu entwickeln?

Die Dissertation umfasst vier Studien. In Studie 1 untersuchten wir die Effektivität eines Nahinfrarotspektroskopie-(NIRS)-basierten NF Paradigmas, realisiert im virtuellen Klassenzimmer, zur Behandlung hoch-impulsiver Universitätsstudenten. In Studie 2 wurde ein 'Standard'-NIRS-basiertes NF Paradigma für ADHS in seinen Einzelteilen überprüft und Empfehlungen für die Weiterentwicklung eines NF Studiendesigns diskutiert sowie Netzwerke, welche mit Erfolg und Misserfolg bei NF Trials assoziiert sind, untersucht. In Studie 3 untersuchten wir das post-Fehlerverhalten bei ADHS-Patienten im Vergleich zu gesunden Kontrollpersonen (healthy controls, HC). Im Einzelnen betrachteten wir die 'P300-informed' NIRS-Analyse um die Unterschiede zwischen ADHS-Patienten und HC genauer zu untersuchen. Zusätzlich wurde die Studie konzipiert um die Effektivität von EEG/NIRS für den zukünftigen Einsatz in einem multimodalen NF Studiendesign für ADHS zu überprüfen. In Studie 4 wird das Studiendesign für eine großangelegte und momentan laufende Studie mit Schulkindern mit ADHS beschrieben, in welcher das Paradigma aus Studie 1 zum Einsatz kommt. Dies liefert eine erste Zusammenfassung über die umfangreiche Datenerhebung, welche zu neuen Erkenntnissen bei NIRS-NF bei ADHS führen soll.

Contents

1 General Introduction

1.1 Attention-deficit Hyperactivity Disorder

Attention-deficit hyperactivity disorder (ADHD), according to the fifth installment of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V), is broadly characterized by three behavioral domains: inattention, hyperactivity and impulsivity (American Psychiatric Association, 2013). The prevalence of the disorder is widespread, with childhood estimates worldwide ranging around 5% (Polanczyk, 2007). In roughly half of these cases, ADHD persists into adulthood, leading to a persisting prevalence of around 2.5% in adults (Faraone et al., 2015). The disorder is problematic in both life stages for differing reasons. In children and adolescents, behavior at home and in social situations can put a strain on parental relationships and make it difficult to form lasting friendships (Wehmeier et al., 2010). Generally, as an impact of ADHD must be present in at least two life domains in order to be classified, in addition to the home, the scholastic setting is also a problem for many children and adolescents (Daley and Birchwood, 2010). In adults, there are corollary problems in romantic relationships and in the workplace or in collegiate studies (Bruner et al., 2015; Shaw-Zirt et al., 2005). There is clearly a need to treat this prevalent, and often psychosocially debilitating disorder. However, we must first examine the relevant symptomatic of the disorder.

1.1.1 Impulsivity and Self-regulation

Clinical diagnoses of ADHD must adhere to strict standards, and with good reason, as a diagnosis of ADHD can lead to a stigma for the child and parents. However, in the scientific world, when trying to better understand the disorder, it can be helpful to isolate the various symptom axes. This is why, despite the categorical taxonomic classifications of ADHD as a disorder differing from normal that are supplied by both the DSM-V and the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (Doernberg and Hollander, 2016), many researchers choose to target a particular dimension of ADHD, and thus seek subclinical populations in order to more clearly isolate the desired behaviors (Balázs and Keresztény, 2014). Some studies have thus targeted the inattentive behaviors of ADHD (Seidman, 2006) while others examine the hyperactive and impulsive dimensions (Lijffijt et al., 2005). This makes sense, as even within the taxonomical constraints of the DSM-V, there is a separation between the primarily inattentive and primarily hyperactive/impulsive presentations, thus lending weight to the argument that studying the behavioral symptoms individually might lead to improved understanding and treatment development (American Psychiatric Association, 2013).

In fact, there exist several prevailing etiological models of ADHD that focus on inhibitory dysfunction as the driving force behind ADHD, at least within the hyperactive/impulsive subtype (Tannock, 1998). Perhaps the most popular model in that vein, the behavioral inhibition model developed by Barkley (1997), seeks to establish a unified theory of ADHD, which purports that not only hyperactive/impulsive symptoms of ADHD, but also symptoms of inattention are caused by deficits in behavioral inhibition. Secondary deficits are then focused around a set

of four executive functions: working memory, self-regulation, internalization of speech, and reconstitution. In the context of this dissertation, self-regulation is the key deficit intertwining the four manuscripts. Self-regulation refers to the ability of the person to regulate motivation, arousal, and affect (Diamond, 2013). As three of the studies involved are neurofeedback studies, self-regulation is vital to the process of being able to regulate one's own brain state. In the fourth study on post-error behavior, self-regulation is crucial in order to detect and act on erroneous behavior (Shiels and Hawk, 2010). Barkley argues that inattention is caused by poor inhibition in the sense that paying attention is seen as an action that must be maintained over time. ADHD patients are thereby not able to self-regulate and inhibit the prepotent response of switching to a more salient stimulus. In this dissertation, we examine how NF treatments might be optimized in order to treat impulsivity and secondarily inattention via targeting, above all, the prefrontal cortex (PFC). As we will see in the coming chapter, the PFC, and especially the dorsolateral PFC (DLPFC), are vital to cognitive control, inhibition, and stimulus response.

1.1.2 Impulsivity and the Brain

Impulsivity is often framed as an inability to inhibit very strong, or prepotent behavioral responses to external stimuli (Bari and Robbins, 2013; Logan et al., 1997). It is a behavioral trait strongly linked to impairments in the frontal lobe, with dysfunctional frontal lobe activity and frontal lobe excisions both contributing to this exophenotype (Bari and Robbins, 2013; Fallgatter and Herrmann, 2001). Impulsivity is of course not exclusive to ADHD and, in this sense, there is more weight lent to the argument that ADHD should be studied dimensionally, rather than categorically. Highly-impulsive individuals, for example, share similar intermediate phenotypes in regards to prefrontal brain deficits, particularly on tasks that involve response inhibition, and also exophenotypes in the form of a behavioral inability to inhibit responses.

Bari and Robbins (2013) discuss impulsivity at length, separating it into two broad categories: one of cognitive impulsivity and one of behavioral impulsivity, such as response inhibition. Response inhibition is perhaps the most-studied subset of motor-impulsivity related behavior problems that affect both highly impulsive, but healthy, subjects and patients with ADHD. Both groups share particularly strong similarities on a behavioral level when it comes to withholding prepotent responses, or responses that have been conditioned to be automatic through repetition (Aichert et al., 2012). Tasks that condition a prepotent response and then require the subject to inhibit this response include the stop-signal task (SST), the Stroop task, the continuous performance task (CPT), and the go/no-go task. The SST and the go/no-go task, in particular, have been widely studied in both groups on both the behavioral and physiological levels. In a typical go/no-go task the subject must respond as quickly as possible to *go* signals, which make up typically 70-80% of trials. In this way, these responses become prepotent, or automated, which interferes with the inhibitory response triggered by the *no-go* trials, wherein the subject should withhold their response.

The main measure of response inhibition in the go/no-go task is commission errors, or errors in which the subject should have withheld a response, but did not. Children and adults with

ADHD typically make more commission errors than their healthy peers (Cubillo et al., 2012; Geburek et al., 2013; Rubia et al., 2007; Smith et al., 2004). Response inhibition on the go/no-go task correlates positively with increased ratings of impulsivity on the Barratt Impulsive Scale (BIS;BIS-11) (Patton et al., 1995; Aichert et al., 2012). (Weidacker et al., 2016) found that high scores on the BIS subsection *Cognitive Complexity* inversely correlate with an ability to inhibit a response on a go/no-go task with added working memory load. Behavioral deficits on go/no-go tasks are also seen in highly-impulsive subgroups such as binge-eaters – where accuracy in the *no-go* condition again scaled inversely with BIS score (Hege et al., 2014) and binge-drinkers (Henges and Marczinski, 2012) where amount of drinks in a single sitting predicted response inhibition on a go/no-go task as well as general impulsivity on the BIS.

Development of impulse control is a process that happens over time from childhood to adulthood. There is, of course, a large heterogeneity in developmental patterns, but in the end it comes down to a competition between two distinct brain systems competing for control – a *socioemotional system* pushing the child or adolescent to experiment with novel stimuli and a *cognitive control* system that attempts to regulate potentially dangerous impulses (Casey et al., 2008; Shulman et al., 2016; Steinberg, 2008). The lateral prefrontal cortex and its associated connections with, in particular, striatal, parietal, motoric, and other frontal areas, seems to be responsible for the development of this cognitive control network, and thus the ability to control impulse (for comprehensive reviews see (Cubillo et al., 2012; Rubia et al., 2013)). Not surprisingly, ADHD is often characterized by sweeping deficits within these cognitive control areas, in particular the lateral prefrontal cortex (Booth et al., 2005; Cubillo et al., 2010; Rubia et al., 2009, 2010, 2014). The prefrontal regions of ADHD patients also seem to be subject to a delayed maturation rate relative to healthy controls (Shaw et al., 2007), further highlighting the imortance of this cortical area in the cognitive disabilities associated with the disorder.

Regarding highly impulsive subjects, Farr et al. (2012) found a negative correlation between trait impulsiveness and activation of the middle frontal cortex as well as between trait impulsiveness and functional connectivity between the MFC and DLPFC. Ding et al. (2014) reported that highly impulsive internet gamers have to activate the cognitive control network more strongly, relative to healthy controls, to achieve similar behavioral results on a go/no-go task. Horn et al. (2003) found a similar result in a general population of impulsive individuals. It would seem then, that targeting a clinical intervention to improve prefrontal brain activation could prove beneficial in lowering impulsive behavioral tendencies. The question is, is it better to target a specific disorder with categorical boundaries, like ADHD, or to focus on one dimension, and perhaps increase the generalizability of a treatment to a larger population. Firstly, we will examine the existing treatments most commonly used for ADHD.

1.2 ADHD Treatment

In the United States of America (USA), by far the most common treatment used for ADHD is psychopharmaceuticals, the most commonly used of which are stimulants such as Methylphenidate, of which Ritalin is probably the best known (Faraone et al., 2015). In Europe, there is a ten-

dency to attempt to treat ADHD with non-pharmacological interventions first, such as behavioral modification, dietary interventions, or neurocognitive interventions, such as neurofeedback (Gevensleben et al., 2014a). The American tendency to treat ADHD with medication is largely driven by efficacy: stimulant-based medications are highly effective, particularly for impulsivity, in the short-term, the time-frame of which is most often explored in clinical studies (Faraone and Glatt, 2010). This is in fitting with the mindset of the United States, where the desire for quick fixes that gloss over underlying problems is rampant. In Europe, due to cultural differences, it may be more desirable to treat underlying causes, and this takes more time to address (Taylor et al., 2004). Nevertheless, the prevalence of stimulant treatment remains over 50% (Hodgkins et al., 2013). However, short term reduction of symptoms is likely not the best metric for evaluating a treatment modality. Other factors should be considered, such as treatment longevity, patient experience, potential side effects, and particularly, the treatment of impulsivity without a diagnosis.

1.2.1 Medicinal Complications

Medications, particularly stimulant-based, are very effective at treating ADHD in the short term, but their effects stop immediately after the pill has run its course (Faraone and Glatt, 2010; Gevensleben et al., 2014a,b). There are many problems associated with this aspect of medication: for example children with ADHD/parents often forget to take/administer the medication. There is a stigma associated with being medicated, as the children/adolescents must take the medication at school, and their peers inevitably discover this and often ridicule the youth (Rappaport et al., 2000; Findling and Dogin, 1998). Beyond this aspect, the children or adolescents themselves are in a stage of life where they are beginning to form an identity, and to be seen as someone sick, or one that relinquishes autonomy to a drug that is often forced on them by parents or physicians, is often more devastating than those same parents and physicians are willing to consider (Barry, Clarke, Johnstone, Mccarthy, & Selikowitz, 2009; Floersch, 2003). The same parents often feel validated by the prescription of medication, that at last a professional has confirmed that their child has a real problem (Floersch, 2003). If this is not enough of a reason to avoid or reduce the use of medication for treating ADHD, there are a plethora of side effects that cause between 8–25% of subjects to discontinue use (Clavenna and Bonati, 2014). The most common of these adverse effects are loss of appetite, headache, and bodily pain. The loss of appetite can lead to significant bodily changes over time. In the seminal Multimodal Treatment Study of Children With ADHD study, stimulant medication led to significant decreases in height and weight over the course of three years of use, compared to healthy controls (Swanson et al., 2007). Add in the high comorbidity of substance use disorder in adolescents and adults with ADHD, and the high abuse potential of, again stimulant-based medicines, and it is clear that alternatives are needed (Winhusen et al., 2011). Furthermore, these medications are not approved to treat highly impulsive symptomatology without a diagnosis of ADHD, meaning that another population of individuals currently does not have any form of treatment.

1.2.2 Alternative Treatments

Tossing medication aside, we are left with several other treatments options: behavioral therapy, cognitive interventions, diet changes, and neurofeedback. Diet changes seem to be more or less ineffective, or lack conclusive evidence to support their use (Faraone et al., 2015). Behavioral therapy is one of the best supported non-pharmacological methods for managing ADHD, but works rather from the perspective of improving the parent-child relationship, and meta-analysis show limited effectiveness in treating the symptoms of ADHD (Faraone et al., 2015). This leaves cognitive interventions and neurofeedback, both of which presumably exert their effects by influencing the complex neuropathological deficits associated with the disorder. Furthermore, as both work on a symptom-level, they are both potentially suitable to treat impulsivity that exists outside of a clinical diagnosis. Cognitive trainings target one or more neurocognitive deficits such a working memory or response inhibition, and train the subject to improve these specific deficits. This in turn is expected to reduce ADHD or impulsive symptomatology, but the main effects are seen as a reduction of the deficits that are trained, with only a small effect on ADHD symptomatology (Cortese et al., 2016a). Furthermore, cognitive trainings are rather treatments by proxy, and do not attempt to target the neurophysiological deficits that may exist. For these reasons, the current dissertation focuses on NF trainings as the best way of moving forward in the treatment of ADHD and impulsivity.

1.2.3 Neurofeedback Overview

NF is, in principle, a simple concept: a particular area of the brain, or parameter of the brain is recorded, and this parameter is fed back to the subject in the form of an auditory or visual cue. The subject should then try to regulate this brain parameter by regulating the cue as per the instructions of the administrator or task. Successful regulation is rewarded via feedback at the end of each trial. Decades of research has given us particular brain regions that are underactive or underdeveloped in ADHD (and also impulsivity), and we aim to strengthen those via repeated trials of NF over repeated sessions. A good analogy for the process of an NF training is a bodybuilder going to the gym: she performs several reps of a particular muscle in one session, and after many sessions she can lift more and more weight with that muscle. The concept would be the same for the NF subject, with the end goal of having a stronger (more flexible) trained brain area or parameter. The problem, as we will shortly see, is that there is no consensus on what the best NF training protocol is; there are as many variables to change as types of potatoes in the world.

1.2.4 EEG Protocols for Neurofeedback

In a recent metanalysis of NF in ADHD, Cortese et al. (2016a) showed that, when taking into account only blinded, 'high-quality'-controlled NF protocols, there was no difference in the effect of NF to that of the control condition. This is a loaded result, as there are many variables that we could pick apart within this analysis alone. For example, the 'high-quality'-control can refer to either sham or active control groups, both of which can refer to multiple different

protocols. In the sham category, one could use another subject's feedback trials as the feedback signal, or one could feed-back a random signal. In an active control group, any number of different brain locations or structures could be fed back, or another, non-cortex feedback could be implemented, such as EMG biofeedback. The myriad possibilities for a 'high-quality' control group alone highlights yet another problem: are they even high quality control groups? Sham control groups can be detrimental to the subjects' motivation, as the subjects become frustrated with their inability to control the training, which, over so many feedback training sessions, can have lasting negative effects (Birbaumer et al., 1991; Gevensleben et al., 2014a; Holtmann et al., 2014a). Active control groups are then perhaps the better choice when trying to achieve a control group that best controls for nonspecific variables, while still retaining the motivation of the participant. Some active control conditions, like EMG-based biofeedback, have even been shown to improve symptomatology (Barth et al., 2017b; Holtmann et al., 2014b). For these reasons, in our lab and in the studies contained in this dissertation, we always use an EMG-based, *musculus supraspinatus*-based control group.

Another critical aspect of NF study design that becomes clear when looking at the Cortese et al. (2016) meta-analysis, is the sheer number of different EEG-protocols available. The authors compare theta-beta frequency (FTB)-based NF, slow-cortical potential (SCP)-based NF, and combined SCP-FTB. Then there are a host of micro-parameters which are also different between the studies, such as: number of feedback sessions, individualized or standardized frequency bandwidths, placement of reference and feedback electrodes, type of reference used, visual presentation of feedback, reinforcement of feedback, etc. It is clear that the field of NF in ADHD lacks a standardized protocol, and furthermore, it is not clear if this standardized protocol should be carried out in EEG, or using another feedback type altogether. For example, recent evidence from advancements in NF protocols, for example using blood-oxygen level dependent (BOLD) feedback from either fMRI or NIRS, offers a promising alternative to EEG.

1.2.5 BOLD-based NF Approaches

Bblood-oxygen level dependent (BOLD)-based NF interventions got their start as fMRI-based protocols. Real-time fMRI NF (rtfMRI) trainings utilize the metabolic process of BOLD activation from a wide range of possible brain areas, due to their superior spatial resolution compared to EEG (Kim and Birbaumer, 2014). Studies have already proven the successful use of rtfMRI in several clinical conditions including major depressive disorder (Zotev et al., 2016), stroke rehabilitation (Wang et al., 2017), schizophrenia, chronic pain (Kim and Birbaumer, 2014) and also in ADHD (Alegria et al., 2017; Zilverstand et al., 2017). The dorsal anterior cingulate cortex (ACC) and the right inferior frontal gyrus (IFG) were used, respectively, as feedback targets for adults and children with ADHD. Only the right IFG method was successful in attenuating ADHD symptomatology, but both studies showed an increase in cognitive functioning specific to the training. Furthermore, rtfMRI was shown to be effective in successful modulation of the DLPFC in just 5 sessions of training for healthy controls (Sherwood et al., 2016). This is much faster than that of typical control of EEG, which takes 10-15 sessions to control. As the

DLPFC is a critical area in ADHD, training it in a faster manner is advantageous. However, while rtfMRI has a superior spatial resolution to that of EEG, it is temporally much slower, so that 'real-time' is rather an unfitting title. rtfMRI has a temporal resolution of between 1-6 seconds, to account for the temporal delay of the hemodynamic response (Albert et al., 2017), while EEG is typically in the range of milliseconds.

One possible way to find a trade-off between the temporal resolution of EEG and the spatial resolution of fMRI is to use NIRS-based NF, which sits nicely in the middle of both methods, having a temporal resolution of 10 Hz and a spatial resolution of $\sim$2 cm (Mehta and Parasuraman, 2013). In our working group, several NIRS-based NF paradigms have been developed, including for schizophrenia, anxiety, ADHD, and even sub-clinical populations such as highly-impulsive adults (Blume et al., 2017; Ehlis et al., 2018a; Hudak et al., 2017). Within the realm of ADHD alone, four studies have either been developed or have already yielded positive results. A pilot study testing the efficacy of DLPFC-based NIRS-based NF in children with ADHD yielded significant positive symptomatic reductions for the NIRS-based NF group, compared to only trends for positive symptomatic reductions in the SCP-based EEG group and the EMG-based control group (Marx et al., 2015). This indicated that NIRS-NF is potentially more quickly effective than SCP-based EEG NF. The next step was to test the efficacy of this NF paradigm in adults with ADHD. This study again showed general symptom improvements that, again, were not differentiable from EEG or EMG-based trainings (Barth et al., 2017a). In fact, this study was the basis for the second manuscript contained within this dissertation (Hudak et al., 2018), and serves to expose how certain flaws within the study design can lead to unreliable results.

1.2.6 Methodological Advancements in NF Approaches

The first and fourth manuscripts in the context of this dissertation concern the final two NIRS-NF studies with ADHD. Both studies seek to advance the possibilities of NF design through implementation of the feedback signal in a virtual classroom. NF studies in general suffer from the artificial laboratory environment where the subjects learn to regulate their brain via feedback in the form of an often arbitrary tone or visual stimulus, such as a ball moving up and down on a screen or a thermometer (Arns et al., 2009). Transfer is the concept of taking what one learns in the NF training and applying it to a real-life setting (Strehl, 2014). For example, the school classroom is an area where both subjects with ADHD and highly impulsive adults suffer from their symptoms (Faraone et al., 2015; Spinella and Miley, 2003). However, subjects may not be able to transfer the skills that they learned in the laboratory into the classroom. To this end, we developed a virtual-reality (VR)-based classroom wherein the subjects use their DLPFC activity to control the level of lighting with the classroom. Instead of controlling the height of an object in 2-D on the computer screen, they need to control how dark or bright the classroom is. This method of feedback works to directly facilitate transfer to the classroom setting while also being non-invasive enough to prevent immersion in the classroom experience. Within our classroom, common distractions seen in daily school life were also implemented to further facilitate this

transfer effect. Furthermore, the teacher in the classroom gives verbal feedback dependent on the success of each feedback trial. The first study in this dissertation concerns a pilot study with highly impulsive college students, in which we aimed to test the efficacy of the VR environment for implementing an NF design. The fourth study in the dissertation is a full-scale study with children with ADHD and contains 3 groups: a 3D-NIRS group, a 2D-NIRS group, and an EMG control group. In this way, we can systematically test if 3D immersive feedback is more advantageous than its 2D counterpart. However, this study is not yet completed, and so it is only partially discussed in this dissertation.

In addition to the implementation of VR, these two studies also examined pre-post brain activation patterns for the first time in NIRS-NF in ADHD, where before there was only subjective assessment of symptoms. In addition, the previously mentioned NF study on adults with ADHD in our lab served as the dataset for the second study in this dissertation, where we examine functional connectivity differences between failed and successful NF trials. This was the first look at such an analysis in NF and led to some interesting implications for study design, which will be discussed in the general discussion.

1.3 Enhancing Information using Multimodal Measurements

In keeping with the dissertation theme of methodological advancements in NIRS-based NF in ADHD, the third and final completed study took advantage of the ability of NIRS to be combined with EEG, and therefore functions as a pilot study for multimodal NF treatment and temporal resolutions. This means that combining them can provide a means to enhance information in the analysis. So-called EEG-informed fMRI analyses have taken advantage of recent advancements in denoising of EEG data to procure event-related potentials (ERP) from single trials that can then be convoluted with the hemodynamic response function used to model single BOLD trials of fMRI to better predict the real BOLD response (Debener et al., 2006). This stems from more basic research which has shown strong correlations between certain ERPs and BOLD activation in certain brain areas. For example, occipital alpha oscillations were anti-correlated with BOLD activation in the visual cortex. Interestingly, these same alpha oscillations correlated positively with activation (increased deoxygenated hemoglobin) in NIRS measurements as well (Moosmann et al., 2003), showing that the two methods can be combined. Another study by Calhoun et al. (2005) showed that performing a combined independent component analysis of the P300 ERP, an ERP strongly elicited in response to target stimuli, and fMRI could reveal precise spatiotemporal patterns of the P300 in a performance-monitoring task. Interestingly, there was a cascade of activation, starting in the brainstem and working its way through frontal and finally temporal lobes. This again shows the utility of combining spatial sensitivity with temporal precision. ERP latencies (Bénar et al., 2007), amplitudes (Debener, 2005), EEG phase coherence (Jann et al., 2009), and EEG frequency band power (Scheeringa et al., 2009) have all been used as part of an EEG-informed analysis, where the respective single trial parameters have been used to inform the fMRI regression, extracting more information than in the basic regression. However, no study to date has made use of an EEG-informed NIRS regression analysis, although the

concept is essentially the same.

1.3.1 Post-error Monitoring in ADHD

Error monitoring, or the constant vigilance for the errors that we make, is a classical problem for people with ADHD. This makes sense when we consider that impulsivity might cause errors and inattention might lead to their oversight. According to Shiels and Hawk (2010), there are two main issues with error-monitoring when it comes to ADHD: early error detection and later-error evaluation. Both problems then make it difficult for the person with ADHD to subsequently engage cognitive control and adapt to the error. This may lead to subsequent problems on post-error trials, but research is in its nascent stages. For example, Yordanova et al. (2011) examined the post-error behavior of juvenile subjects with ADHD and found that they committed more errors following errors and also displayed a higher reaction time variability, indicating some flaw in corrective behavior following errors. Amazingly, to date there are no studies looking at post-error behavior in an adult ADHD population and, perhaps even more importantly, post-error brain functioning. Ehlis et al. (2018b) studied the post-error behavior of adults with ADHD in a combined NIRS-EEG monitored Eriksen flanker task with an embedded go/no-go element. The task was designed to produce a high number of both go and nogo errors to allow for the investigation of post-error behavior. In their initial manuscript, Ehlis et al. (2018b) reported ERP findings only, with a reduced error-related negativity/correct-response negativity (ERN/Ne / CRN) for the ADHD group, compared to healthy controls, in error/correct trials. The ERN/Ne is a large negative event-related potential (ERP) and it is assumed to reflect an automated early error-monitoring process that signals to control regions to be vigilant on the succeeding trial, which has been localized to the ACC in the medial prefrontal cortex (Herrmann et al., 2004). The CRN is a smaller form of the ERN/Ne that occurs on correct trials and has been found to have similar cortical origins (Vidal et al., 2000). The diminished amplitudes in the ADHD group seem to indicate that there was less action-monitoring than in healthy controls. No significant behavioral differences were observed between the groups; however, there were trends for more errors committed and larger reaction time variability (SDRT) in the ADHD group. To further elucidate these group differences, and particularly to look for post-error differences between the groups, we decided to utilize the NIRS data in combination with EEG to maximize information.

In addition to a general-linear model-based regression of NIRS data, and in keeping with the theme outlined in chapter 1.3, we also performed an EEG-informed NIRS regression analysis, in which we used the single-trial wavelet-denoised P300 amplitudes to predict the BOLD response in concurrent trials. This was achieved in effect by modulating the trigger function with the amplitudes of the P300 for each trial and thereby adjusting the height of the hemodynamic response function (HRF) to fit each trial individually. As the P300 is an ERP related to vigilance following a stimulus, with larger amplitudes indicating greater allocated attentional resources, adjusting the HRF to fit the size of the P300 should result in a better fit to cognitive control areas such as the prefrontal cortex and sensorimotor regions that are involved in response preparation.

2 Objectives and Expected Output of the Thesis

The present dissertation seeks to optimize NIRS-based NF paradigms for the treatment of ADHD through examining the following questions: 1) How can NIRS-based NF paradigms be implemented in VR?; 2) What are the underlying mechanisms responsible for a successful NF trial and paradigm?; 3) How can we combine methods to improve NF paradigms even further? To accomplish these aims, we began simply with a VR implementation of already established NIRS-based NF paradigms for ADHD. Studies 1 and 4 seek to implement the first VR-based NF for impulsivity and ADHD that occurs within a virtual-classroom, with the main aim of increasing transfer to a very difficult environment for ADHD and impulsivity. NF itself is chiefly concerned with training cognitive control and an established method for doing so in ADHD. Study 2 takes an already completed NIRS-based NF study of ADHD and seeks to find network activity related to successful and failed NF trials. This is the beginning of the methodological advancement question: how can we improve NF paradigms so that more subjects have a higher success rate, and hopefully decreased symptomatology? Furthermore, we know now that ADHD symptoms are not the result of a single brain region, but rather of networks of activity, so should we be training these instead? This study presented us with several problems during the analysis that allowed us to take an even deeper look into NF study design, and subsequently highlight common problems in design that can easily be avoided. Study 3 looks again at networks involved in successful and failed trials, but from a post-error perspective. Furthermore, it acts as a pilot study to a future multimodal NF paradigm using both NIRS and EEG.

3 General Discussion

Neurofeedback trainings have shown varying success in the treatment of ADHD. A long history of EEG-driven NF paradigms has a rather moderate effect size in the treatment of symptoms, which decreases when considering double-blind, robustly-controlled experimental designs. The intent of this dissertation is to further advance NF technology, attempt to further NIRS-based NF designs and to introduce virtual reality-based designs to NF. We emphasize the importance of being intentional with small aspects of the study design that can have a big influence on the efficacy of the training. Additionally, we explore EEG/NIRS multimodal measurements of cognitive networks within ADHD and how these networks may influence NF trainings.

In order to accomplish the above aims, **Study 1** introduced a new method in DLPFC-based NIRS-based NF integrated in a virtual reality classroom for the first time. The study showed that control of the feedback parameter is indeed possible in a VR environment, and also that the NF seemed to combat impulsiveness. **Study 2** deconstructed the standard NIRS-based NF paradigm for ADHD and looked at proper recommendations for moving forward with NF study design while also examining networks involved in failure and success of NF trials. **Study 3** looked at post-error behavior in ADHD versus HC. Specifically, we explore a P300-informed NIRS analysis intended to fine-tune the differences between ADHD and HC. It is additionally designed to test the efficacy of EEG/NIRS for future use within a multimodal NF study design for ADHD. **Study 4** presents the study design for a large and ongoing study using the paradigm from Study 1 with schoolchildren with ADHD. It offers a glimpse into the massive data collection that should give new insights into NIRS-NF in ADHD.

3.1 Summary of the Individual Studies

Study 1 established and tested a virtual reality classroom setting for the implementation of a NIRS-based NF training. This was the first study to embed the feedback signal of an NF design naturalistically within a virtual reality classroom, using the brightness of the lighting in the room as a feedback display. Highly impulsive college students in the experimental group controlled the overhead lighting using the preprocessed average relative oxygenated hemoglobin (O_2Hb) concentration from their DLPFC. In the biofeedback control group (BF), also composed of matched highly impulsive subjects, electromyography from the *musculus supraspinatus* was used to control the overhead lighting. Subjects were measured pre-training on a go/no-go task, an n-back task, and a stop-signal task. They then underwent 8 sessions of NF or BF training and were assessed once again on the measures listed before. Subjects in the experimental group committed significantly less commission errors, or responses when the cue was to withhold the response, in the post- versus the pretest. This portion of the task required subjects to inhibit prepotent behavioral responses and is especially difficult for impulsive subjects. During the same task, these same subjects showed an increased pre versus post brain activation (increased O_2Hb) in the left hemisphere feedback channels (DLPFC) on the no-go portion of the go/no-go task. Furthermore, the extent to which individual commission errors decreased was significantly correlated with the ability of the subject to learn the NF paradigm. The control group showed

neither a decrease in commission errors, nor an increase of DLPFC activity on the go/no-go task. Neither group showed any behavioral or brain changes during the n-back task, but this was believed to be a ceiling effect, as both groups achieved nearly 100% response rate on the pretest. Finally, in the stop-signal task, there was no effect of stop signal reaction time, but subjects in the experimental group showed a significant decrease in reaction time variability from pre- to posttest.

Study 2 took a deeper look into the construction of a neurofeedback algorithm. The study was essentially an analysis performed on a full dataset from a NIRS-based NF study collected in our working group with adult subjects with ADHD. The subjects (age M = 30.37 years, SD = 9.25; 6 female) underwent 30 sessions of prefrontal NIRS-based NF training and were compared against control groups of both EEG and EMG. The analysis started out with the intent of comparing successful versus failed NF trials and rests preceding trials via a functional connectivity analysis, with the intent of comparing networks involved in successful trials and the rests preceding them. However, during the course of the analysis process, there appeared a series of concerning phenomena related to the study design that both forced a new version of the analysis, and unveiled interesting insights into the effects of commonly used NF parameters.

The first problem was that when we grouped all activation and all deactivation trials together, we noticed that the grand average trials had interesting properties: in the baseline phase wherein the baseline for the upcoming trial is calculated, there was a sharp decrease in feedback channel activity just before activation trials and a sharp increase in feedback channel activity just before deactivation trials. This phenomenon made it easier for subjects to achieve success in both trial types. Upon further examination, it was discovered that the cause of this baseline bias was an improperly randomized trial presentation order. Incongruent trial presentation, such as an activation trial following a deactivation trial or vice versa, was much more likely than a congruent trial presentation, such as an activation trial following an activation trial. This preparatory bias was further confirmed via a greater success rate in incongruent trials versus congruent trials during the first half of NF sessions. Patients adapted behaviorally to this presentation bias, with an equal performance on congruent and incongruent trials in the second half of NF sessions, but the hemodynamic preparation did not go away.

Another knock-on effect of this hemodynamic bias was that we were not able to assess activation and deactivation trials separately, due to the bias strongly affecting functional connectivity calculations. Therefore, we grouped the trial types together for the analysis, which allowed us to look at networks responsible for general feedback control. Interestingly, during failed trials and rests, bilateral DLPFC connectivity was significantly higher, whereas DLPFC–right parietal connectivity was marginally significant. In successful trials and rests, increased right IFG–left parietal connectivity was observed. All of these regions, bilateral DLPFC, IFG, and parietal areas are part of the FCPN, meaning this network is at least partially activated during successful trials. However, in the failed and marginally failed networks, there is also activation of the FCPN, though the regions connected all involved the right DLPFC, which is crucially not part of the feedback channels (the right side was completely composed of IFG). We believe this was due to the Common Average Reference (CAR) that was implemented in the NF algorithm.

The CAR is calculated by subtracting the common average of all channels from the feedback channels, and through this calculation, network activity that could be beneficial for the feedback control is actually punished by being subtracted out of the feedback calculation. Since the right DLPFC was not involved in the feedback, its helpful activity hurt the chances of achieving success during the trials. We then discuss practical ways to improve NF design.

Study 3 examined post-error behavior in adults with ADHD versus healthy controls (HC). To achieve this end, a complex Eriksen Flanker Task with an embedded go/no-go element was implemented during a combined NIRS-EEG measurement. We compared the behavior in trials following errors (post-error trials) versus trials following correct trials (post-correct trials). Importantly, we calculated both a general linear model-based regression analysis of the NIRS data during post-error trials, and pioneered an EEG-informed regression of the NIRS data, in which the P300 amplitude from concurrent trials informed the amplitude of the hemodynamic response function for each trial. Behaviorally, there were no significant post-error differences between the groups, though ADHD subjects showed an overall trend for more errors committed. However, in the general linear model-based analysis approach, the HC group, compared to the ADHD group, showed significantly more activation in the right DLPFC in post-error as compared to post-correct trials. This was an expected finding, as the DLPFC activity is expected to increase in healthy populations following an error, to implement cognitive control for the next trial. Interestingly, when we performed the EEG-informed analysis, these group differences became even more pronounced, showing stronger activation for HC in bilateral DLPFC, right IFG, and premotor, motor, and sensorimotor cortices. These areas are critically involved in both response preparation and cognitive control. Upon further inspection, the ADHD group showed no difference in significance between the EEG-informed and the standard analysis, while the HC group showed more frontal, temporal and parietal activation overall in the EEG-informed analysis. We drew the conclusion that the P300 has a different developmental trajectory in ADHD and may be more likely to have a subcortical generator than in healthy controls. At the same time, adults with ADHD continue to display prefrontal cortical activation deficits, which may influence their overall cognitive control. Still, the lack of behavioral differences between the groups suggests that adults with ADHD may have learned to compensate for their deficits over the lifespan. The successful integration of EEG/NIRS shows us that the additional level of brain process information derived may be beneficial in a multimodal NF design.

Study 4 assesses virtual classroom-based NIRS-based NF as a treatment modality for schoolchildren with ADHD. The 3D-NIRS NF is compared against both a 2D-NIRS NF and a 3D-EMG control group. In this way, we will be able to assess if 3D is helpful or harmful as an NF environment. Furthermore, we will assess a horde of measures pre- and post-NF: (neuro-)cognitive, scholastic, ADHD symptomatology, and training measures will all be assessed. As this study is not yet completed, it will not be discussed per se, but rather referred to as a logical future direction of virtual classroom-based NF therapy.

3.2 Optimization of NIRS-based Neurofeedback Trainings for ADHD

This dissertation has as its chief concern the optimization of NIRS-based neurofeedback trainings for ADHD. NF trainings are not new in the realm of ADHD, but there is much evidence contributing to doubt in the efficacy of feedback trainings. Cortese et al. (2016a) found that, in a meta-analysis of all NF trials that were properly controlled (sham-control; double-blinded) there was no benefit of the NF treatment over that of the control condition. Thibault et al. (2017) argue that NF itself is rather a superplacebo, which means that essentially it has no medical effect on symptom reduction, but because patients have such strong beliefs in its effects, it nevertheless results in a significant reduction in symptoms. This effect can, however, not be differentiated from proper controls. However, in both studies, the focus is on established EEG NF protocols. These protocols typically rely on frequency bandwidths such as theta/beta, where both bandwidths are relatively wide. Such protocols will then have highly varying effects between subjects, who often have an individualized frequency within each band that they respond to optimally. Thibault et al. (2016) argue that EEG may suffer from exactly these problems, and that the spatially superior fMRI and NIRS offer possible future directions for the advancement of NF protocols in ADHD. Furthermore, fMRI-based NF regulation can be learned very quickly in comparison to EEG, an effect that has also been shown in ADHD (Alegria et al., 2017; Kim and Birbaumer, 2014). However, fMRI NF is cumbersome: it is expensive, very susceptible to motion artifacts, and limiting for certain psychiatric diseases (Thibault and Raz, 2016). NIRS offers a much more ergonomic method of providing BOLD-based NF at the cost of spatial resolution. However, as we can see from the studies in this dissertation, it is promising and effective.

Study 1 showed that NIRS-based NF can be effective not only on a symptomatic level, but also that subjects can learn to control the feedback in a short amount of time. Furthermore, the ability of the subjects to control their prefrontal hemodynamics correlated strongly with a reduction in impulsive symptoms. Crucially, we see that in as few as eight sessions of NF over two weeks, both a reduction in impulsive errors and a concurrent increase in activation of the trained feedback area during a cognitively difficult task was possible.

In terms of NF optimization, this study was important for several reasons. Firstly, we saw that there was a link between learning to regulate O_2Hb in the left DLPFC and a reduction in impulsive symptomatology. While the IFG is typically the region most strongly associated with impulsive behavior (Arnsten and Rubia, 2012; Cubillo and Rubia, 2010), in particular within ADHD, several studies carried out in our lab group have shown that regulation of the DLPFC leads to impulsive symptom reduction. The dataset used as the basis for **Study 2** comes from a large NIRS-NF study in adults with ADHD. The only significant improvement in terms of ADHD symptomatology for subjects in the NIRS group, compared with the subjects in the EEG and EMG groups, was a reduction in impulsive symptoms (Barth et al., 2017a). The results of this study have not been formally published, so all we have is conference proceedings. However, as we will see, the feedback channels in this study also contained bilateral IFG, and in fact did not include the right DLPFC at all. Marx et al. (2015) also found an improvement in impulsive

symptoms in a study comparing NIRS-NF with EEG- and EMG-BF in children with ADHD. This study found a reduction in all ADHD symptoms for the NIRS group, but also trends for symptom reduction in the EMG and EEG groups. Taken together, these studies indicate that (left)-DLPFC-based NIRS-NF training seems to reduce impulsive symptomatology whenever it is present, whether this is due to direct influence from the DLPFC, a secondary influence of the DLPFC on neighboring structures, the inclusion of the IFG within the feedback areas, or network activity including the DLPFC is perhaps a future direction that studies should explore.

Interestingly, although there was a correlation between learning to regulate the NF parameter and symptom reduction (**Study 1**), there was no significant learning rate effect for the experimental group. That is, the NIRS group did not learn to systematically control the NF paradigm, while the EMG control group did. This is likely due to the complexity of the two feedback sources. EMG is relatively easy to control as it is a specific body movement pattern that can reliably be controlled once it has been learned, while the NIRS feedback is much more complex as it relies on abstract thought patterns that are unique to the individual. However, the data suggest that on an individual level, the training can be quickly learned. This is further reinforced by the fact that posttest feedback channel activation was increased for the NIRS-NF group during the nogo task, which requires response inhibition. Still, it seems that more than eight sessions of feedback are needed to reliably control prefrontal hemodynamics in NIRS-NF, and for this reason we also increased the number of sessions to 15 in **Study 4**.

Perhaps even more exciting is that with the success of **Study 1**, we know that both NIRS-NF as well as EMG-BF can be implemented inside of a virtual classroom. This was not the first study with virtual classroom-based NF. In fact, Cho et al. (2004) implemented such a design with highly impulsive and inattentive, but sub-clinical adolescents. They observed pre-post differences in impulsivity and inattention but never recorded pre-post brain differences. This becomes even more problematic when we consider that the control group used was a waiting control, so non-specific factors of the NF training can not be ruled out to have had an influence. The study also does not implement the NF naturalistically, but rather uses a separate module, wherein the subjects leaves the classroom setting and a dinosaur egg breaks open with successful regulation. In this sense, **Study 1** was integral in integrating an NF modality that was both salient and not distracting from the overall experience of being in a classroom.

In order to facilitate learning of a behavior, the learning environment should be as similar as possible to the environment where the learned behavior will be recalled (Gershman et al., 2017). As both highly impulsive adults (Spinella and Miley, 2003) and children with ADHD (Daley and Birchwood, 2010) struggle in classroom settings, a classroom would be the ultimate place to conduct a NF training. Strehl (2014) even suggested that bringing the NF training into the classroom would be the best way to facilitate transfer of the learned feedback skill into daily life, where it is urgently needed. While technology may have reached a point where such a training is possible, subjects cannot simply conduct a training while instruction is going on. Therefore, the virtual classroom provides a valuable tool with which to facilitate transfer, but still has the advantages of being carried out in a controlled laboratory setting. Furthermore, **Study 4** makes an attempt to flesh out the full capabilities of the virtual classroom from **Study**

1. In **Study 1** we added triggers to many different trial types that allow us to assess a wide range of data. In **Study 1**, there was no separation of trials distinguishing between trials containing distractions and trials without. Distractions included other students turning and waving, coughing, a truck driving by outside, a lady coming briefly into the room and then leaving, etc. Again, distractions add to the ecological validity of the virtual classroom, and are an important addition for children with ADHD, who have been shown to perform worse in a virtual classroom setting when distractors are introduced (Adams et al., 2009). Since, however, distractors are a part of real classroom environments, their addition, or rather the addition of their triggers, in **Study 4** will allow us to see if training with distractors can indeed improve over many sessions.

Another change which is made in the transition from **Study 1** to **Study 4** is that feedback will be presented in a multimodal format. In a study underlying the importance of verbal feedback on the success of the NF training, (Siniatchkin et al., 2000) found that presenting intentionally incorrect feedback but still verbally reinforcing, via the trainer, correct NF trials, led to successful feedback learning. Strehl (2014) suggests that the verbal reinforcement may actually be a stronger mechanism for learning than the continuous operant conditioning provided by the feedback signal itself. For these reasons, it is often suggested that the trainer cannot be removed from the NF paradigm. However, in the age of automation, there is a demand for exactly that. In **Study 4**, we implement verbal feedback from the virtual teacher when a trial is successful, in addition to the existing smilies presented in **Study 1**. While this cannot replace the function of the NF trainer immediately, it can begin to explore if this is a viable modality for feedback. Eventually, the feedback from the teacher may be combined with automatic artifact detection methods to provide specific feedback when the subject is using an artifact to control the signal. In this way, the NF trainer may be gradually fazed out of NF trainings, which would be particularly useful, with the emergence of do-it-yourself or portable imaging methods, in making NF trainings more accessible to the general public.

Still another advancement from **Study 1** to **Study 4** is the implementation of an adaptive feedback calculation. If subjects perform particularly well in the first block (of 3) in the training, then the remaining two blocks will be more difficult, i.e. require more activation or deactivation for a longer period of time to achieve a successful trial. Conversely, if the subject performs particularly poorly during the first block, the subsequent trials have a decreased threshold for success. This takes advantage of a concept from learning theory called 'shaping', where behavior is viewed as a series of small steps that gradually sum up to the behavior itself (Skinner, 1958). In the context of NF trainings, shaping is used to aid learning of the feedback, both by making it easier to achieve success when the subject is less successful, and harder to achieve success when they are more successful (Sherlin et al., 2011). We believe that the addition of the adaptive feedback calculation will promote learning in both users that struggle with gaining control of the scenario and users that find the scenario unchallenging and need extra motivation. This should also help to address the problem seen is **Study 1**, where participants in the more difficult NIRS group were not all able to gain control of the feedback parameter.

Finally, **Study 4** adds a second experimental group to the study that was not present in

Study 1. The 2D NIRS-based NF group is exactly the same as the virtual reality NIRS-NF group, however the subjects experience the virtual classroom on a computer monitor in 2D. There are many advantages to adding this group to the experiment, for example, we will be able to test whether the virtual reality component is helpful to the learning of the feedback. While the sense of presence in a virtual environment has been shown to be positively correlated with performance on emotional tasks within the environment (Aymerich-Franch, 2010; Riva et al., 2007), no relationship was found between the sense of presence in a virtual classroom and virtual cognitive tests taken within that classroom (Nolin et al., 2016). However, a study comparing a standard continuous performance task (CPT) with a virtual classroom-based CPT showed that the virtual classroom-based scenario was better at differentiating ADHD subjects from healthy controls, in that ADHD subjects made more commission errors and overall errors than controls in the virtual classroom, but not on the standard test (Adams et al., 2009). Therefore, it may be that the virtual classroom that we use is even more distracting for ADHD and leads to worse performance or ability to learn the NF than in the 2D group. It may also be that the 3D group performs worse in the scenario but has a better transfer to daily life, and thus a reduction in teacher and parent rated symptoms, which would be the ultimate goal of the training. Either way, the addition of this group adds an exciting element to **Study 4**.

In summary, **Study 1** served as the basis for the subsequent full implementation of the virtual-reality classroom used in **Study 4**. The study showed that control of the feedback parameter was possible even when implemented in a virtual classroom. Subjects were able, at least on an individual level, to learn the feedback parameter in the NIRS-based NF group, and on a group level in the EMG group. Better control of the feedback parameter in the NIRS group led to a reduction in impulsive behavior and a greater activation in the feedback channels (DLPFC), while these effects were not seen in the control group. Therefore, there appears to also be a specificity to our NF training. **Study 4** advances the paradigm created in **Study 1** in myriad ways, and will allow for the exploration of many questions related to virtual reality and NF trainings. However, **Studies 1 and 4** tread rather lightly in the realm of NF study design, and there remain several questions about NIRS-based NF to answer. In the following section, we take a look at how **Studies 2 and 3** can be used to draw conclusions about basic NF study design as well as complex algorithm creation, in the hopes of creating a more efficient NF training for ADHD.

3.3 Networks of Impulsivity

The NIRS-based NF protocols that we have developed in our working group seem to be specifically good at treating impulsivity symptoms (Barth et al., 2017b; Hudak et al., 2017; Marx et al., 2015). Particularly interesting is the fact that the DLPFC is used as a feedback region in each of these studies, though in Barth et al. (2017b), the IFG is also heavily involved. Typically, the DLPFC is associated with a host of executive functions like working memory, attention focusing, and declarative memory or planning, but is not directly involved in response inhibition like its prefrontal counterpart, the IFG (Cubillo et al., 2012; Peña-Gómez et al., 2012). However, the

DLPFC is a critical component of the frontoparietal control network (FPCN), also known as the cognitive control network, and, due to its fortuitous positioning amongst several other critical prefrontal structures, and its vast connections to regions involved in motor response, it also has a role as mediator in response inhibition (Cieslik et al., 2013; Miller and Cohen, 2001). The FPCN is thought to act as a bridge between two other critical networks that are also diminished in ADHD, the dorsal attention network (DAN) and the default mode network (DMN) (Castellanos and Proal, 2012). The DLPFC is also active in the right DAN, a network that is responsible for coordinating attentional processes for the task at hand (Fox et al., 2006; Rubia et al., 2014). Spreng et al. (2013) even suggest that the DLPFC may be a common link between the DAN and the FPCN. The DLPFC has been shown to be the region with most overlap with the ventral attention system (Fox et al., 2006), as well as the influence that the two attention systems and the FPCN have on the various other systems involved in daily life (i.e. motor system, visual system). In summary, the DLPFC serves as a crucial top-down controller of human behavior (Castellanos and Proal, 2012). In ADHD, there is evidence that the DLPFC is coactivated along with the DMN, which is in direct contrast to healthy controls, where the two are traditionally anticorrelated (Hoekzema et al., 2014). In fact there is a great deal of evidence that the FPCN and the DMN are not strictly anticorrelated in ADHD, but that both infringe upon the other's domain, so that during tasks of cognitive control, the ADHD patient may rather be resting, and *vice versa* (Posner et al., 2014). It is thought that the DLPFC is directly at play here as faulty coordinator, failing to signal to the DAN and the FPCN that they should be inactive during DMN activity (Rubia et al., 2014). Therefore, training the vital DLPFC via NF paradigms makes sense when attempting to combat ADHD, and in particular impulsive symptoms.

However, as we can see in the results of **Study 2**, there are some potentially serious pitfalls to training a region rather than a network, particularly when the CAR is used. In the functional connectivity analysis of successful versus unsuccessful NF trials and rests, we saw that failed trials and rests exhibited a stronger bilateral DLPFC connectivity and a stronger DLPFC to parietal connectivity, whereas in successful NF trials and rests, there was a stronger IFG to parietal connectivity. At first glance, this does not appear to make much sense, as the DLPFC is a critical region in the FPCN, so it should be involved in successful regulation. However, the right DLPFC was not involved in the feedback channels in this study, and because of the CAR, its activation, as well as the activation of the parietal lobes, was subtracted from feedback channel activity. This result demonstrates that one needs to take extra care when selecting a reference region or technique. The CAR is used in many NF studies, both EEG and NIRS-based, as a way to control for probeset or scalp-wide artifacts resulting from biting, head movement, scalp muscle tension, respiration, or other global physiological sources that are not true brain signals (Blume et al., 2017; Egner and Gruzelier, 2004; Hudak et al., 2017; Marx et al., 2015; Mayer et al., 2015; Mayer and Arns, 2016; Schönenberg et al., 2017). The CAR is often used because it is computationally tractable and can easily be applied without slowing the NF algorithm (Nunez and Srinivasan, 2006). However, as Nunez and Srinivasan (2006) argue, the CAR biases the desired signal as a function of the total number of channels or optodes. The more channels or optodes, and the greater the surface of the brain that is covered, the less bias from the CAR

will dilute the true brain signal or interest. Conversely, when one is computing a relatively small number of channels or optodes, such as in an NF experiment, and in particular when these channels are spread over a restricted surface area of the brain, the CAR can potentially be deleterious. While this may seem relatively pedestrian, many NF studies fail to report sufficiently clear methodology such as exact reference (Cortese et al., 2016b). Based on evidence that we have gathered in our lab, we cannot conclude that the CAR is entirely counterproductive. Hudak et al. (2017)(**Study 1**) and Marx et al. (2015) both tried to isolate the DLPFC as the main feedback source using the CAR and both studies showed positive reductions in impulsive symptomatology, with Marx et al. showing global ADHD symptom reduction. However, both studies still had room to improve, for example only impulsive symptoms were improved in Hudak et al., while Marx et al. showed no significant differential between EEG and NIRS-based treatment modalities. Therefore, it may be that isolated DLPFC training works but could perhaps be optimized. Optimization of the reference-feedback relationship could be achieved by either selecting a study-specific reference, or potentially using a connectivity-based feedback algorithm that implements one of the large-scale brain networks outlined by Castellanos and Proal (2012) as particularly affected in ADHD. I will discuss this idea in more detail in Chapter 3.4 *Methodological Innovations*.

3.4 Methodological Innovations

The previously discussed studies (**Studies 1, 2, and 4**) have all implemented some sort of methodological advancement in terms of NF design: **Studies 1 and 4** work toward implementing virtual reality to improve the ecological validity of NF, while **Study 2** attempts to define some standards for the design of NF algorithms in NIRS-based NF paradigms for ADHD. **Study 3** takes methodological innovations one step further and uses the additional information provided by combining NIRS with concurrent EEG, to take full advantage of all temporal and spatial information available. In the normal general linear model based NIRS analysis, we see attenuated activation in the right DLPFC in ADHD subjects compared to healthy controls in post-error trials compared to post-correct trials. In the contrasted brain maps, it is clear to see that this is a result of specifically more activation in post-error trials in HC compared to ADHD, rather than showing egregious differences in the post-correct trials. This means that the DLPFC is underactivated for ADHD subjects in trials following errors. This makes sense because while the anterior cingulate cortex (ACC) is the region most typically associated with error monitoring, the DLPFC is often signaled following errors to facilitate the coordination of cognitive control and attention in subsequent trials (Botvinick et al., 2001; Kerns, 2004; King et al., 2010; Larson et al., 2016). This result is in line with the theory of DLPFC as linchpin between the various task-positive and default mode networks described above in Chapter 4.3 *Networks of Impulsivity*. The ADHD specific result further supports the idea that, in ADHD, the DLPFC is faulty in its role as mediator between the various task-positive and task-negative networks (Rubia et al., 2014). Interestingly, Cieslik et al. (2013) found differential functional connectivity networks between the anterior and posterior parts of the DLPFC and relevant cog-

nitive control regions. They find that the anterior DLPFC is strongly connected with the ACC during error monitoring tasks with a high cognitive demand, such as in the paradigm in **Study 3**. The region of DLPFC that we found to be more active in HC compared to ADHD is indeed in this anterior part, and thus our result may help to elucidate the disconnect between the DLPFC and ACC in ADHD, though the ACC is not directly measurable using NIRS.

Study 3's combined regression analysis of NIRS, using P300 amplitudes to shape individual hemodynamic response functions for concurrent trials, elucidated even more interesting network problems in ADHD, namely that there appears to be a disconnect between the P300 and cortical processes for ADHD subjects, but not for HC. More specifically, the HC group showed a greater activation than ADHD in bilateral DLPFC, IFG, and sensorimotor and motor cortices on post-error compared to post-correct trials. This effect is even more interesting when looking at the single group brain activation maps. The ADHD group shows no changes in activation for either condition when modulating the HRF using P300 amplitudes, whereas the HC group shows a stronger activation across much of the brain limited to the post-error condition. We believe that this difference reflects a delayed and perhaps special maturation of cortical and subcortical structures in ADHD, and thus perhaps also a differential generator of the P300. The exact origins of the P300 are not yet well understood, but it seems that a diffuse mix of tempo-parietal cortex as well as thalamus and hippocampus are responsible for generating the P300 in HC (Key et al., 2005). However, in ADHD, the P300 experiences a maturational delay in that amplitudes are reduced, compared to HC, in juvenile subjects, but these amplitudes then normalize as the subjects become adults (Barry et al., 2003; Lazzaro et al., 1997; Prox et al., 2007). This could be a potential explanation for our result, especially when we consider that adults also show differential deficits compared to children. A meta-analysis of similarly complex inhibition tasks like ours showed subcortical deficits (thalamus and basal ganglia) in children with ADHD compared to HC, while their adult counterparts display deficits in sensorimotor cortex and IFG, both areas elucidated after refining our analysis with P300 amplitudes (Hart et al., 2013). Indeed, our refined analysis is also in line with previous research that shows that adults with ADHD have reduced IFG and DLPFC-frontal-striatal connections on highly complex inhibition tasks. Instead, ADHD subjects tend to compensate with increased occipital and cerebellar activation (Cubillo and Rubia, 2010; Rubia et al., 2014). While we did not measure these areas in our study, this compensatory activity may have been present and responsible for the lack in behavioral differences between the groups. Taken together, our results seem to suggest that there may be a differential maturation, and potentially source and function, of the P300 in ADHD compared to HC.

Furthermore, we again see the DLPFC underactivated in ADHD, this time in conjunction with elements of the motor network and following erroneous responses. The underactivation is in line with behavioral deficits seen post-error in adolescent ADHD (Yordanova et al., 2011). That the motor and sensorimotor areas are also implicated might suggest that there is a strained relationship between the DLPFC and the motor network in ADHD. This would be line with deficits in the FPCN, which is typically impaired in ADHD in motor inhibition tasks (Castellanos and Proal, 2012). The FPCN is perhaps not as efficient in signaling the motor network to remain

vigilant as it is in HC. While the motor network has not been sufficiently studied in ADHD, general impairment is suggested in the motor cortex (Castellanos and Proal, 2012). Again, the DLPFC seems to be a very big problem in many aspects of ADHD behavior.

3.5 Limitations and Future Implications

While the studies forming this dissertation each have their own set of limitations, I will focus here only on the global limitations of the scientific impact of this dissertation, as the individual limitations have already been listed. Firstly, this dissertation seeks to provide guidelines on how to properly build NF studies for ADHD using NIRS-based designs. However, the studies come from a range of different angles: 3 of the studies are NF studies, but one of them is with ADHD children, one with ADHD adults, and one with highly impulsive adults without ADHD. This makes it obviously difficult to draw conclusions about the effects of NF on one of those groups, though all the groups have particular similarities. Indeed, the DLPFC is implicated as deviant in all three groups, and that was the target of each training. Again, in each of these three studies, the sample size was relatively small. This is routinely a problem in NF studies because of the sheer number of hours required to measure one subject. Unfortunately, this makes the field of NF research difficult to advance; so many variables exist that can change in an NF study design, but the length of NF studies means that one needs to be pragmatic in the variables one changes. **Study 3**, does not explicitly looks at NF, but still shows us some interesting relations between the DLPFC and other task-positive networks during a task requiring great cognitive control. Finally, the method of NIRS itself is one limitation in this dissertation. While it has many advantages, NIRS has strong limitations in that it cannot access deeper cortical, or subcortical structures, and that is temporally not as precise as EEG. In **Study 3**, we attempt to combat the latter weakness with a combined measurement, and the results are fruitful in that sense. The spatial resolution, however, does provide an obstacle when one considers that we have learned that networks are vital in ADHD, and some networks contain subcortical structures.

Indeed, in future studies, it would be important to consider the methodology used. As we have seen in this dissertation, it is difficult to completely isolate the DLPFC, as there is always network activity in play during NF trials. Perhaps one way to counter this problem going forward is to build algorithms that take whole network activity into account. For example, one could use a functional connectivity-based algorithm as the feedback source, with the FPCN as the target network. **Study 2** shows that the FPCN seems to be activated anyway during feedback trials. Only a handful of studies have so far been able to implement connectivity-based feedback and they include theoretical healthy control experiments (Koush et al., 2013; Zilverstand et al., 2014), smoking cessation (Kim et al., 2015), stroke (Liew et al., 2016), and anxiety (Zhao et al., 2018). These initial studies have shown promise in being perhaps more effective on behavioral outcomes than pure activation-based methods. However, all these studies were performed with fMRI, and the access that fMRI has to subcortical structures makes it a more viable method for connectivity-based NF. fMRI has a slower temporal resolution than NIRS, meaning there is more time to implement the more complex connectivity-based calculations, however one idea could be

to perform a NIRS-based investigation of the supra-cortical elements of the FCPN and perhaps adjust the number of feedback calculation points to aid the processing of the algorithm. This would combine the potential advantages of connectivity-based NF with the ergonomic advantages of NIRS. Of course, a multimodal feedback source could also be an option for ADHD treatment. For example, as we saw in **Study 4**, there was a disconnect in ADHD subjects between the P300 as measured by EEG, and the cortical activations measured by NIRS. One potential step forward would be to implement both a regulation of the P300 via EEG-NF and a regulation of the accessible FPCN elements via NIRS as previously discussed. This could be implemented relatively easily within a 3D classroom environment, for example, with each feedback source controlling a specific element in the classroom. The P300 feedback could be used to control the onset of the NIRS-based connectivity feedback, for example. A more nuanced knowledge of the relation between the P300 and cortical activation in ADHD would, however, be required to implement such a complex design. It remains to be seen how important the implementation of Virtual Reality will be in NF design for ADHD. It has the potential to be more distracting than necessary, and only after **Study 4** is completed will we have a better idea. It is, however, relatively clear, as evidenced by Study 2, that a few standard procedures should be implemented in all NF designs. The CAR should be used only very intentionally to avoid any deleterious effects of potentially positive network activity not part of the feedback signal. Furthermore, all future NF studies should be controlled for randomization errors. While this might seem like a redundant measure, the clearly observed baseline bias from **Study 2** had severe effects on the outcome of the study. The NF study was designed with commercial software that is used in many published and searchable NF studies. This same baseline bias can come about as the result of intentionally stratifying the presentation of activation and deactivation trials, and thus such designs must also be carefully considered when implementing a local baseline. Finally, **Study 3** shows us how beneficial it can be to combine information from multiple imaging modalities. The combination of NIRS-EEG analysis has not often been done and the possibilities for using ERPs to inform NIRS regression are endless, particularly when one considers potential future use in building NF studies.

3.6 Conclusion

The overarching goal of this dissertation was to take a deeper look inside of the design of NF studies for ADHD, particularly using NIRS. We looked at all aspects of design, from more superficial elements like the randomization of trials or the stratification of feedback trial type, to deeper issues within the design of NF algorithms. We focused on ADHD as being a disorder characterized strongly by impulsivity in terms of response inhibition. We looked at brain networks responsible for producing, or rather failing to prevent, impulsive behavior, and how the DLPFC is a critical linchpin in such systems. And finally, we suggested ways in which we might train these networks and generally build more efficient NF paradigms for the treatment of ADHD.

This was a dissertation rich in the development and exploitation of methodological inter-

ventions. We pioneered the use of the virtual reality classroom for the implementation of NF experiments and we used unconventional methodology to examine what is actually being fedback in an NF experiment. Furthermore, we were able to break down the important elements of NF design. Finally, we were able to harvest critical information that would otherwise have gone missing by combining the methods of EEG and NIRS. Taken together, this dissertation provides a guideline for moving forward in the development of NIRS-based NF for ADHD. In fact, much of what we discuss in this dissertation can be transferred to other NF studies as well, both in terms of different imaging modalities and in terms of designing paradigms for different psychiatric disorders.

References

Adams, R., Finn, P., Moes, E., Flannery, K., and Rizzo, A. S. (2009). Distractibility in Attention/Deficit/ Hyperactivity Disorder (ADHD): The Virtual Reality Classroom. *Child Neuropsychology*, 15(2):120–135.

Aichert, D. S., Wöstmann, N. M., Costa, A., Macare, C., Wenig, J. R., Möller, H.-j., Rubia, K., and Ettinger, U. (2012). Associations between trait impulsivity and prepotent response inhibition. *Journal of Clinical and Experimental Neuropsychology*, 34(10):1016–1032.

Albert, J., Sánchez-Carmona, A. J., Fernández-Jaén, A., and López-Martín, S. (2017). Neurofeedback for ADHD: a Critical Review and Suggested Future Directions. *Current Developmental Disorders Reports*, 4(3):86–93.

Alegria, A. A., Wulff, M., Brinson, H., Barker, G. J., Norman, L. J., Brandeis, D., Stahl, D., David, A. S., Taylor, E., Giampietro, V., and Rubia, K. (2017). Real-time fMRI neurofeedback in adolescents with attention deficit hyperactivity disorder. *Human Brain Mapping*, 38(6):3190–3209.

Arns, M., de Ridder, S., Strehl, U., Breteler, M., and Coenen, A. (2009). Efficacy of Neurofeedback Treatment in ADHD: The Effects on Inattention, Impulsivity and Hyperactivity: A Meta-Analysis. *Clinical EEG and Neuroscience*, 40(3):180–189.

Arnsten, A. F. T. and Rubia, K. (2012). Neurobiological circuits regulating attention, cognitive control, motivation, and emotion: Disruptions in neurodevelopmental psychiatric disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(4):356–367.

Aymerich-Franch, L. (2010). Presence and emotions in playing a group game in a virtual environment: the influence of body participation. *Cyberpsychology, behavior and social networking*, 13(6):649–654.

Balázs, J. and Keresztény, Á. (2014). Subthreshold attention deficit hyperactivity in children and adolescents: a systematic review. *European Child & Adolescent Psychiatry*, 23(6):393–408.

Bari, A. and Robbins, T. W. (2013). Inhibition and impulsivity: behavioral and neural basis of response control. *Progress in neurobiology*, 108:44–79.

Barkley, R. a. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, 121(1):65–94.

Barry, R. J., Clarke, A. R., and Johnstone, S. J. (2003). A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography. *Clinical Neurophysiology*, 114(2):171–183.

Barth, B., Hudak, J., Blume, F., Marx, A., Mayer, K., Strehl, U., Ehlis, A., and Fallgatter, A. (2017a). NIRS neurofeedback in patients with ADHD. *European Neuropsychopharmacology*, 27:S563.

Barth, B., Mayer, K., Strehl, U., Fallgatter, A. J., and Ehlis, A. C. (2017b). EMG biofeedback training in adult attention-deficit/hyperactivity disorder: An active (control) training? *Behavioural Brain Research*, 329(April):58–66.

Bénar, C. G., Schön, D., Grimault, S., Nazarian, B., Burle, B., Roth, M., Badier, J. M., Marquis, P., Liegeois-Chauvel, C., and Anton, J. L. (2007). Single-trial analysis of oddball event-related potentials in simultaneous EEG-fMRI. *Human Brain Mapping*, 28(7):602–613.

Birbaumer, N., Elbert, T., Rockstroh, B., Daum, I., Wolf, P., &., and Canavan, A. (1991). Clinical-psychological treatment of epileptic seizures: a controlled study. In Ehlers, A., editor, *Perspectives and promises in clinical psychology*, pages 81–96. Plenum Press, New York, 1st edition.

Blume, F., Hudak, J., Dresler, T., Ehlis, A.-C., Kühnhausen, J., Renner, T. J., and Gawrilow, C. (2017). NIRS-based neurofeedback training in a virtual reality classroom for children with attention-deficit/hyperactivity disorder: study protocol for a randomized controlled trial. *Trials*, 18(1):41.

Booth, J. R., Burman, D. D., Meyer, J. R., Lei, Z., Trommer, B. L., Davenport, N. D., Li, W., Parrish, T. B., Gitelman, D. R., and Mesulam, M. M. (2005). Larger deficits in brain networks for response inhibition than for visual selective attention in attention deficit hyperactivity disorder (ADHD). *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 46(1):94–111.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict Monitoring and Cognitive Control Despite the importance of these efforts to characterize the func- tion of cognitive control, most of them share an important limita- tion in scope. Most current theories focus nearly exclusively on the. *Psychological Review*, 108(3):624–652.

Bruner, M. R., Kuryluk, A. D., and Whitton, S. W. (2015). Attention-deficit/hyperactivity disorder symptom levels and romantic relationship quality in college students. *Journal of American College Health*, 63(2):98–108.

Calhoun, V., Adali, T., Pearlson, G., and Kiehl, K. (2005). Neuronal chronometry of target detection: Fusion of hemodynamic and event-related potential DATA. *2005 IEEE Workshop on Machine Learning for Signal Processing*, 30:239–244.

Casey, B., Getz, S., and Galvan, A. (2008). The adolecent brain. *Developmental Review*, 28(1):62–77.

Castellanos, F. X. and Proal, E. (2012). Large-scale brain systems in ADHD: beyond the prefrontal–striatal model. *Trends in Cognitive Sciences*, 16(1):17–26.

Cho, B.-H., Kim, S., Shin, D. I., Lee, J. H., Min Lee, S., Young Kim, I., and Kim, S. I. (2004). Neurofeedback Training with Virtual Reality for Inattention and Impulsiveness. *CyberPsychology & Behavior*, 7(5):519–526.

Cieslik, E. C., Zilles, K., Caspers, S., Roski, C., Kellermann, T. S., Jakobs, O., Langner, R., Laird, A. R., Fox, P. T., and Eickhoff, S. B. (2013). Is there one DLPFC in cognitive action control? Evidence for heterogeneity from Co-activation-based parcellation. *Cerebral Cortex*, 23(11):2677–2689.

Clavenna, A. and Bonati, M. (2014). Safety of medicines used for ADHD in children: a review of published prospective clinical trials. *Archives of disease in childhood*, pages 1–7.

Cortese, S., Ferrin, M., Brandeis, D., Holtmann, M., Aggensteiner, P., Daley, D., Santosh, P., Simonoff, E., Stevenson, J., Stringaris, A., Sonuga-Barke, E. J., Asherson, P., Banaschewski, T., Brandeis, D., Buitelaar, J., Coghill, D., Cortese, S., Daley, D., Danckaerts, M., Dittmann, R. W., Döpfner, M., Ferrin, M., Hollis, C., Holtmann, M., Konofal, E., Lecendreux, M., Rothenberger, A., Santosh, P., Sergeant, J. A., Simonoff, E., Sonuga-Barke, E. J., Soutullo, C., Steinhausen, H., Stevenson, J., Stringaris, A., Taylor, E., van der Oord, S., Wong, I., and Zuddas, A. (2016a). Neurofeedback for Attention-Deficit/Hyperactivity Disorder: Meta-Analysis of Clinical and Neuropsychological Outcomes From Randomized Controlled Trials. *Journal of the American Academy of Child & Adolescent Psychiatry*, 55(6):444–455.

Cortese, S., Ferrin, M., Brandeis, D., Holtmann, M., Aggensteiner, P., Daley, D., Santosh, P., Simonoff, E., Stevenson, J., Stringaris, A., Sonuga-Barke, E. J., Asherson, P., Banaschewski, T., Buitelaar, J., Coghill, D., Danckaerts, M., Dittmann, R. W., Döpfner, M., Hollis, C., Konofal, E., Lecendreux, M., Rothenberger, A., Sergeant, J. A., Soutullo, C., Steinhausen, H. C., Taylor, E., van der Oord, S., Wong, I., and Zuddas, A. (2016b). Neurofeedback for Attention-Deficit/Hyperactivity Disorder: Meta-Analysis of Clinical and Neuropsychological Outcomes From Randomized Controlled Trials. *Journal of the American Academy of Child and Adolescent Psychiatry*, 55(6):444–455.

Cubillo, A., Halari, R., Ecker, C., Giampietro, V., Taylor, E., and Rubia, K. (2010). Reduced activation and inter-regional functional connectivity of fronto-striatal networks in adults with childhood Attention-Deficit Hyperactivity Disorder (ADHD) and persisting symptoms during tasks of motor inhibition and cognitive switching. *Journal of Psychiatric Research*, 44(10):629–639.

Cubillo, A., Halari, R., Smith, A., Taylor, E., and Rubia, K. (2012). A review of fronto-striatal and fronto-cortical brain abnormalities in children and adults with Attention Deficit Hyperactivity Disorder (ADHD) and new evidence for dysfunction in adults with ADHD during motivation and attention. *Cortex*, 48(2):194–215.

Cubillo, A. I. and Rubia, K. (2010). Structural and functional brain imaging in adult attention-deficit/hyperactivity disorder. *Expert review of neurotherapeutics*, 10(4):603–620.

Daley, D. and Birchwood, J. (2010). ADHD and academic performance: why does ADHD impact on academic performance and what can be done to support ADHD children in the classroom? *Child: Care, Health and Development*, 36(4):455–464.

Debener, S. (2005). Trial-by-Trial Coupling of Concurrent Electroencephalogram and Functional Magnetic Resonance Imaging Identifies the Dynamics of Performance Monitoring. *Journal of Neuroscience*, 25(50):11730–11737.

Debener, S., Ullsperger, M., Siegel, M., and Engel, A. K. (2006). Single-trial EEG–fMRI reveals the dynamics of cognitive function. *Trends in Cognitive Sciences*, 10(12):558–563.

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1):135–168.

Ding, W.-n., Sun, J.-h., Sun, Y.-W., Chen, X., Zhou, Y., Zhuang, Z.-g., Li, L., Zhang, Y., Xu, J.-r., and Du, Y.-s. (2014). Trait impulsivity and impaired prefrontal impulse inhibition function in adolescents with internet gaming addiction revealed by a Go/No-Go fMRI study. *Behavioral and Brain Functions*, 10(1):20.

Doernberg, E. and Hollander, E. (2016). Neurodevelopmental Disorders (ASD and ADHD): DSM-5, ICD-10, and ICD-11. *CNS Spectrums*, 21(4):295–299.

Egner, T. and Gruzelier, J. H. (2004). The Temporal Dynamics of Electroencephalographic Responses to Alpha / Theta Neurofeedback Training in Healthy Subjects. *Journal of Neurotherapy*, 8(1):37–41.

Ehlis, A.-C., Barth, B., Hudak, J., Storchak, H., Weber, L., Kimmig, A.-C. S., Kreifelts, B., Dresler, T., and Fallgatter, A. J. (2018a). Near-Infrared Spectroscopy as a New Tool for Neurofeedback Training: Applications in Psychiatry and Methodological Considerations. *Japanese Psychological Research*, 60(4):225–241.

Ehlis, A.-C., Deppermann, S., and Fallgatter, A. J. (2018b). Performance monitoring and post-error adjustments in adults with attention-deficit/hyperactivity disorder: an EEG analysis. *Journal of psychiatry & neuroscience : JPN*, 43(5):170118.

Fallgatter, A. J. and Herrmann, M. J. (2001). Electrophysiological assessment of impulsive behavior in healthy subjects. *Neuropsychologia*, 39(3):328–333.

Faraone, S. V., Asherson, P., Banaschewski, T., Biederman, J., Buitelaar, J. K., Ramos-Quiroga, J. A., Rohde, L. A., Sonuga-Barke, E. J. S., Tannock, R., and Franke, B. (2015). Attention-deficit/hyperactivity disorder. *Nature Reviews Disease Primers*, 1:15020.

Faraone, S. V. and Glatt, S. J. (2010). A Comparison of the Efficacy of Medications for Adult Attention-Deficit/Hyperactivity Disorder Using Meta-Analysis of Effect Sizes. *The Journal of Clinical Psychiatry*, 71(6):754–763.

Farr, O. M., Hu, S., Zhang, S., and Li, C. s. R. (2012). Decreased saliency processing as a neural measure of Barratt impulsivity in healthy adults. *NeuroImage*, 63(3):1070–1077.

Findling, R. L. and Dogin, J. W. (1998). Psychopharmacology of ADHD: children and adolescents. *The Journal of clinical psychiatry*, 59(suppl7):42–49.

Floersch, J. (2003). The Subjective Experience of Youth Psychotropic Treatment. *Social Work in Mental Health*, 1(4):51–69.

Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., and Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences*, 103(26):10046–10051.

Geburek, A. J., Rist, F., Gediga, G., Stroux, D., and Pedersen, A. (2013). Electrophysiological indices of error monitoring in juvenile and adult attention deficit hyperactivity disorder (ADHD)-A meta-analytic appraisal. *International Journal of Psychophysiology*, 87(3):349–362.

Gershman, S. J., Pouncy, H. T., and Gweon, H. (2017). Learning the Structure of Social Influence. *Cognitive Science*, 41:545–575.

Gevensleben, H., Kleemeyer, M., Rothenberger, L. G., Studer, P., Flaig-Röhr, A., Moll, G. H., Rothenberger, A., and Heinrich, H. (2014a). Neurofeedback in ADHD: further pieces of the puzzle. *Brain topography*, 27(1):20–32.

Gevensleben, H., Moll, G. H., Rothenberger, A., and Heinrich, H. (2014b). Neurofeedback in attention-deficit/hyperactivity disorder - different models, different ways of application. *Frontiers in human neuroscience*, 8(October):846.

Hart, H., Radua, J., Nakao, T., Mataix-Cols, D., and Rubia, K. (2013). Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: Exploring task-specific, stimulant medication, and age effects. *JAMA Psychiatry*, 70(2):185–198.

Hege, M. a., Stingl, K. T., Kullmann, S., Schag, K., Giel, K. E., Zipfel, S., and Preissl, H. (2014). Attentional impulsivity in binge eating disorder modulates response inhibition performance and frontal brain networks. *International journal of obesity (2005)*, 39(April):1–30.

Henges, A. L. and Marczinski, C. A. (2012). Impulsivity and alcohol consumption in young social drinkers. *Addictive Behavior*, 2(37):217–220.

Herrmann, M. J., Römmler, J., Ehlis, A. C., Heidrich, A., and Fallgatter, A. J. (2004). Source localization (LORETA) of the error-related-negativity (ERN/Ne) and positivity (Pe). *Cognitive Brain Research*, 20(2):294–299.

Hodgkins, P., Setyawan, J., Mitra, D., Davis, K., Quintero, J., Fridman, M., Shaw, M., and Harpin, V. (2013). Management of ADHD in children across Europe: Patient demographics, physician characteristics and treatment patterns. *European Journal of Pediatrics*, 172(7):895–906.

Hoekzema, E., Carmona, S., Ramos-Quiroga, J. A., Richarte Fernández, V., Bosch, R., Soliva, J. C., Rovira, M., Bulbena, A., Tobeña, A., Casas, M., and Vilarroya, O. (2014). An independent components and functional connectivity analysis of resting state fMRI data points to neural network dysregulation in adult ADHD. *Human Brain Mapping*, 35(4):1261–1272.

Holtmann, M., Pniewski, B., Wachtlin, D., Wörz, S., and Strehl, U. (2014a). Neurofeedback in children with attention-deficit/hyperactivity disorder (ADHD) – a controlled multicenter study of a non-pharmacological treatment approach. *BMC Pediatrics*, 14(1):202.

Holtmann, M., Sonuga-Barke, E., Cortese, S., and Brandeis, D. (2014b). Neurofeedback for ADHD. A Review of Current Evidence. *Child and adolescent psychiatric clinics of North America*, 23:789–806.

Horn, N. R., Dolan, M., Elliott, R., Deakin, J. F. W., and Woodruff, P. W. R. (2003). Response inhibition and impulsivity: An fMRI study. *Neuropsychologia*, 41(14):1959–1966.

Hudak, J., Blume, F., Dresler, T., Haeussinger, F. B., Renner, T. J., Fallgatter, A. J., Gawrilow, C., and Ehlis, A.-C. (2017). Near-Infrared Spectroscopy-Based Frontal Lobe Neurofeedback Integrated in Virtual Reality Modulates Brain and Behavior in Highly Impulsive Adults . *Frontiers in Human Neuroscience*, 11(September):425.

Hudak, J., Rosenbaum, D., Barth, B., Fallgatter, A. J., and Ehlis, A.-C. (2018). Functionally disconnected: A look at how study design influences neurofeedback data and mechanisms in attention-deficit/hyperactivity disorder. *PLOS ONE*, 13(8):e0200931.

Jann, K., Dierks, T., Boesch, C., Kottlow, M., Strik, W., and Koenig, T. (2009). BOLD correlates of EEG alpha phase-locking and the fMRI default mode network. *NeuroImage*, 45(3):903–916.

Kerns, J. G. (2004). Anterior Cingulate Conflict Monitoring and Adjustments in Control. *Science*, 303(5660):1023–1026.

Key, A. P. F., Dove, G. O., and Maguire, M. J. (2005). Linking Brainwaves to the Brain: An ERP Primer. *Developmental Neuropsychology*, 27(2):183–215.

Kim, D.-Y., Yoo, S.-S., Tegethoff, M., Meinlschmidt, G., and Lee, J.-H. (2015). The Inclusion of Functional Connectivity Information into fMRI-based Neurofeedback Improves Its Efficacy in the Reduction of Cigarette Cravings. *Journal of Cognitive Neuroscience*, 27(8):1552–1572.

Kim, S. and Birbaumer, N. (2014). Real-time functional MRI neurofeedback: a tool for psychiatry. *Current opinion in psychiatry*, 27(5):332–336.

King, J. A., Korb, F. M., von Cramon, D. Y., and Ullsperger, M. (2010). Post-Error Behavioral Adjustments Are Facilitated by Activation and Suppression of Task-Relevant and Task-Irrelevant Information Processing. *Journal of Neuroscience*, 30(38):12759–12769.

Koush, Y., Rosa, M. J., Robineau, F., Heinen, K., W. Rieger, S., Weiskopf, N., Vuilleumier, P., Van De Ville, D., and Scharnowski, F. (2013). Connectivity-based neurofeedback: Dynamic causal modeling for real-time fMRI. *NeuroImage*, 81:422–430.

Larson, M. J., Clayson, P. E., Keith, C. M., Hunt, I. J., Hedges, D. W., Nielsen, B. L., and Call, V. R. (2016). Cognitive control adjustments in healthy older and younger adults: Conflict adaptation, the error-related negativity (ERN), and evidence of generalized decline with age. *Biological Psychology*, 115:50–63.

Lazzaro, I., Anderson, J., Gordon, E., Clarke, S., Leong, J., and Meares, R. (1997). Single trial variability within the P300 (250-500 ms) processing window in adolescents with attention deficit hyperactivity disorder. *Psychiatry Research*, 73:91–101.

Liew, S. L., Rana, M., Cornelsen, S., Fortunato De Barros Filho, M., Birbaumer, N., Sitaram, R., Cohen, L. G., and Soekadar, S. R. (2016). Improving Motor Corticothalamic Communication after Stroke Using Real-Time fMRI Connectivity-Based Neurofeedback. *Neurorehabilitation and Neural Repair*, 30(7):671–675.

Lijffijt, M., Kenemans, J. L., Verbaten, M. N., and van Engeland, H. (2005). A meta-analytic review of stopping performance in attention-deficit/hyperactivity disorder: deficient inhibitory motor control? *Journal of Abnormal Psychology*, 114(2):216–222.

Logan, G. D., Schachar, R. J., and Tannock, R. (1997). Impulsivity and Inhibitoey Control. *Psychological Science*, 8(1):60–64.

Marx, A.-M., Ehlis, A.-C., Furdea, A., Holtmann, M., Banaschewski, T., Brandeis, D., Rothenberger, A., Gevensleben, H., Freitag, C. M., Fuchsenberger, Y., Fallgatter, A. J., and Strehl, U. (2015). Near-infrared spectroscopy (NIRS) neurofeedback as a treatment for children with attention deficit hyperactivity disorder (ADHD) – a pilot study.

Mayer, K. and Arns, M. (2016). Electroencephalogram Neurofeedback: Application in ADHD and Epilepsy. *Psychiatric Annals*, 46:594–600.

Mayer, K., Wyckoff, S. N., Fallgatter, A. J., Ehlis, A.-C., and Strehl, U. (2015). Neurofeedback as a nonpharmacological treatment for adults with attention-deficit/hyperactivity disorder (ADHD): study protocol for a randomized controlled trial. *Trials*, 16(1):174.

Mehta, R. K. and Parasuraman, R. (2013). Neuroergonomics: a review of applications to physical and cognitive work. *Frontiers in Human Neuroscience*, 7(December):1–10.

Miller, E. K. and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, 24:167–202.

Moosmann, M., Ritter, P., Krastel, I., Brink, A., Thees, S., Blankenburg, F., Taskin, B., Obrig, H., and Villringer, A. (2003). Correlates of alpha rhythm in functional magnetic resonance imaging and near infrared spectroscopy. *NeuroImage*, 20(1):145–158.

Nolin, P., Stipanicic, A., Henry, M., Lachapelle, Y., Lussier-Desrochers, D., Rizzo, A. S., and Allain, P. (2016). ClinicaVR: Classroom-CPT: A virtual reality tool for assessing attention and inhibition in children and adolescents. *Computers in Human Behavior*, 59:327–333.

Nunez, P. L. and Srinivasan, R. (2006). *Electric Fields of the Brain*. Oxford University Press.

Patton, J. H., Stanford, M. S., and Barratt, E. S. (1995). Patton Factor Structure of the BIS.pdf.

Peña-Gómez, C., Sala-Lonch, R., Junqué, C., Clemente, I. C., Vidal, D., Bargalló, N., Falcón, C., Valls-Solé, J., Pascual-Leone, Á., and Bartrés-Faz, D. (2012). Modulation of large-scale brain networks by transcranial direct current stimulation evidenced by resting-state functional MRI. *Brain Stimulation*, 5(3):252–263.

Polanczyk, G. (2007). The Worldwide Prevalence of ADHD: A Systematic Review and Metaregression Analysis. *American Journal of Psychiatry*, 164(6):942.

Posner, J., Park, C., and Wang, Z. (2014). Connecting the dots: A review of resting connectivity MRI studies in attention-deficit/hyperactivity disorder. *Neuropsychology Review*, 24(1):3–15.

Prox, V., Dietrich, D. E., Zhang, Y., Emrich, H. M., and Ohlmeier, M. D. (2007). Attentional processing in adults with ADHD as reflected by event-related potentials. *Neuroscience Letters*, 419(3):236–241.

Rappaport, N., Chubinsky, P., and Jellinek, M. S. (2000). The meaning of psychotropic medications for children, adolescents, and their families. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39(9):1198–1200.

Riva, G., Mantovani, F., Capideville, C. S., Preziosa, A., Morganti, F., Villani, D., Gaggioli, A., Botella, C., and Alcañiz, M. (2007). Affective Interactions Using Virtual Reality: The Link between Presence and Emotions. *CyberPsychology & Behavior*, 10(1):45–56.

Rubia, K., Alegria, A., and Brinson, H. (2014). Imaging the ADHD brain: disorder-specificity, medication effects and clinical translation. *Expert Review of Neurotherapeutics*, 14(5):519–538.

Rubia, K., Halari, R., Cubillo, A., Mohammad, A. M., Scott, S., and Brammer, M. (2010). Disorder-specific inferior prefrontal hypofunction in boys with pure attention-deficit/hyperactivity disorder compared to boys with pure conduct disorder during cognitive flexibility. *Human Brain Mapping*, 31(12):1823–1833.

Rubia, K., Halari, R., Smith, A. B., Mohammad, M., Scott, S., and Brammer, M. J. (2009). Shared and disorder-specific prefrontal abnormalities in boys with pure attention-deficit/hyperactivity disorder compared to boys with pure CD during interference inhibition and attention allocation. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 50(6):669–678.

Rubia, K., Lim, L., Ecker, C., Halari, R., Giampietro, V., Simmons, A., Brammer, M., and Smith, A. (2013). Effects of age and gender on neural networks of motor response inhibition: From adolescence to mid-adulthood. *NeuroImage*, 83:690–703.

Rubia, K., Smith, a., and Taylor, E. (2007). Performance of children with attention deficit hyperactivity disorder (ADHD) on a test battery of impulsiveness. *Child neuropsychology*, 13(3):276–304.

Scheeringa, R., Petersson, K. M., Oostenveld, R., Norris, D. G., Hagoort, P., and Bastiaansen, M. C. M. (2009). Trial-by-trial coupling between EEG and BOLD identifies networks related to alpha and theta EEG power increases during working memory maintenance. *NeuroImage*, 44(3):1224–1238.

Schönenberg, M., Wiedemann, E., Schneidt, A., Scheeff, J., Logemann, A., Keune, P. M., and Hautzinger, M. (2017). Neurofeedback, sham neurofeedback, and cognitive-behavioural group therapy in adults with attention-deficit hyperactivity disorder: a triple-blind, randomised, controlled trial. *The Lancet Psychiatry*, 0366(17):1–12.

Seidman, L. J. (2006). Neuropsychological functioning in people with ADHD across the lifespan. *Clinical Psychology Review*, 26(4):466–485.

Shaw, P., Eckstrand, K., Sharp, W., Blumenthal, J., Lerch, J. P., Greenstein, D., Clasen, L., Evans, A., Giedd, J., and Rapoport, J. L. (2007). Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(49):19649–54.

Shaw-Zirt, B., Popali-Lehane, L., Chaplin, W., and Bergman, A. (2005). Adjustment, Social Skills, and Self-Esteem in College Students With Symptoms of ADHD. *Journal of Attention Disorders*, 8(3):109–120.

Sherlin, L. H., Arns, M., Lubar, J., Heinrich, H., Kerson, C., Strehl, U., and Sterman, M. B. (2011). Neurofeedback and Basic Learning Theory: Implications for Research and Practice. *Journal of Neurotherapy*, 15(4):292–304.

Sherwood, M. S., Kane, J. H., Weisend, M. P., and Parker, J. G. (2016). Enhanced control of dorsolateral prefrontal cortex neurophysiology with real-time functional magnetic resonance imaging (rt-fMRI) neurofeedback training and working memory practice. *NeuroImage*, 124:214–223.

Shiels, K. and Hawk, L. W. (2010). Self-regulation in ADHD: The role of error processing. *Clinical Psychology Review*, 30(8):951–961.

Shulman, E. P., Smith, A. R., Silva, K., Icenogle, G., Duell, N., Chein, J., and Steinberg, L. (2016). The dual systems model: Review, reappraisal, and reaffirmation. *Developmental Cognitive Neuroscience*, 17:103–117.

Siniatchkin, M., Kropp, P., and Gerber, W. D. (2000). Neurofeedback - The significance of reinforcement and the search for an appropriate strategy for the success of self-regulation. *Applied Psychophysiology Biofeedback*, 25(3):167–175.

Skinner, B. F. (1958). Diagramming Schedules of Reinforcement. *Journal of the Experimental Analysis of Behavior*, 1(1):67–68.

Smith, J. L., Johnstone, S. J., and Barry, R. J. (2004). Inhibitory processing during the Go/NoGo task: An ERP analysis of children with attention-deficit/hyperactivity disorder. *Clinical Neurophysiology*, 115(6):1320–1331.

Spinella, M. and Miley, W. M. (2003). Impulsivity and academic achievement in college students. *College Student Journal*, 37(4):545–549.

Spreng, R. N., Sepulcre, J., Turner, G. R., Stevens, W. D., and Schacter, D. L. (2013). Intrinsic Architecture Underlying the Relations among the Default, Dorsal Attention, and Frontoparietal Control Networks of the Human Brain. *Journal of Cognitive Neuroscience*, 25(1):74–86.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review*, 28(1):78–106.

Strehl, U. (2014). What learning theories can teach us in designing neurofeedback treatments. *Frontiers in human neuroscience*, 8(November):894.

Swanson, J. M., Elliott, G. R., Greenhill, L. L., Wigal, T., Arnold, L. E., Vitiello, B., Hechtman, L., Epstein, J. N., Pelham, W. E., Abikoff, H. B., Newcorn, J. H., Molina, B. S., Hinshaw, S. P., Wells, K. C., Hoza, B., Jensen, P. S., Gibbons, R. D., Hur, K., Stehli, A., Davies, M., March, J. S., Conners, C. K., Caron, M., and Volkow, N. D. (2007). Effects of stimulant medication on growth rates across 3 years in the MTA follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46(8):1015–1027.

Tannock, R. (1998). Attention Deficit Hyperactivity Disorder: Advances in Cognitive, Neurobiological, and Genetic Research. *Journal of Child Psychology and Psychiatry*, 39(1):65–99.

Taylor, E., Doepfner, M., Sergeant, J., Asherson, P., Banaschewski, T., Buitelaar, J., Coghill, D., Danckaerts, M., Rothenberger, A., Sonuga-Barke, E., Steinhausen, H.-C., and Zuddas, A. (2004). European clinical guidelines for hyperkinetic disorder ? first upgrade. *European Child & Adolescent Psychiatry*, 13(S1).

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

Thibault, R. T., Lifshitz, M., and Raz, A. (2017). Neurofeedback or neuroplacebo? *Brain*, 140(4):862–864.

Thibault, R. T. and Raz, A. (2016). When can neurofeedback join the clinical armamentarium? *The Lancet Psychiatry*, 3(6):497–498.

Vidal, F., Hasbroucq, T., Grapperon, J., and Bonnet, M. (2000). Is the 'error negativity' specific to errors? *Biological Psychology*, 51(2-3):109–128.

Wang, T., Mantini, D., and Gillebert, C. R. (2017). The potential of real-time fMRI neurofeedback for stroke rehabilitation: A systematic review. *Cortex*, 107:148–165.

Wehmeier, P. M., Schacht, A., and Barkley, R. A. (2010). Social and Emotional Impairment in Children and Adolescents with ADHD and the Impact on Quality of Life. *Journal of Adolescent Health*, 46(3):209–217.

Weidacker, K., Whiteford, S., Boy, F., and Johnston, S. J. (2016). Response Inhibition in the Parametric Go/No-Go Task and its Relation to Impulsivity and Subclinical Psychopathy. *Quarterly journal of experimental psychology (2006)*, 0218(May):1–52.

Winhusen, T. M., Lewis, D. F., Riggs, P. D., Davies, R. D., Adler, L. A., Sonne, S., and Somoza, E. C. (2011). Subjective Effects, Misuse, and Adverse Effects of Osmotic-Release Methylphenidate Treatment in Adolescent Substance Abusers with Attention-Deficit/Hyperactivity Disorder. *Journal of Child and Adolescent Psychopharmacology*, 21(5):455–463.

Yordanova, J., Kolev, V., Albrecht, B., Uebel, H., Banaschewski, T., and Rothenberger, A. (2011). May posterror performance be a critical factor for behavioral deficits in attention-deficit/hyperactivity disorder? *Biological Psychiatry*, 70(3):246–254.

Zhao, Z., Yao, S., Li, K., Sindermann, C., Zhou, F., Zhao, W., and Becker, B. (2018). Real-time functional connectivity-based neurofeedback of amygdala-frontal pathways reduces anxiety. *bioRxiv*.

Zilverstand, A., Sorger, B., Slaats-Willemse, D., Kan, C. C., Goebel, R., and Buitelaar, J. K. (2017). fMRI Neurofeedback Training for Increasing Anterior Cingulate Cortex Activation in Adult Attention Deficit Hyperactivity Disorder. An Exploratory Randomized, Single-Blinded Study. *Plos One*, 12(1):e0170795.

Zilverstand, A., Sorger, B., Zimmermann, J., Kaas, A., and Goebel, R. (2014). Windowed correlation: A suitable tool for providing dynamic fMRI-based functional connectivity neurofeedback on task difficulty. *PLoS ONE*, 9(1).

Zotev, V., Yuan, H., Misaki, M., Phillips, R., Young, K. D., Feldner, M. T., and Bodurka, J. (2016). Correlation between amygdala BOLD activity and frontal EEG asymmetry during real-time fMRI neurofeedback training in patients with depression. *NeuroImage: Clinical*, 11:224–238.

4 List of Publications and Author's Contributions

4.1 Accepted Manuscripts

Study 1 Near-Infrared Spectroscopy-Based Frontal Lobe Neurofeedback Integrated in Virtual Reality Modulates Brain and Behavior in Highly Impulsive Adults

Justin Hudak: study design, data collection/analysis, manuscript writing/editing.
Friederike Blume: study design, data collection, manuscript editing
Thomas Dresler: study design, study conception, data analysis, manuscript editing
Florian Hauessinger: data analysis
Tobias Renner: study design and consultation, manuscript editing
Andreas Fallgatter: study design, provision of resources, manuscript editing
Caterina Gawrilow: study design and consultation, manuscript editing
Ann-Christine Ehlis: study design and consultation, manuscript editing

Study 2 Combined near-infrared spectroscopy / EEG measurement elucidates functional deficits in post-error processing in adults with ADHD

Justin Hudak: study conception, data collection/analysis, manuscript writing/editing.
David Rosenbaum: study design, data analysis, manuscript editing
Beatrix Barth: study design, data collection, manuscript editing
Andreas Fallgatter: provision of resources, manuscript editing
Ann-Christine Ehlis: study design and consultation, manuscript editing

4.2 Submitted Manuscripts

Study 3 Combined near-infrared spectroscopy / EEG measurement elucidates functional deficits in post-error processing in adults with ADHD

Justin Hudak: study conception, data collection/analysis, manuscript writing/editing.
Saskia Deppermann: study design, data collection, manuscript editing
Andreas Fallgatter: provision of resources, manuscript editing
Ann-Christine Ehlis: study design and consultation, manuscript writing/editing

4.3 Supplementary Manuscripts

Study 4 NIRS-based neurofeedback training in a virtual reality classroom for children with attention-deficit/hyperactivity disorder: study protocol for a randomized controlled trial

Friederike Blume: study design/conception,primary manuscript writing/editing

Justin Hudak: study design/conception, manuscript writing/editing.
Thomas Dresler: study design/conception, manuscript editing
Ann-Christine Ehlis: study design/conception, manuscript editing
Jan Kuehnhausen: study design/conception, manuscript editing
Tobias Renner: study design/conception, manuscript editing
Caterina Gawrilow: study design and consultation, manuscript editing

5 Publications

5.1 Accepted Manuscripts

frontiers in Human Neuroscience

ORIGINAL RESEARCH
published: 04 September 2017
doi: 10.3389/fnhum.2017.00425

Near-Infrared Spectroscopy-Based Frontal Lobe Neurofeedback Integrated in Virtual Reality Modulates Brain and Behavior in Highly Impulsive Adults

Justin Hudak[1], Friederike Blume[1], Thomas Dresler[1,2], Florian B. Haeussinger[2], Tobias J. Renner[3], Andreas J. Fallgatter[1,2,4], Caterina Gawrilow[1,5,6] and Ann-Christine Ehlis[1,2]*

[1] LEAD Graduate School & Research Network, University of Tübingen, Tübingen, Germany, [2] Department of Psychiatry and Psychotherapy, University Hospital Tübingen, Tübingen, Germany, [3] Department of Child and Adolescence Psychiatry, University Hospital Tübingen, Tübingen, Germany, [4] Center for Integrative Neuroscience, University of Tübingen, Tübingen, Germany, [5] Department of Psychology, University of Tübingen, Tübingen, Germany, [6] Center for Individual Development and Adaptive Education of Children at Risk, Goethe University Frankfurt, Frankfurt, Germany

OPEN ACCESS

Edited by:
Soledad Ballesteros,
Universidad Nacional de Educación a Distancia (UNED), Spain

Reviewed by:
Sabrina Brigadoi,
University of Padua, Italy
Hasan Ayaz,
Drexel University, United States

***Correspondence:**
Justin Hudak
justin.hudak@lead.uni-tuebingen.de

Received: 12 May 2017
Accepted: 08 August 2017
Published: 04 September 2017

Citation:
Hudak J, Blume F, Dresler T, Haeussinger FB, Renner TJ, Fallgatter AJ, Gawrilow C and Ehlis A-C (2017) Near-Infrared Spectroscopy-Based Frontal Lobe Neurofeedback Integrated in Virtual Reality Modulates Brain and Behavior in Highly Impulsive Adults. Front. Hum. Neurosci. 11:425. doi: 10.3389/fnhum.2017.00425

Based on neurofeedback (NF) training as a neurocognitive treatment in attention-deficit/hyperactivity disorder (ADHD), we designed a randomized, controlled functional near-infrared spectroscopy (fNIRS) NF intervention embedded in an immersive virtual reality classroom in which participants learned to control overhead lighting with their dorsolateral prefrontal brain activation. We tested the efficacy of the intervention on healthy adults displaying high impulsivity as a sub-clinical population sharing common features with ADHD. Twenty participants, 10 in an experimental and 10 in a shoulder muscle-based electromyography control group, underwent eight training sessions across 2 weeks. Training was bookended by a pre- and post-test including go/no-go, n-back, and stop-signal tasks (SST). Results indicated a significant reduction in commission errors on the no-go task with a simultaneous increase in prefrontal oxygenated hemoglobin concentration for the experimental group, but not for the control group. Furthermore, the ability of the subjects to gain control over the feedback parameter correlated strongly with the reduction in commission errors for the experimental, but not for the control group, indicating the potential importance of learning feedback control in moderating behavioral outcomes. In addition, participants of the fNIRS group showed a reduction in reaction time variability on the SST. Results indicate a clear effect of our NF intervention in reducing impulsive behavior possibly via a strengthening of frontal lobe functioning. Virtual reality additions to conventional NF may be one way to improve the ecological validity and symptom-relevance of the training situation, hence positively affecting transfer of acquired skills to real life.

Keywords: NIRS, neurofeedback, virtual reality, impulsivity, ADHD

INTRODUCTION

Impulsivity refers to the inability to inhibit behavioral responses to urges created by external stimuli as well as internal desires, often brought about by the current environment. It is a ubiquitous behavioral trait found in healthy individuals as well as those with developmental disorders such as attention-deficit/hyperactivity disorder (ADHD), substance-use disorders, binge eating disorders, and others (Whiteside and Lynam, 2001; Bari and Robbins, 2013). Individual impulsive episodes, such as drunk driving, can negatively impact the lives of the impulsive individual, as well as the lives of others. On neuropsychological tasks, impulsive behavior is associated with certain types of errors, typically on conditions requiring inhibitory control. For example, the more impulsive an individual is, the more commission errors [i.e., false alarms (FA)] they make on go/no-go tasks (Aichert et al., 2012; Weidacker et al., 2016). Impulsive subgroups such as binge eaters (Hege et al., 2014) and binge drinkers (Henges and Marczinski, 2012) also make more FA than healthy controls.

From a neuroscientific perspective, impulsivity is strongly linked with dysfunctional frontal lobe activity and frontal lobe excisions (Fallgatter and Herrmann, 2001; Bari and Robbins, 2013). Development of impulse control is the result of maturation of the cognitive control network (CCN; Casey et al., 2008; Steinberg, 2008; in Shulman et al., 2016) which consists of the lateral prefrontal cortex and its connectivity with other frontal, striatal, motoric, and parietal regions (for comprehensive reviews see Cubillo et al., 2012; Rubia et al., 2013). Highly impulsive subgroups require a stronger activation of the CCN than healthy controls to achieve comparable response inhibition (Horn et al., 2003; Ding et al., 2014). Additionally, evidence for negative correlations between trait impulsiveness and activation as well as connectivity in prefrontal brain structures has been provided (Farr et al., 2012). Furthermore, there is evidence that the bilateral dorsolateral prefrontal cortex (dlPFC) may be involved in inhibitory control as transcranial direct current stimulation (tDCS) of the left dlPFC led to improved inhibitory control on a go/no-go task in participants with ADHD (Soltaninejad et al., 2015).

Neurofeedback (NF), a therapeutic technique in which participants are tasked with regulating their own brain activity, is used as a way to effect long-term change in abnormal brain activity (Arns et al., 2013). Thereby, electroencephalography (EEG)-based NF protocols have shown promise in reducing impulsive symptoms in ADHD (Gevensleben et al., 2012, 2014a). However, these protocols have had mixed effects, particularly as they are often based on brain-frequency imbalances that are highly heterogeneous within subjects (Holtmann et al., 2014). A recently emerging NF protocol for ADHD using functional near-infrared spectroscopy (fNIRS) to measure the blood oxygenation level dependent (BOLD) response within the dlPFC has several potential advantages over traditional EEG protocols (Marx et al., 2015).

Compared to EEG, fNIRS has improved spatial resolution and better correspondence of channel to underlying brain region, as well as reduced sensitivity to movement-based artifacts, making it ideal for NF training of circumscribed brain areas in motorically restless individuals (e.g., ADHD patients, children, etc.). Furthermore, evidence from BOLD-based NF paradigms suggest that they yield effects faster than their EEG-based counterparts. In a pilot study with children with ADHD, significant symptom improvements were found after only 12 sessions of fNIRS-based dlPFC training (Marx et al., 2015). Sherwood et al. (2016) found that – in healthy subjects – achieving control of the BOLD response in the dlPFC is possible after just five sessions of real-time functional magnetic resonance imaging (fMRI) NF training. Current EEG protocols, on the other hand, require between 25 and 50 sessions to realize significant effects (for a review and meta-analysis see Begemann et al., 2016). However, despite the promise of BOLD-based protocols as a potential treatment for impulsivity, such protocols still need to translate from laboratory to real-world settings.

Neurofeedback treatment is often criticized for its lack of ecological validity. Simply put, strategies of brain regulation learned in a lab setting may not translate well into the real world. Those with impulsivity struggle in the classroom where academic achievement is negatively correlated to impulsivity severity (Spinella and Miley, 2003). Therefore, any effective strategies developed in NF therapy should ultimately be applied in the classroom (or a similar real-world) setting, a concept known as transfer (e.g., Strehl, 2014). However, NF protocols – at this point – cannot be utilized in a real scholastic setting as they require large and delicate equipment, and students need to concentrate on the current lesson. An increasingly viable option, virtual reality (VR), has been used for assessment of clinical symptoms of ADHD in the classroom (Muhlberger et al., 2016) and with an EEG-based NF protocol designed to reduce inattentive and impulsive behavior in adolescents displaying behavioral problems (Cho et al., 2004). In the latter study, the VR group showed the greatest improvement following NF training on attention-related tasks relative to both a control group and a 2-D classroom group, but no difference in impulsivity. However, this study was controlled with a waiting group, thus not ruling out non-specific effects of NF training, such as continuous performance monitoring, reinforcement of compliance, and the idea that one is being treated by a sophisticated technology and professional (Gevensleben et al., 2012, 2014b). Furthermore, the NF was a separate module, not incorporated into the experience of the class itself.

Based on these findings, we developed a virtual classroom-based fNIRS NF protocol (for study design see Blume et al., 2017) in order to directly facilitate transfer of NF training effects to the classroom. Importantly, feedback is delivered in the form of gentle dimming or brightening of the overhead lighting which does not distract the participant from the experience of being in a classroom. In the present study, we implement a 2 week accelerated protocol in highly impulsive young adults, consisting of eight training sessions (one per day) which were bookended by a pre- and a post-test to assess behavioral changes during a go/no-go, n-back, and stop-signal task (SST). Changes in frontal lobe function were also assessed during the go/no-go and n-back tasks using fNIRS. To control for the previously mentioned non-specific effects of the NF training, we used bilateral musculi

supraspinatus-based electromyography (EMG) biofeedback (BF) (see Marx et al., 2015; Mayer et al., 2015). This method has been successfully used in the aforementioned studies as a control for NIRS-based NF. Sham-based NF control groups (e.g., targeting putatively unrelated brain areas) invite ethical concerns, as training random areas may have unforeseeable negative effects on the participant, who is often recruited on the premise that the training will be helpful to their condition (Holtmann et al., 2014). Furthermore, participants sometimes become aware that they are part of some sham conditions (particularly if the sham feedback contains data completely unrelated to the current training situation, e.g., training data of another participant), or even assume they are part of one when they are not, leading to both drop-outs and reduced motivation, a critical aspect for any successful NF training (Birbaumer et al., 1991; Gevensleben et al., 2014a). As we did not explicitly inform participants that EMG BF was a control condition, they were less susceptible to this motivation loss.

We hypothesize that the fNIRS-based NF group will show an improvement in dlPFC activity during the cognitive tasks (go/no-go and n-back) relative to the EMG-based control group following the treatment program. We also expect the NF group to show a reduction in FA (go/no-go task) as well as reduced stop-signal reaction time (SSRT) on the SST from pre- to post-test measurement (as measures of response inhibition). As secondary outcomes, we expect reaction time (RT) and RT variability [standard deviation of the reaction time (SDRT)] to decrease for the NF group on all tasks, as the dlPFC plays a role in a multitude of executive functions. The expected neurocognitive improvements following frontal lobe focused fNIRS-based NF in a virtual training environment would confirm the general feasibility of a combination of NF with virtual training scenarios which could – in the long run – increase the ecological validity of NF interventions.

MATERIALS AND METHODS

Participants

We recruited 22 students from the University of Tübingen out of a larger group of potential participants who had completed the Barratt Impulsiveness Scale (BIS; Barratt, 1959) using an online format. Based on their high BIS scores ($M_{BIS} = 85.75$, $SD_{BIS} = 9.36$), these students were selected and invited to an in-person screening for ADHD [according to criteria from the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR; Sass et al., 2003)] using two subtests from the Homburger ADHD Scale for Adults (HASE; Rösler et al., 2008), the German versions of the Wender Utah Rating Scale (WURS-K) and the ADHD-Self Assessment Scale (ADHS-SB). Participants meeting the criteria for an indication of ADHD under this context (WURS-K $>$ 30 and ADHS-SB $>$ 18) were excluded from the study and informed about the outpatient ADHD program at the Department of Psychiatry and Psychotherapy at the University Hospital Tübingen ($n = 1$). The remaining participants ($n = 21$; nine female, $M_{Age} = 23.4$, $SD_{Age} = 2.8$) reported no history of serious or chronic illness, neurological, or psychiatric disorders.

This study was approved by the Ethics Committee of the Medical Faculty of the University and the University Hospital of Tübingen and all procedures were in accordance with the Helsinki Declaration of 1975, as revised in 2013. Participants provided written informed consent and were compensated with 100 Euros for completing the duration of the training including pre- and post-measurements (10 sessions, 1 h each, over 2 weeks). One participant dropped out of the study due to feeling ill from the VR and was payed pro-rata of 10 Euro per hour participated.

Study Design

The study followed a randomized, controlled experimental design. Participants were randomized (10 participants in each group) to either eight fNIRS-based NF (experimental) or eight EMG-based BF (control) sessions taking course daily over two weeks (Tuesday to Friday in the first week, Monday to Thursday in the second week). We randomized without stratifying for any other variables. Groups did not differ significantly in gender (NF: 4 female, 6 male; BF: 5 female, 5 male; Fisher's exact test, $p = 0.50$), or in age ($M_{BF} = 22.9$, $SD = 2.88$; $M_{NF} = 23.9$, $SD = 2.77$; $t(18) = 0.80$, $p = 0.44$). The pre-test and post-test were exactly the same and included a go/no-go task, an n-back task, and an SST. The pre-test took place on the Monday of the first week, while the post-test occurred on the Friday of the second week. Order of the pre- and post-test measures was counter-balanced between subjects.

Virtual Classroom Scenario

The participants were seated and wore the Oculus Rift (Oculus Rift, United States[1]) VR head-mounted display (HMD). The HMD rendered a virtual classroom developed by KatanaSim (KatanaSim, Germany[2]) with animated students and a teacher. The participants' point of view was seated first-person, facing the teacher (**Figure 1**). The participant had a full 360° view from the desk seat, with other students seated nearby. The task was to control the brightness of the lighting in the classroom. When an upward-pointing arrow was shown on the chalkboard, the participant was required to "activate" in order to make the light brighter. When the arrow pointed downward, the participant was required to "deactivate" in order to make the light darker. Briefly, activation requires higher output compared to baseline from the respective feedback source, while deactivation requires reduced output compared to baseline (see below for more details on fNIRS and EMG activation/deactivation protocols). Importantly, participants were not told, in either condition, how to regulate the lighting in the classroom, they were instructed simply to try to increase the lighting in the room when the arrow pointed upward and to decrease the lighting when the arrow pointed downward. In this way, only the positive or negative feedback they received from the scenario should have enforced their learning of the feedback parameter. The probability that a trial was activation (arrow up) was 50% in sessions 1–4 and 80% in sessions 5–8. More activation was encouraged in the second half of the scenario, as more upregulation of the prefrontal cortex is

[1] www.oculusvr.com

[2] www.katanasim.com

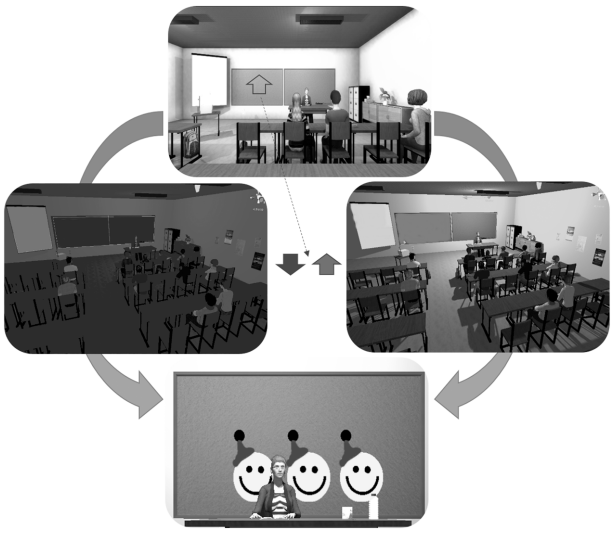


FIGURE 1 | The virtual reality classroom scenario. The top image depicts the view from the participant's head-mounted display (HMD). An arrow pointing either up or down was displayed on the chalkboard. If the arrow pointed downward, the participant should decrease the lighting in the classroom. If it pointed up, the participant should increase the lighting in the classroom. If the participant performed the task adequately, they would receive one to three smileys, presented on the chalkboard, based on the duration of success.

associated with stronger inhibitory control (Rubia et al., 2013; Soltaninejad et al., 2015). Participants were confronted with distractions within the scenario (e.g., students turning around or cell phones ringing) from the second half of each session until the end.

Before each trial, a baseline and threshold of light fluctuation were calculated to determine the point at which the classroom light was balanced between fully bright and fully dark and the range within which it could fluctuate. Following successful activation or deactivation – when the signal was 60, 70, or 80% of the time above or below the baseline, respectively – the participant was rewarded with one, two, or three smiley faces, respectively, on the chalkboard.

Each session was comprised of three blocks, the first and the last being 12 min in length while the second, the transfer block, was 8 min. In the transfer block, the light's brightness was fixed, meaning that the only feedback came at the end of each trial. Trial number and length varied depending on the feedback source and will be discussed in the following sections.

fNIRS

Functional near-infrared spectroscopy records change in oxygenated (O_2Hb) and deoxygenated (HHb) hemoglobin relative to a baseline; the amount of local O_2Hb infers the amount of local brain activation, via the process of hemodyamic coupling, wherein increases of cortical activation lead to increases in O_2Hb and decreases in HHb (Haeussinger et al., 2014). The ETG-4000 continuous Optical Topography System (Hitachi Medical Co., Japan) was used for pre- and post-tests as well as NF sessions.

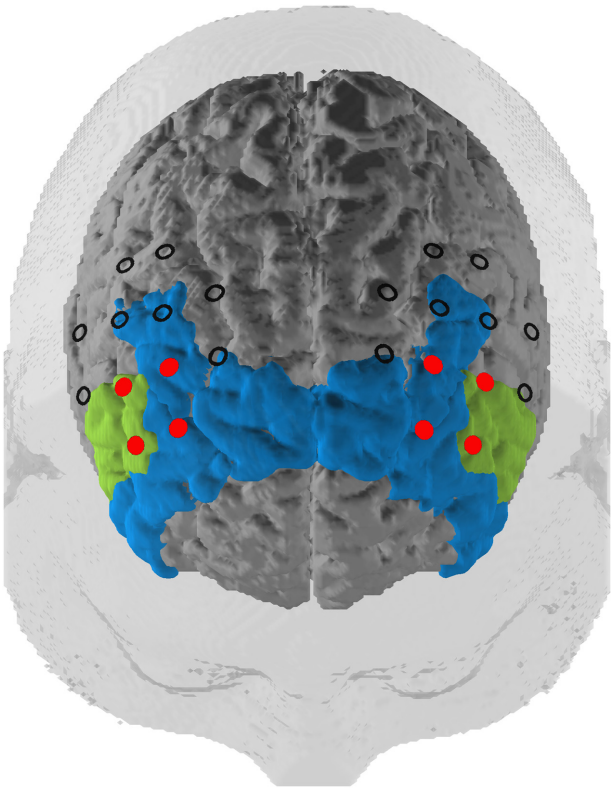


FIGURE 2 | Probeset. Depiction of the target regions of the neurofeedback training (in red) and the rest of the optode array. The optodes cover a space mostly located in the dlPFC (blue, Brodmann areas 9 and 46) but also extending to the inferior frontal gyrus (green, Brodmann area 45) according to a virtual registration method (see subsection Functional Near-Infrared Spectroscopy Data).

Our optode montage featured two 3 × 3 optode arrays centered with the innermost channel of the front row of each array placed on F3 (left hemisphere) and F4 (right hemisphere) of the international 10–20 EEG system (Jasper, 1958). Source–detector distances were kept at 3 cm. The optode arrays were rotated 45° laterally along the transversal plane so that the innermost four channels in the two frontal rows were oriented over the left and right dlPFC (**Figure 2**). The third optode array was a 3 × 5 arrangement where the most superior and lateral optode on the left and right of the array were oriented on P3 and P4, respectively. Subtending the parieto-occipital cortex, this probeset was used exclusively for common average (CA) reference, a signal correction method (see below).

fNIRS Feedback Signal and Trials

The feedback target was the average amplitude of O_2Hb within the bilateral dlPFC (see Marx et al., 2015). The raw fNIRS signal was sampled at 10 Hz and preprocessed in MATLAB version 9.0 (The MathWorks Inc., United States). A moving average Kalman filter with a 5 s sliding window was then applied to the data. Finally, we used a CA artifact removal method used in previous NF designs serving as a

basis for this design (Marx et al., 2015; Mayer et al., 2015). This method was preferred because of its ability to remove probeset-wide effects from individual channels (Heinzel et al., 2013). For the CA, the raw average of all 46 channels was subtracted from the raw average of the eight emitter–detector channel pairings over the dlPFC in order to limit the influence of artifacts – e.g., superficial blood flow, head and jaw movements, and respiration – on the hemodynamic response in the feedback channels. All preprocessing occurred online.

The fNIRS trials were 30 s in duration with a 5 s baseline period. Relative O_2Hb concentration higher than baseline led to brightening of the lights; concentration lower than baseline led to dimming. Trials were divided into three blocks (**Figure 3**). The first and last blocks contained 12 trials and subsequent rests of 20 s duration. The middle block contained eight trials and rests and was used as the transfer block, wherein no continuous feedback was provided, though participants were still given feedback at the end of the trial. There was no jittering of intertrial intervals.

EMG Feedback Signal and Trials

Monopolar EMG, with a sampling rate of 1000 Hz, provided feedback from the bilateral *musculus supraspinatus* for the control group (see Mayer et al., 2015). The signal was referenced to the right mastoid and was grounded on the left mastoid. The data stream was bandpass filtered between 80 and 300 Hz. The resulting signal was then normalized via a maximum output and a resting output, for which the participant flexed both muscles maximally for 10 s and sat completely at rest for 10 s, respectively. At each time point, feedback was equivalent to:

$$\text{Feedback Index} = \text{R}-\text{L},$$

where R and L were the right and left normalized muscle outputs, respectively, given by:

$$\text{R(L)} = \frac{\text{Right (Left) EMG Signal} - \text{Average Resting Baseline Right (Left)}}{\text{Average Maximal Muscle Output Right (Left)}}.$$

Therefore, more tensing of right muscle led to brightening; more tensing of left muscle led to dimming. Baseline for each trial was an average of the last 2 s of the resting feedback signal.

The EMG trials were 15 s in duration with a 2 s baseline period. Relative muscular feedback index higher than baseline led to brightening of the lights; feedback index lower than baseline led to dimming. Trials were divided into three blocks (**Figure 3**). The first and last blocks contained 24 trials and subsequent rests of 10 s duration. The middle block contained 16 trials and rests and was used as the transfer block, wherein no contingent feedback was provided.

Pre- and Post-measures

Go/no-go and n-Back Task

The go/no-go and the n-back tasks were programmed in Presentation version 18.0 (Neuro Behavioral Systems, United States) following previously published protocols (Mayer et al., 2015; see also Ehlis et al., 2008). We recorded fNIRS during both tasks. Briefly, the go/no-go task consisted of alternating go and no-go blocks (four repetitions each) separated by rest blocks, each block lasting 30 s. In the "go" condition, participants were asked to respond as fast as possible to each stimulus. In the "no-go" condition, participants were instructed to withhold their response on no-go trials (here: presentation of the letter "N"; 25% of trials). Dependent variables were RT, SDRT, FA, and omission errors.

The n-back task consisted of three blocks each of 2-back (high working memory load), 1-back (low working memory load), and 0-back (control) (block length: 30 s; separated by 30 s rest periods). In the 2- (1-)back task, the participants were instructed to press the space bar as quickly as possible whenever the current letter was the same as the letter two letters (one letter) back. In the 0-back task, the participant was instructed to respond when the letter "O" appeared on the screen. Dependent variables were RT, SDRT, and correct hits.

Stop-Signal Task

The SST followed the protocol described in Verbruggen et al. (2008). The task consisted of one practice block and three 3-min verum blocks wherein the participant should respond to the direction of an arrow pointing on the screen as quickly as possible. In roughly 25% of trials, the arrow would turn blue,

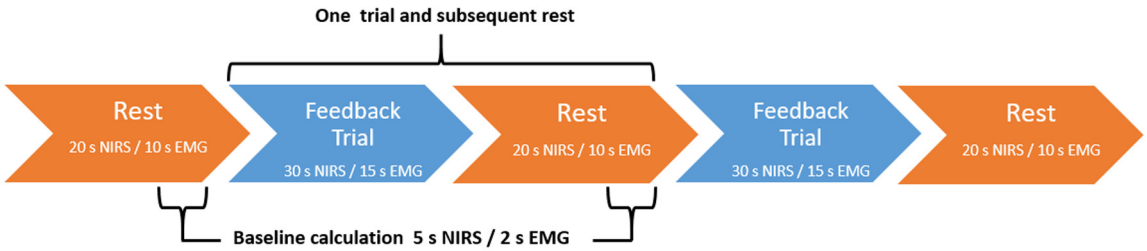


FIGURE 3 | Feedback block design. fNIRS-based NF blocks consisted of either 12 NF trials and subsequent rest trials (continuous feedback blocks one and three) or 8 NF trials and subsequent rests (transfer block two). EMG-based BF blocks consisted of either 24 BF trials and subsequent rests (continuous feedback blocks one and three) or 16 BF trials and subsequent rests (transfer block two). In both conditions, blocks always began with a rest trial.

indicating the participant should withhold their response, after a variable stop-signal delay (SSD) that started at 250 ms and increased or decreased by 50 ms depending on if they failed or succeeded to stop, respectively. Dependent variables in the SST included the SSRT – a measure of behavioral inhibition – RT, and SDRT. The SST was added as a secondary measure for behavior. We did not record simultaneous fNIRS with this measure.

Analysis

Functional Near-Infrared Spectroscopy Data

All analysis was performed using MATLAB. In order to analyze fNIRS data, we used subroutines programmed in our research group, adapted for fNIRS from the Statistics Parametrical Mapping toolbox for MATLAB (SPM8; Friston et al., 1994). Raw signals were bandpass filtered between 0.01 and 0.2 Hz to remove unwanted physiological artifacts such as heartbeat and respiration. Next, channels exceeding three times the within-subject standard deviation over the course of the measurement were interpolated (see Hagen et al., 2014) using a Gaussian distribution with the O_2Hb values of proximal channels given a higher weighting than distal ones; less than 10% of all channels were interpolated. We then applied a wavelet-based transform (Molavi and Dumont, 2012) to detect and correct motion artifacts that were still part of the data. We used the hmrMotionCorrectWavelet algorithm from the Homer2 fNIRS analysis package for MATLAB with the standard motion artifact detection threshold of 1.5 *SD* above the interquartile range of the data (Huppert et al., 2009). Finally, a block-related average amplitude was calculated for each channel using an interval of 0–60 s after block onset with a 10-s baseline correction. Linear detrending was applied to remove slow drifts in the data. Finally, average amplitudes over the duration of the task blocks (0–30 s) were calculated.

Region of Interest (ROI)

We mapped fNIRS channels to corresponding, underlying cortical areas based on a virtual registration method (Rorden and Brett, 2000; Singh et al., 2005; Tsuzuki et al., 2007). The left and right dlPFC regions of interest (ROIs) consisted of the channels that we used for the NF training. These channels are concentrated in Brodmann Areas 9, 45, and 46. This includes the dlPFC and also slightly expands into the inferior frontal gyrus (IGF; see **Figure 2**).

Rate of Learning and Correlation with Primary Outcome Variables

Additionally, we analyzed the success of the participants in obtaining control of the feedback parameter. Our success rate was calculated as the average percentage of time spent in the correct direction of the desired feedback (above or below the baseline, for activation vs. deactivation trials, respectively) for the duration of the trial. An average was calculated for all trials from the first week (four sessions) and the second week (four sessions). The rate of learning was calculated as the average of the second week minus the average of the first week. Rate of learning was then correlated with the primary outcome variables of FA rate in the no-go task and average amplitude of O_2Hb of the feedback channels during the no-go task. Similar metrics were created in order to compute the correlations: pre–post FA errors were computed for each subject, to give a metric of individual improvement. Similarly, a post–pre average amplitude of O_2Hb of the feedback channels was computed to reflect difference in activation after the training. In the event of significant correlations in one or more groups, we computed a pseudo-permutation test ($n = 10{,}000$ permutations), permuting the group assignment while keeping within-subject correlation pairs intact, to determine a significant difference between groups. The number of permutations in which the permuted group difference in ρ value was larger than the verum group difference in ρ value was divided by the total number of permutations to create a *p*-value.

Statistical Analysis

To evaluate the statistical significance of pre–post changes in O_2Hb and HHb in the go/no-go and n-back tasks, we conducted $2 \times 2 \times 2 \times 2(3)$ repeated measures analyses of variance (ANOVAs), with the between-subjects factor treatment group (NIRS vs. EMG) and the within-subject factors of time (pre vs. post), ROI (left dlPFC vs. right dlPFC), and condition (n-back (3): 2-, 1-, and 0-back; go/no-go (2): go and no-go). For behavioral data, repeated measures ANOVAs were performed using the same factors excluding ROI. When data violated the assumption of sphericity, Greenhouse-Geisser corrected values were reported. For significant main and interaction effects, two-tailed Student's *t*-tests were employed for *post hoc* analyses (paired or independent samples, as appropriate). In cases where the assumption of normality was violated, we used two-tailed Mann–Whitney *U* tests or Wilcoxon signed-rank tests, respectively.

ROI Specificity

In order to determine specificity of ROIs we used pseudo-permutations tests, wherein the mean difference in the average amplitudes from pre to post measurement for a given verum ROI (vROI) for all participants was compared to a pseudo-ROI (pROI) composed of an equal number of randomly chosen NIRS channels. $N = 10{,}000$ permutations of pROI were calculated and the resulting *p*-value was the sum of trials in which the resulting statistic from the vROI was greater than the permuted statistic from the pROI.

RESULTS

Behavioral Data

Only significant results related to the hypotheses are reported here. For a full summary of behavioral data, see **Table 1**.

Go/no-go

False alarm errors in the go/no-go task showed a trend with a large effect size for a measurement time*group interaction effect

TABLE 1 | Behavioral data from pre- and post-test in the two experimental groups.

	Pre-test		Post-test	
Task	**NIRS group means (±*SD*)**	**EMG group means (±*SD*)**	**NIRS group means (±*SD*)**	**EMG group means (±*SD*)**
Go/no-go				
Go RT (ms)	300.0 (20.3)	289.6 (47.1)	290.5 (18.5)	293.5 (43.3)
Go SDRT (ms)	90.9 (43.1)	90.4 (39.4)	77.7 (18.1)	144.5 (99.6)
Go omission errors	0.2 (0.4)	2.9 (5.7)	0.8 (0.9)	1.6 (2.5)
No-go FA errors	4.8 (2.4)	4.8 (2.7)	2.6 (1.3)	6.0 (5.2)
No-go RT (ms)	434.6 (30.4)	417.4 (30.4)	438.3 (48.3)	411.3 (38.0)
No-go SDRT (ms)	82.0 (37.6)	81.0 (38.8)	71.1 (23.6)	79.6 (30.3)
N-back				
2-back Hit rate	0.95 (0.06)	0.93 (0.11)	0.98 (0.04)	0.93 (0.11)
1-back Hit rate	1 (0)	0.99 (0.03)	0.98 (0.08)	0.97 (0.04)
0-back Hit rate	1 (0)	0.99 (0.03)	1 (0)	0.96 (0.06)
2-back RT (ms)	554.2 (65.8)	485.8 (77.8)	550.3 (74.0)	433.5 (66.4)
1-back RT (ms)	473.6 (61.0)	411.7 (46.8)	491.2 (78.8)	427.2 (76.2)
0-back RT (ms)	423.9 (48.3)	388.7 (35.1)	450.0 (49.2)	418.8 (83.6)
2-back SDRT (ms)	173.1 (57.5)	163.8 (50.2)	171.7 (69.5)	87.9 (25.9)
1-back SDRT (ms)	122.6 (58.1)	71.8 (19.1)	123.5 (81.5)	82.7 (36.9)
0-back SDRT (ms)	96.6 (18.3)	139.7 (50.1)	139.7 (50.4)	113.3 (55.1)
Stop-signal task				
SSRT (ms)	223.5 (36.8)	224.2 (53.9)	232.2 (55.8)	223.6 (55.7)
Go trial RT (ms)	659.2 (212.8)	543.4 (96.9)	605.1 (186.5)	568.9 (102.5)
Go trial SDRT (ms)	160.8 (64.9)	125.6 (46.1)	124.1 (60.6)	145.7 (61.0)

SD, standard deviation; RT, reaction time; SDRT, standard deviation of the reaction time; FA, false alarms; SSRT, stop-signal reaction time.

($F(1,18) = 4.08$, $p = 0.059$, $\eta^2 = 0.185$). *Post hoc* Wilcoxon signed-rank tests revealed a reduction of FA errors from pre to post measurement in the experimental group ($M_{pre} = 4.8$, $SD_{pre} = 2.4$; $M_{post} = 2.6$, $SD_{post} = 1.3$; $Z = -2.57$, $p = 0.01$), but not in the control group ($M_{pre} = 4.8$, $SD_{pre} = 2.7$; $M_{post} = 6.0$, $SD_{post} = 5.2$; $Z = -0.30$, $p = 0.77$) (**Figure 4A**). No other interaction effects were observed.

Rate of Learning

A one-sample Kolmogorov–Smirnov test rejected the null hypothesis that the learning rates for the first half and second half of the experimental and control groups, respectively, followed a normal distribution ($D = 0.65, 0.65, 0.64, 0.65$, $N = 10$ each, and $p < 0.05$ each). Therefore, Wilcoxon signed-rank tests and Spearman correlations were calculated. For the experimental group, there was no significant difference between first half and second half performance, but a medium effect size indicating better second half performance ($Z = 1.48$, $p = 0.13$, $r = 0.33$). There was, however, a significant difference between first and second half performance for the control group ($Z = 2.68$, $p = 0.013$, $r = 0.60$), indicating a significantly better performance in the second week with a large effect size.

The rate of learning of both groups failed to correlate significantly with post–pre changes in average O_2Hb concentration in feedback channels ($|\rho| < 0.224$, $p > 0.05$). The rate of learning in the experimental group, however, correlated strongly with size of pre–post reduction in FA ($\rho = 0.75$, $p = 0.013$; see **Figure 4B**). Rate of learning in the control group did not correlate with pre–post reduction in FA ($\rho = -0.24$, $p = 0.508$). The resulting pseudo-permutation test concluded that there was a significant group difference ($p = 0.015$).

N-Back Task

No significant behavioral interaction effects were observed. Hit rates for each condition were nearly 100% in the pre-test. Furthermore, no FA errors were made in this task. A ceiling effect was evident for this task.

Stop-Signal Task

Reaction time variability yielded a significant interaction effect of measurement time*group ($F(1,18) = 5.39$, $p = 0.03$, $\eta^2 = 0.231$), with the experimental group showing significantly reduced RT variability following the training ($M_{pre} = 160.78$ ms, $SD_{pre} = 64.88$; $M_{post} = 124.13$, $SD_{post} = 60.60$; $t(9) = 2.48$, $p = 0.035$). The control group showed no difference between measurements ($M_{pre} = 125.55$ ms, $SD_{pre} = 46.13$; $M_{post} = 145.70$, $SD_{post} = 61.04$; $t(9) = 1.04$, $p = 0.328$).

fNIRS Data

Go/no-go O_2Hb

We observed a main effect of task ($F(1,18) = 11.92$, $p = 0.003$, $\eta^2 = 0.398$, mean amplitudes: $M_{go} = 0.005$, $SD = 0.033$ mm*mol/l, $M_{no\text{-}go} = -0.005$, $SD = 0.029$ mm*mol/l) and an interaction effect of time*task*ROI*group ($F(1,18) = 5.63$, $p = 0.029$, $\eta^2 = 0.238$). This interaction was caused by a pre to post increase in O_2Hb amplitudes of the left dlPFC in the experimental group during the no-go task ($M_{pre} = -0.029$, $SD = 0.035$ mm*mol/l; $M_{post} = 0.010$,

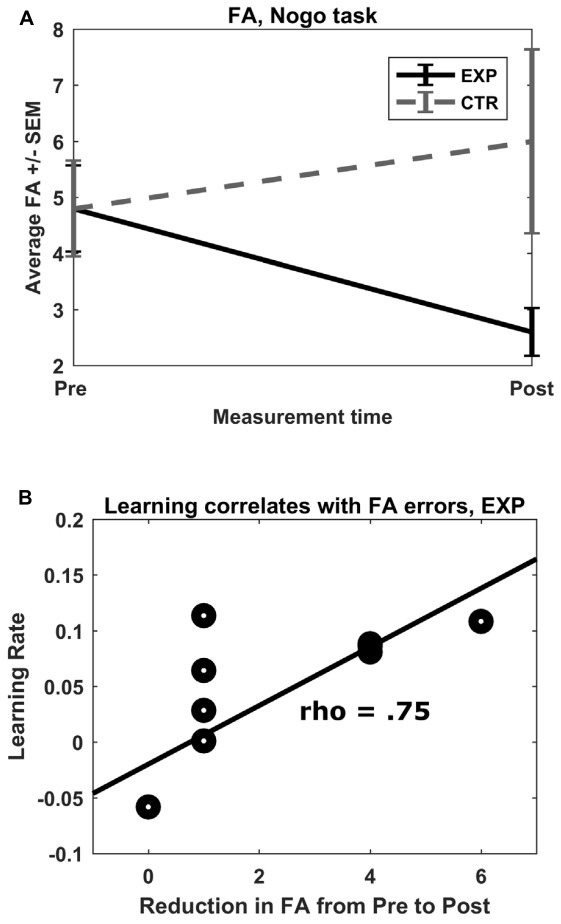


FIGURE 4 | (A) No-go FA errors. Mean total FA errors for the no-go condition shown for both groups for both pre- and post-test. *p = 0.01. **(B)** Rate of learning correlation with FA reduction. In the experimental group, there was a strong correlation between rate of learning of the feedback parameter (prefrontal oxygenation) and pre–post reduction in FA errors committed.

$SD = 0.040$ mm*mol/l; $t(9) = -3.63$, $p = 0.005$; see **Figure 5**). In the control group of the same condition, time, and ROI, there was no significant change ($M_{pre} = 0.006$, $SD = 0.017$ mm*mol/l; $M_{post} = -0.006$, $SD = 0.031$ mm*mol/l; $t(9) = 1.15$, $p = 0.281$). All other *post hoc* comparisons failed to reach significance ($|t(9)| < 1.837$, $p > 0.1$). The permutation test indicated that this ROI was indeed the focal point for the increase in brain activation. The resulting p-value was equal to $p = 0.003$, indicating that there is high spatial specificity to the activation, located in the left dlPFC.

Go/no-go HHb

We observed no main effects, only an interaction effect of task*hemisphere ($F(1,18) = 5.79$, $p = 0.027$, $\eta^2 = 0.243$). *Post hoc* testing indicated that there was a trend toward a significant difference in HHb activation between the left and right hemisphere in the "go" condition ($M_{left} = -0.005$, $SD = 0.023$; $M_{right} = -0.013$, $SD = 0.026$; $t(9) = 2.07$, $p = 0.052$).

N-Back O_2Hb

We observed no main effects or significant interaction effects (all $|F(2,36)| < 2.50$; all $p > 0.11$).

N-Back HHb

We observed a trend for a main effect of task ($F(1.39,24.93) = 3.75$, $p = 0.052$, $\eta^2 = 0.173$; mean amplitudes: $M_{2Back} = -0.011$, $SD = 0.023$; $M_{1Back} = -0.007$, $SD = 0.021$; $M_{0Back} = 0.001$, $SD = 0.024$). Again, the indication is a higher activation in tasks with a higher working load. We also observed a trend for a main effect of time ($F(1,18) = 3.26$, $p = 0.088$, $\eta^2 = 0.153$; $M_{Pre} = -0.008$, $SD = 0.026$; $M_{Post} = -0.003$, $SD = 0.021$), indicating a marginal decrease in activation across all tasks from pre to post measurement time. No other main effects or interaction effects were observed.

DISCUSSION

The present study was designed to test the efficacy of a novel neurofeedback intervention (fNIRS-based frontal lobe NF in a virtual classroom environment) with the ultimate aim of reducing ADHD symptoms in schoolchildren by increasing their ability to regulate prefrontal cortex activity (Blume et al., 2017). Here, we focused on the effects of this newly developed NF protocol in a sample of highly impulsive young adults, a sub-clinical risk population that exhibits many of the behavioral abnormalities also seen in patients with ADHD (e.g., Herrmann et al., 2009). In this proof-of-concept study, we were primarily interested in first, whether the fNIRS-based NF group would show increased cortical activation in feedback channels during frontal lobe/impulsivity-related tasks (go/no-go and n-back), following focused training of these channels and second, whether the fNIRS-based NF group would show a reduction in impulsive behaviors (go/no-go, n-back, SST).

During a go/no-go task, we observed a significant increase compared to a pre-training baseline in cortical O_2Hb concentration in the left dlPFC of the experimental (fNIRS) group only. During the same task, we observed a concurrent and significant reduction in FA errors of the same group. Importantly, this reduction in FA errors correlated significantly with the rate of learning of the experimental subjects but not the control subjects. Additionally, we observed a reduction in RT variability on the SST for the experimental group. We observed no group differences in either cortical activation or behavior on the n-back task. The lack of a group difference after training on this task is likely due to the study specifically focusing on the recruitment of highly impulsive students. There is no evidence to suggest that highly impulsive participants have explicit deficits in working memory. In fact, in a study examining the correlations between trait impulsivity (as measured by BIS self-report) and performance on various neurocognitive tasks, no significant correlation was found between trait impulsivity and working memory performance, while trait impulsivity correlated strongest with go/no-go errors (Keilp et al., 2005). Furthermore, task accuracy reflected a ceiling effect from the pre-test, indicating that the task was not difficult for these

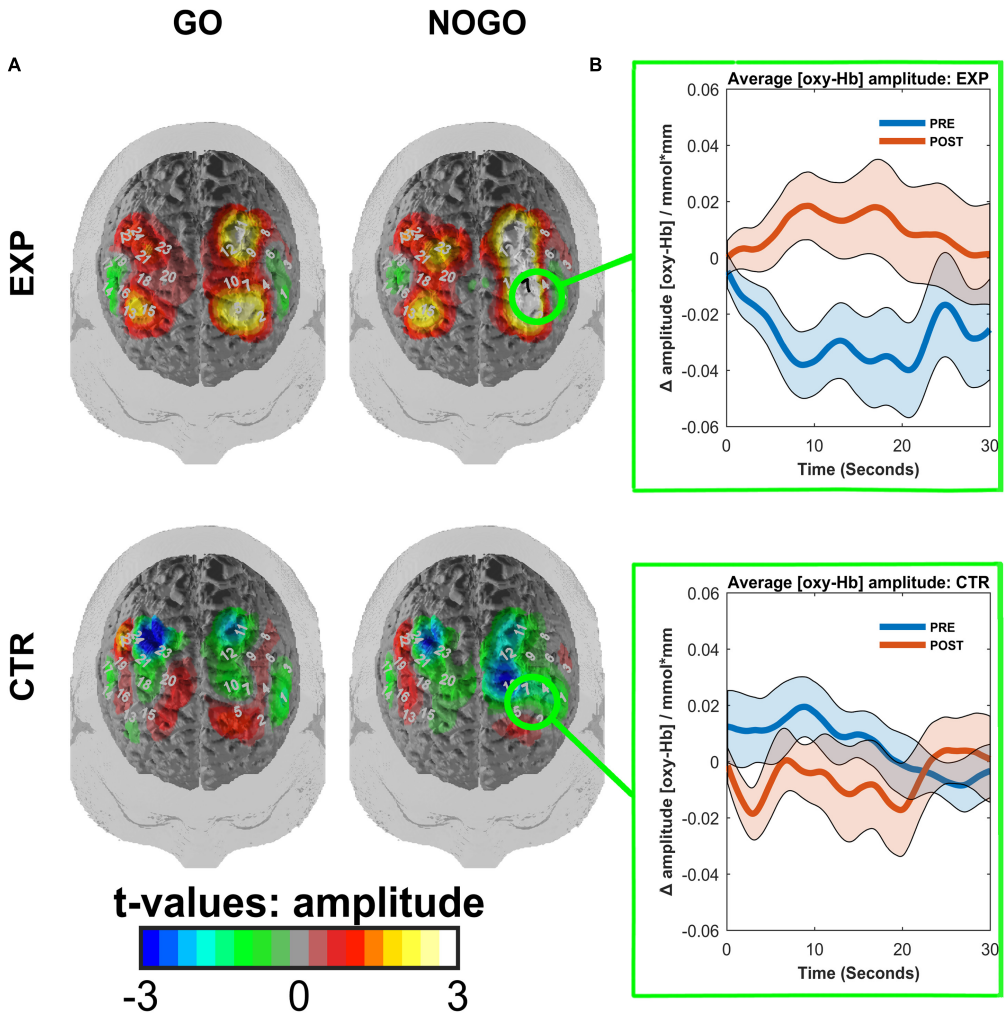


FIGURE 5 | (A) Contrasted t-maps of the average amplitudes of O_2Hb for different blocks of the go/no-go task. Contrasts represent post-test values minus pre-test values. CTR represents the EMG control group, while EXP represents the experimental fNIRS group. *T*-values were obtained by *t*-tests corrected for multiple comparisons with the Armitage–Parmar correction. Positive channels indicate stronger activation in the post-relative to the pre-test. Significant channels are depicted in black. **(B)** ROI event-related averages. Circled regions from **(A)** indicate the left dlPFC ROI for which the event-related average of O_2Hb ± standard error of the mean (SEM) is depicted for both pre-(blue) and post-(red) tests.

subjects. Therefore, despite the potential benefit to working memory that training the dlPFC might imbue, in our case there may have been no deficit to correct. Lastly, HHb data showed no differences in activation in either task. These results make sense in the context of the NF training; since O_2Hb was trained, the hypothesis would be that O_2Hb and not HHb would show the strongest pre–post effects. In addition, O_2Hb is more sensitive to detection of changes than HHb (Strangman et al., 2002).

False alarm errors, or incorrect go-responses to no-go stimuli, represent a failure to exhibit response inhibition (Aichert et al., 2012), an impulsive trait that subjects with ADHD share with highly impulsive participants. A reduction for the experimental group and not for the control group suggests that the fNIRS intervention was effective in reducing impulsive behavior as specified. The strong O_2Hb correlation observed between a reduction in FA errors and the rate of learning within the experimental group, but not within the control group, further illustrates the importance of specificity in NF training. The goal of actually learning to control the feedback parameter is often overlooked in NF studies, where the rate of obtained control is rarely reported (Zuberer et al., 2015). Interestingly, the control group showed a significant improvement between the first and second week in regulating the feedback parameter while the experimental group did not. This likely has to do with the comparable ease of the EMG feedback; once one learns the correct movement, it can relatively easily be replicated every trial. The fNIRS feedback is likely more complex, as there is no right or wrong way to achieve the feedback parameter, and sustaining oxygenation of the dlPFC over time is strenuous. Given this complexity, the medium effect size observed in the fNIRS learning rate is encouraging, and may simply mean that more sessions are needed to fully gain control. Moreover, for

the specific sample investigated and trained here (i.e., highly impulsive subjects), frontal lobe alterations have been shown as a central neurophysiological correlate, so it is perhaps not surprising that improving control over this area of the brain seems to have been particularly difficult. However, this behavioral effort seems to pay dividends, as we see that the more control impulsive subjects were able to gain over the activation of their dlPFC, the fewer FA errors they made, whereas the successful learning of the EMG parameter had little effect. This result supports the findings of an fNIRS study that sought to differentiate the roles of the medial and lateral prefrontal cortex during a go/no-go task. The bilateral middle frontal gyrus (i.e., the dlPFC) was responsible for error monitoring during the motor inhibition segment of the go/no-go task (Rodrigo et al., 2014). Our results indicate that the combination of both correct feedback parameter (i.e., frontal lobe focused) and successful learning of that parameter, not one or the other in isolation, is important to the feedback's overall success.

The task-specific increase in prefrontal oxygenation coinciding with a reduction in FA errors suggests that – following the frontal fNIRS training – the highly impulsive participants were able to recruit more cognitive resources, particularly from the dlPFC, during this task, leading to improved performance. Whether or not this was intentional is a matter of debate, but the goal of NF interventions remains to train implicit activation of brain activity through operant and classical conditioning (Strehl, 2014). Therefore, it seems that the participants were able to transfer skills learned either implicitly or explicitly from the training into a performance situation. Furthermore, this increase in cortical activation was both task- (no-go) and region-specific (left dlPFC). While there was no increase in activation in the right dlPFC, the left-specific increase as well as the increase in inhibitory control are in line with the tDCS study of Soltaninejad et al. (2015) who used cathodal stimulation over the left dlPFC of adolescents with ADHD and observed a decrease in FA errors. While the literary consensus places the locus of inhibitory control within the right dlPFC, inferior prefrontal, premotoric, and striatal brain structures (Aron et al., 2004, 2014; Bari and Robbins, 2013; Obeso et al., 2013), the left dlPFC shares strong functional connectivity with the above-mentioned areas (Ridderinkhof et al., 2004; Aron et al., 2014). Moreover, the dlPFC does not seem to be directly responsible for inhibitory control, but rather functions as a higher order mechanism that organizes the relevant brain structures above when attention control or increased working memory capacity is needed, in particular for oddball or complex no-go tasks (Criaud and Boulinguez, 2013). Because our go/no-go paradigm could be considered oddball, with an occurrence of no-go stimuli in only 25% of trials, it may be that the extra dlPFC resources recruited were used for focusing attention, rather than inhibitory control *per se*. Indeed, the reduction in SDRT seen in the SST also indicates an increase in attentional resources, possibly also mediated by an increase in prefrontal brain activity, though NIRS data were not available for this task. Increases in SDRT are generally considered to be related to lapses in attention (Alderson et al., 2007), though Kirkeby and Robinson (2005) found SDRT to be inversely correlated with trait impulsivity. Still, this does not rule out the idea that our impulsive sample also suffered from inattentiveness.

Treatment effects for both impulsivity and possibly inattention are encouraging from a translational perspective regarding potential use of our NF design with an ADHD population. We chose the dlPFC as a NF site because of its involvement in general top-down cognitive control, and the realization of significant training effects in impulsivity and possible inattention suggests that the protocol may be useful for an ADHD population. Several reasons lead us to be hopeful of even greater effects in a current study in our lab with ADHD schoolchildren (Blume et al., 2017). First, the sample size of this study was small. Only large effects could be detected, and with a greater sample size, we would expect to see effects in a wide range of other cognitive and behavioral deficits. Secondly, the training was compact and about half the number of training sessions we would recommend (and currently use) for a clinical ADHD project. As far as we know, this is the shortest number of training sessions to produce effects in brain activation and behavior that was adequately controlled for specificity. Cho et al. (2004) also used a 2 week, eight session NF paradigm with EEG and found training effects for inattention and impulsivity, but they did not have an adequate control group (waiting group), and additionally, did not measure differences in brain activity pre and post. Lastly, but most importantly, children have a greater capacity for brain plasticity than adults (Kolb and Gibb, 2011). For children with ADHD, this capacity is even more pronounced within the dlPFC, a region that develops particularly late for them (Rubia et al., 2013). Given the current study's results, we would expect even greater improvements within a child population.

The current study was limited by several factors, which we hope to improve upon in a second study with children with an ADHD diagnosis (Blume et al., 2017). The sample size was small which limited data analysis. Our aim was to test the viability of an immersive VR NF paradigm, and it appears that the full classroom immersion did not detract from the ability of the participants to regulate their brain activity. There was a difference between experimental groups in pre-test no-go activation, with the experimental group showing less activation than the control group. Small groups, even with proper randomization, have a much greater chance of having differing baseline measurements simply due to sampling error (Marshall, 1996). The larger the group, the smaller the chance of pre-baseline differences due to a random sampling error. As NF studies require large time and monetary investments per participant, and the aim of our study was to ultimately test the efficacy of VR NF, we chose 10 participants per group as a balance between power and realism. For technical reasons, we did not have triggers to compare the extent to which participants were able to regulate their brain activity across sessions, something that will be improved in the next study. While we used distractors in the current study, there was no way to compare trials in which a distractor occurred to trials in which they did not. Furthermore, we lack a comparison of the effects of the immersive VR NF paradigm to a 2-D version. In an ongoing study with children with ADHD (Blume et al., 2017),

we include a 2-D group that still uses lighting in the classroom as the feedback source, but the child sees the classroom on a normal computer monitor. In this way, we will be able to determine if immersive NF is actually more effective for the transfer of the learned regulation. Furthermore, the classroom itself is only one of many possible VR NF designs. Virtual reality scenarios coupled with NF are limited only to the imagination and relevance to a certain psychological disorder. Virtual reality NF with subjects with social phobias, for example, could be integrated within a potentially stressful social situation, like a bar or dinner party, furthering the ecological validity of the treatment while also avoiding an exposition-driven therapeutic approach that cannot be as easily controlled.

Considering these limitations and the relative ease with which they could be improved upon going forward, it seems that VR NF is a very promising modality for the treatment of behavioral disorders with known pathophysiological alterations.

ETHICS STATEMENT

This study was approved by the Ethics Committee of the Medical Faculty of the University and the University Hospital of Tübingen and all procedures were in accordance with the Helsinki Declaration of 1975, as revised in 2013.

AUTHOR CONTRIBUTIONS

All authors have approved of the final version of this manuscript. JH study design, data collection and analysis, and manuscript preparation; FB study design, data collection, and manuscript preparation; TD study design and manuscript preparation; FH data analysis; TR study design and manuscript preparation; AF study design and manuscript preparation; CG study design and manuscript preparation; and A-CE study design, data analysis, and manuscript preparation.

FUNDING

This research was funded by the LEAD Graduate School & Research Network (GSC1028), a project of the Excellence Initiative of the German federal and state governments. JH and FB are doctoral candidates and TD is a Junior Research Group Leader at the LEAD Graduate School & Research Network. A-CE was partly supported by the Interdisciplinary Center for Clinical Research (IZKF) Tübingen (Junior Research Group, grant 2115-0-0). We acknowledge support by the Deutsche Forschungsgemeinschaft and the Open Access Publishing Fund of the University of Tübingen.

ACKNOWLEDGMENTS

The authors would like to give additional thanks to those who contributed with measurements, in particular Marc Werth, Silke M. Bieck, and Nana Ambs. Sandru Razvan was influential in editing and structuring of the manuscript. Additionally, uncompromising support was given by Betti Schopp and Ramona Taeglich in coordinating measurements and training. They would also like to thank the participants for their participation and a very involved study.

REFERENCES

Aichert, D. S., Wöstmann, N. M., Costa, A., Macare, C., Wenig, J. R., Möller, H., et al. (2012). Associations between trait impulsivity and prepotent response inhibition. *J. Clin. Exp. Neuropsychol.* 34, 1016–1032. doi: 10.1080/13803395.2012.706261

Alderson, R. M., Rapport, M. D., and Kofler, M. J. (2007). Attention-deficit/hyperactivity disorder and behavioral inhibition: a meta-analytic review of the stop-signal paradigm. *J. Abnorm. Child Psychol.* 35, 745–758. doi: 10.1007/s10802-007-9131-6

Arns, M., Heinrich, H., and Strehl, U. (2013). Evaluation of neurofeedback in ADHD: the long and winding road. *Biol. Psychol.* 95, 108–115. doi: 10.1016/j.biopsycho.2013.11.013

Aron, A. R., Robbins, T. W., and Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci.* 8, 170–177. doi: 10.1016/j.tics.2004.02.010

Aron, A. R., Robbins, T. W., and Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: one decade on. *Trends Cogn. Sci.* 18, 177–185. doi: 10.1016/j.tics.2013.12.003

Bari, A., and Robbins, T. W. (2013). Inhibition and impulsivity: behavioral and neural basis of response control. *Prog. Neurobiol.* 108, 44–79. doi: 10.1016/j.pneurobio.2013.06.005

Barratt, E. S. (1959). Anxiety and impulsiveness related to psychomotor efficiency. *Percept. Mot. Skills* 9, 191–198. doi: 10.2466/pms.1959.9.3.191

Begemann, M. J. H., Florisse, E. J. R., Van Lutterveld, R., Kooyman, M., and Sommer, I. E. (2016). Efficacy of EEG neurofeedback in psychiatry: a comprehensive overview and meta-analysis. *Transl. Brain Rhythmicity* 1, 19–29. doi: 10.15761/TBR.1000105

Birbaumer, N., Elbert, T., Rockstroh, B., Daum, I., Wolf, P., and Canavan, A. (1991). "Clinical-psychological treatment of epileptic seizures: a controlled study," in *Perspectives and Promises in Clinical Psychology*, ed. A. Ehlers (New York, NY: Plenum Press), 81–96.

Blume, F., Hudak, J., Dresler, T., Ehlis, A.-C., Kühnhausen, J., Renner, T. J., et al. (2017). NIRS-based neurofeedback training in a virtual reality classroom for children with attention-deficit/hyperactivity disorder: study protocol for a randomized controlled trial. *Trials* 18, 41. doi: 10.1186/s13063-016-1769-3

Casey, B., Getz, S., and Galvan, A. (2008). The adolescent brain. *Dev. Rev.* 28, 62–77. doi: 10.1016/j.dr.2007.08.003

Cho, B.-H., Kim, S., Shin, D. I., Lee, J. H., Min Lee, S., Young Kim, I., et al. (2004). Neurofeedback training with virtual reality for inattention and impulsiveness. *Cyberpsychol. Behav.* 7, 519–526. doi: 10.1089/cpb.2004.7.519

Criaud, M., and Boulinguez, P. (2013). Have we been asking the right questions when assessing response inhibition in go/no-go tasks with fMRI? A meta-analysis and critical review. *Neurosci. Biobehav. Rev.* 37, 11–23. doi: 10.1016/j.neubiorev.2012.11.003

Cubillo, A., Halari, R., Smith, A., Taylor, E., and Rubia, K. (2012). A review of fronto-striatal and fronto-cortical brain abnormalities in children and adults with Attention Deficit Hyperactivity Disorder (ADHD) and new evidence for dysfunction in adults with ADHD during motivation and attention. *Cortex* 48, 194–215. doi: 10.1016/j.cortex.2011.04.007

Ding, W., Sun, J., Sun, Y.-W., Chen, X., Zhou, Y., Zhuang, Z., et al. (2014). Trait impulsivity and impaired prefrontal impulse inhibition function in adolescents with internet gaming addiction revealed by a Go/No-Go fMRI study. *Behav. Brain Funct.* 10:20. doi: 10.1186/1744-9081-10-20

Ehlis, A.-C., Bähne, C. G., Jacob, C. P., Herrmann, M. J., and Fallgatter, A. J. (2008). Reduced lateral prefrontal activation in adult patients with attention-deficit/hyperactivity disorder (ADHD) during a working memory task: a functional near-infrared spectroscopy (fNIRS) study. *J. Psychiatr. Res.* 42, 1060–1067. doi: 10.1016/j.jpsychires.2007.11.011

Fallgatter, A. J., and Herrmann, M. J. (2001). Electrophysiological assessment of impulsive behavior in healthy subjects. *Neuropsychologia* 39, 328–333. doi: 10.1016/S0028-3932(00)00115-9

Farr, O. M., Hu, S., Zhang, S., Li, C., and Shan, R. (2012). Decreased saliency processing as a neural measure of Barratt impulsivity in healthy adults. *Neuroimage* 63, 1070–1077. doi: 10.1016/j.neuroimage.2012.07.049

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., and Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210. doi: 10.1002/hbm.460020402

Gevensleben, H., Kleemeyer, M., Rothenberger, L. G., Studer, P., Flaig-Röhr, A., Moll, G. H., et al. (2014a). Neurofeedback in ADHD: further pieces of the puzzle. *Brain Topogr.* 27, 20–32. doi: 10.1007/s10548-013-0285-y

Gevensleben, H., Moll, G. H., Rothenberger, A., and Heinrich, H. (2014b). Neurofeedback in attention-deficit/hyperactivity disorder – different models, different ways of application. *Front. Hum. Neurosci.* 8:846. doi: 10.3389/fnhum.2014.00846

Gevensleben, H., Rothenberger, A., Moll, G. H., and Heinrich, H. (2012). Neurofeedback in children with ADHD: validation and challenges. *Expert Rev. Neurother.* 12, 447–460. doi: 10.1586/ern.12.22

Haeussinger, F. B., Dresler, T., Heinzel, S., Schecklmann, M., Fallgatter, A. J., and Ehlis, A. C. (2014). Reconstructing functional near-infrared spectroscopy (fNIRS) signals impaired by extra-cranial confounds: an easy-to-use filter method. *Neuroimage* 95, 69–79. doi: 10.1016/j.neuroimage.2014.02.035

Hagen, K., Ehlis, A.-C., Haeussinger, F. B., Heinzel, S., Dresler, T., Mueller, L. D., et al. (2014). Activation during the Trail Making Test measured with functional near-infrared spectroscopy in healthy elderly subjects. *Neuroimage* 85(Pt 1), 583–591. doi: 10.1016/j.neuroimage.2013.09.014

Hege, M. A., Stingl, K. T., Kullmann, S., Schag, K., Giel, K. E., Zipfel, S., et al. (2014). Attentional impulsivity in binge eating disorder modulates response inhibition performance and frontal brain networks. *Int. J. Obes.* 39, 353–360. doi: 10.1038/ijo.2014.99

Heinzel, S., Metzger, F. G., Ehlis, A.-C., Korell, R., Alboji, A., Haeussinger, F. B., et al. (2013). Aging-related cortical reorganization of verbal fluency processing: a functional near-infrared spectroscopy study. *Neurobiol. Aging* 34, 439–450. doi: 10.1016/j.neurobiolaging.2012.05.021

Henges, A. L., and Marczinski, C. A. (2012). Impulsivity and alcohol consumption in young social drinkers. *Addict. Behav.* 2, 217–220. doi: 10.1016/j.addbeh.2011.09.013

Herrmann, M. J., Saathoff, C., Schreppel, T. J., Ehlis, A. C., Scheuerpflug, P., Pauli, P., et al. (2009). The effect of ADHD symptoms on performance monitoring in a non-clinical population. *Psychiatry Res.* 169, 144–148. doi: 10.1016/j.psychres.2008.06.015

Holtmann, M., Sonuga-Barke, E., Cortese, S., and Brandeis, D. (2014). Neurofeedback for ADHD: a review of current evidence. *Child Adolesc. Psychiatr. Clin. N. Am.* 23, 789–806. doi: 10.1016/j.chc.2014.05.006

Horn, N. R., Dolan, M., Elliott, R., Deakin, J. F. W., and Woodruff, P. W. R. (2003). Response inhibition and impulsivity: an fMRI study. *Neuropsychologia* 41, 1959–1966. doi: 10.1016/S0028-3932(03)00077-0

Huppert, T. J., Diamond, S. G., Franceschini, M. A., and Boas, D. A. (2009). HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Appl. Opt.* 48, D280. doi: 10.1364/AO.48.00D280

Jasper, H. H. (1958). Report of the committee on methods of clinical examination in electroencephalography. *Electroencephalogr. Clin. Neurophysiol. Suppl.* 10, 370–375. doi: 10.1016/0013-4694(58)90053-1

Keilp, J. G., Sackeim, H. A., and Mann, J. J. (2005). Correlates of trait impulsiveness in performance measures and neuropsychological tests. *Psychiatry Res.* 135, 191–201. doi: 10.1016/j.psychres.2005.03.006

Kirkeby, B. S., and Robinson, M. D. (2005). Impulsive behavior and stimulus-response variability in choice reaction time. *J. Res. Pers.* 39, 263–277. doi: 10.1016/j.jrp.2004.04.001

Kolb, B., and Gibb, R. (2011). Brain plasticity and behaviour in the developing brain. *J. Can. Acad. Child Adolesc. Psychiatry* 20, 265–276.

Marshall, M. N. (1996). Sampling for qualitative research. *Fam. Pract.* 13, 522–526. doi: 10.1093/fampra/13.6.522

Marx, A.-M., Ehlis, A.-C., Furdea, A., Holtmann, M., Banaschewski, T., Brandeis, D., et al. (2015). Near-infrared spectroscopy (NIRS) neurofeedback as a treatment for children with attention deficit hyperactivity disorder (ADHD)—a pilot study. *Front. Hum. Neurosci.* 8:1038. doi: 10.3389/fnhum.2014.01038

Mayer, K., Wyckoff, S. N., Fallgatter, A. J., Ehlis, A.-C., and Strehl, U. (2015). Neurofeedback as a nonpharmacological treatment for adults with attention-deficit/hyperactivity disorder (ADHD): study protocol for a randomized controlled trial. *Trials* 16, 174. doi: 10.1186/s13063-015-0683-4

Molavi, B., and Dumont, G. A. (2012). Wavelet-based motion artifact removal for functional near-infrared spectroscopy. *Physiol. Meas.* 33, 259–270. doi: 10.1088/0967-3334/33/2/259

Muhlberger, A., Jekel, K., Probst, T., Schecklmann, M., Conzelmann, A., Andreatta, M., et al. (2016). The influence of methylphenidate on hyperactivity and attention deficits in children with ADHD: a virtual classroom test. *J. Atten. Disord.* doi: 10.1177/1087054716647480 [Epub ahead of print].

Obeso, I., Robles, N., Marron, E. M., and Redolar-Ripoll, D. (2013). Dissociating the role of the pre-SMA in response inhibition and switching: a combined online and offline TMS approach. *Front. Hum. Neurosci.* 7:150. doi: 10.3389/fnhum.2013.00150

Ridderinkhof, K. R., Van Den Wildenberg, W. P. M., Segalowitz, S. J., and Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn.* 56, 129–140. doi: 10.1016/j.bandc.2004.09.016

Rodrigo, A. H., Di Domenico, S. I., Ayaz, H., Gulrajani, S., Lam, J., and Ruocco, A. C. (2014). Differentiating functions of the lateral and medial prefrontal cortex in motor response inhibition. *Neuroimage* 85, 423–431. doi: 10.1016/j.neuroimage.2013.01.059

Rorden, C., and Brett, M. (2000). Stereotaxic display of brain lesions. *Behav. Neurol.* 12, 191–200. doi: 10.1155/2000/421719

Rösler, M., Retz-Junginger, P., Retz, W. W., and Stieglitz, R. D. (2008). *Homburger ADHS Skalen für Erwachsene [Homburg's ADHD Scales for Adults (HASE)]*. Gottingen: Hogrefe.

Rubia, K., Lim, L., Ecker, C., Halari, R., Giampietro, V., Simmons, A., et al. (2013). Effects of age and gender on neural networks of motor response inhibition: from adolescence to mid-adulthood. *Neuroimage* 83, 690–703. doi: 10.1016/j.neuroimage.2013.06.078

Sass, H., Wittchen, H. U., Zaudig, M., and Houben, I. (2003). *Diagnostisches und Statistisches Manual Psychischer Störungen DSM-IV-TR: Textrevision*, 1st Edn. Bern: Hogrefe Verlag.

Sherwood, M. S., Kane, J. H., Weisend, M. P., and Parker, J. G. (2016). Enhanced control of dorsolateral prefrontal cortex neurophysiology with real-time functional magnetic resonance imaging (rt-fMRI) neurofeedback training and working memory practice. *Neuroimage* 124, 214–223. doi: 10.1016/j.neuroimage.2015.08.074

Shulman, E. P., Smith, A. R., Silva, K., Icenogle, G., Duell, N., Chein, J., et al. (2016). The dual systems model: review, reappraisal, and reaffirmation. *Dev. Cogn. Neurosci.* 17, 103–117. doi: 10.1016/j.dcn.2015.12.010

Singh, A. K., Okamoto, M., Dan, H., Jurcak, V., and Dan, I. (2005). Spatial registration of multichannel multi-subject fNIRS data to MNI space without MRI. *Neuroimage* 27, 842–851. doi: 10.1016/j.neuroimage.2005.05.019

Soltaninejad, Z., Nejati, V., and Ekhtiari, H. (2015). Effect of anodal and cathodal transcranial direct current stimulation on DLPFC on modulation of inhibitory control in ADHD. *J. Atten. Disord.* doi: 10.1177/1087054715618792 [Epub ahead of print].

Spinella, M., and Miley, W. M. (2003). Impulsivity and academic achievement in college students. *Coll. Stud. J.* 37, 545–549.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Dev. Rev.* 28, 78–106. doi: 10.1016/j.dr.2007.08.002

Strangman, G., Culver, J. P., Thompson, J. H., and Boas, D. A. (2002). A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation. *Neuroimage* 17, 719–731. doi: 10.1016/S1053-8119(02)91227-9

Strehl, U. (2014). What learning theories can teach us in designing neurofeedback treatments. *Front. Hum. Neurosci.* 8:894. doi: 10.3389/fnhum.2014.00894

Tsuzuki, D., Jurcak, V., Singh, A. K., Okamoto, M., Watanabe, E., and Dan, I. (2007). Virtual spatial registration of stand-alone fNIRS data to MNI space. *Neuroimage* 34, 1506–1518. doi: 10.1016/j.neuroimage.2006.10.043

Verbruggen, F., Logan, G. D., and Stevens, M. A. (2008). STOP-IT: windows executable software for the stop-signal paradigm. *Behav. Res. Methods* 40, 479–483. doi: 10.3758/BRM.40.2.479

Weidacker, K., Whiteford, S., Boy, F., and Johnston, S. J. (2016). Response inhibition in the parametric Go/No-Go task and its relation to impulsivity and subclinical psychopathy. *Q. J. Exp. Psychol.* 70, 473–487. doi: 10.1080/17470218.2015.1135350

Whiteside, S. P., and Lynam, D. R. (2001). The five factor model and impulsivity: using a structural model of personality to understand impulsivity. *Pers. Individ. Dif.* 30, 669–689. doi: 10.1016/S0191-8869(00)00064-7

Zuberer, A., Brandeis, D., and Drechsler, R. (2015). Are treatment effects of neurofeedback training in children with ADHD related to the successful regulation of brain activity? A review on the learning of regulation of brain activity and a contribution to the discussion on specificity. *Front. Hum. Neurosci.* 9:135. doi: 10.3389/fnhum.2015.00135

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

PLOS ONE

RESEARCH ARTICLE

Functionally disconnected: A look at how study design influences neurofeedback data and mechanisms in attention-deficit/hyperactivity disorder

Justin Hudak[1,2]*, David Rosenbaum[2], Beatrix Barth[2,3], Andreas J. Fallgatter[1,2,4], Ann-Christine Ehlis[1,2]

1 LEAD Graduate School & Research Network, University of Tübingen, Tübingen, Germany, **2** Department of Psychiatry and Psychotherapy, University Hospital Tübingen, Tübingen, Germany, **3** Centre for Integrative Neuroscience, University of Tübingen, Tübingen, Germany, **4** Graduate School of Neural and Behavioral Sciences, University of Tübingen, Tübingen, Germany

* justin.hudak@med.uni-tuebingen.de

Citation: Hudak J, Rosenbaum D, Barth B, Fallgatter AJ, Ehlis A-C (2018) Functionally disconnected: A look at how study design influences neurofeedback data and mechanisms in attention-deficit/hyperactivity disorder. PLoS ONE 13(8): e0200931. https://doi.org/10.1371/journal.pone.0200931

Editor: Lutz Jäncke, University of Zurich, SWITZERLAND

Received: December 8, 2017

Accepted: July 4, 2018

Published: August 10, 2018

Data Availability Statement: All raw NIRS data files that were used for the complete analysis of this study are available from the Data Archiving and Networked Services (DANS) database (accession link: https://doi.org/10.17026/dans-xrk-hnhb).

Funding: The multicenter study was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft [DFG] project numbers FA 361/21-1 and STR 597/7-1). Ann-Christine Ehlis was partly supported by IZKF Tübingen (Junior

Abstract

Neurofeedback (NF) is a form of behavioral therapy used to treat e.g. attention-deficit/hyperactivity disorder (ADHD). Briefly, subjects are fed-back a putatively dysfunctional parameter of their brain activity in real time and must learn to control it in a suggested direction. NF protocols for ADHD have been used in practice for decades, though no clear standards on NF design have been implemented. Furthermore, studies often present only data from the general outcome of the NF treatment and do not look at how exactly the NF paradigm affects brain functionality, or what exactly the NF is training. The current study is two-fold: firstly, we look at how the functional connectivity (FC) patterns within key networks associated with ADHD differ between rests, feedback trials, success and failure in a complete functional near-infrared spectroscopy-based NF experiment on adults with ADHD. Secondly, due to methodological concerns discovered during the analysis of our data, we address important considerations in the design of NF that are often ignored in protocols being used widely in therapy and research today. In particular, we examine the common average reference and its impact on network activity as well as the importance of balancing the randomization in a design. Finally, we discuss how these methodological considerations may have influenced our FC results.

Introduction

Broadly, neurofeedback (NF) is a form of behavioral therapy in which participants must learn to control a particular parameter of their brain activity by monitoring this parameter in real time via auditory, visual or combined feedback. NF therapy has become ubiquitous in modern times, developed for everything from enhancing cognitive activity in healthy populations to treating tinnitus [1,2]. However, NF is a contentious topic in current neuroscientific research [3]. The contention regarding NF protocols arises from the complexity of the human brain;

Research Group, grant 2115-0-0). Justin Hudak was funded by the LEAD Graduate School & Research Network (GSC1028), a project of the Excellence Initiative of the German federal and state governments. We acknowledge the support of the Open Access Publishing Fund of the University of Tübingen.

Competing interests: The authors have declared that no competing interests exist.

human behavior is normally not based on any one parameter of activity, such as NF designs tend to target, but rather on a complex interplay of different brain structures and brain frequencies [4]. NF is now commonly used to treat attention-deficit/hyperactivity disorder (ADHD), a prevalent and disruptive disorder that affects roughly 2.5–5% of adults worldwide [5,6]. Symptoms include abnormal levels of impulsivity, hyperactivity, and inattention. NF with ADHD mirrors the greater NF community in that there are myriad protocols available to treat the same condition, each based on different theories about the neurobiology of the disorder.

Theta-beta frequency-band NF, slow cortical potential (SCP)-based NF, real-time functional magnetic resonance imaging NF (rt-fMRI), and functional near-infrared spectroscopy-based (fNIRS) NF have all been used in studies seeking to treat the behavioral alterations of both adult and juvenile ADHD [7–11]. Effect sizes for treatment are inconclusive, with earlier data suggesting medium to strong effects based on all prospective controlled [12] or only randomized trials [13]. However, a recent meta-analysis by Cortese et al. (2016) concluded that when NF designs are randomized, sham-controlled and double-blinded, the so-called gold standard for NF, they are not conclusively more effective than sham in treating ADHD. They further cite a lack of standardized protocols and the participants' failure to learn the feedback as potential pitfalls concerning the design of the study. It is estimated, for example, that 15–30 percent of participants fail to learn control over the feedback parameter in every feedback study [14], whereby it is assumed that successful learning of control over the feedback parameter will result in the desired behavioral modification [15]. The standardization of each type of ADHD NF protocol, i.e. having the same number of trials, the same distribution of activation/deactivation trials, the same brain region or electrode targeted, would allow for better evaluation of small changes made between the protocols to advance the state of NF with ADHD. Finally, most NF studies in ADHD focus solely on clinical outcome measures, while only a few have additionally looked at how NF paradigms affect brain functionality [16,17] or what exactly is being trained during NF trials. The present study, therefore, focuses on brain activation changes during successful and failed feedback trials and the preceding resting periods, whereby the interplay between single brain regions is specifically considered.

ADHD is considered to be a disorder of network dysfunction on a large scale [18]. Affected networks seem to be as diverse as the symptoms belonging to the disorder itself. Castellanos and Proal [19] identify seven different cognitive networks associated with deficits in ADHD compared to healthy controls, which encompass nearly the entirety of the cortex. The prefrontal cortex, an area widely associated with executive functioning, is typically under-active or under-developed in childhood and adult ADHD [20]. Furthermore, this region is vital to a brain network called the frontoparietal control network (FPCN) where it assumes connections with frontal, striatal, motoric and parietal regions to restrain impulsive behavior and allow focus on cognitively strenuous tasks.

The interplay between the default mode network (DMN) and the FPCN is perhaps the most relevant to task-based behavior in ADHD. In healthy controls, the DMN [21], composed of medial prefrontal cortex, precuneus, and parietal areas, demonstrates strong functional connectivity (FC) during resting states. Furthermore, healthy controls exhibit strong FC in the FPCN during tasks requiring a great deal of cognitive control, and when healthy controls switch between cognitively demanding tasks and rest, there is a clear switch of responsible network. FC is typically anti-correlated in the FPCN and the DMN when cognitive tasks versus resting state are compared [18,21–23]. In individuals with ADHD however, this switch is less clear or non-existent, the FPCN failing to switch to the DMN during rest and vice-versa [18]. This failure to switch causes many problems, namely failure to rest during resting periods and failure to sustain attention [24] or increased errors during cognitive tasks.

A newly emerging NF protocol for ADHD utilizes functional near-infrared spectroscopy (fNIRS) to feedback oxygenated hemoglobin (O_2Hb) activity from the prefrontal cortex, an area traditionally implicated in the disorder [8–10]. O_2Hb activity reflects activation of the underlying brain region and is the chromophore most strongly correlated with the blood-oxygenation level-dependent (BOLD) response synonymous with functional magnetic resonance imaging (fMRI) studies [25]. fNIRS is an optical imaging method that takes advantage of the special properties of near-infrared light, and its interplay with the human skull and brain matter, to image cortical activation. fNIRS affords several advantages to more traditional imaging methods such as EEG and fMRI. In particular, fNIRS provides spatial resolution that is higher than that of EEG raw data and temporal resolution that is higher than that of fMRI, balancing it nicely between the two methods when considering NF experiments [26]. It allows access to cortical hemodynamics in a similar manner to fMRI, but is much cheaper, has a much easier subject preparation phase, and allows the subject to sit in a relatively naturalistic setting, such as a comfortable chair. The recent development of portable fNIRS devices makes the potential for ecological validity greater than ever. Furthermore, fNIRS is less susceptible to motion artifacts than EEG, an ideal advantage conducting NF studies with ADHD subjects [27].

In the current study, we analyzed the differences in connectivity patterns of adults with ADHD between failed and successful NF trials, both in the rest, or preparation phase, preceding the trials and during the trials themselves. NF trials are cognitively active states [1] compared with relatively cognitively-inactive states preceding these trials (resting states). While the difference in brain activation between failure and success in NF has been studied during the course of individual trials in healthy participants [28], to our knowledge there have been no studies evaluating the differences in FC between failed and successful trials, nor in the rests preceding these trials (nor in ADHD). We predicted that, based off of the tendency of subjects with ADHD to have difficulties in switching between cognitively active and restful states, that subjects would show no significant difference in FC between resting periods and active NF, particularly in failed versus successful trials. The NF training analyzed targeted control of prefrontal cortex, therefore we expected enhanced FPCN connectivity during successful NF trials, with less clearly defined patterns during failed trials. During successful rests, we expected connectivity more similar to the DMN, with failed rests displaying FPCN activity due to a failure of task-switching.

However, as we will see in the coming methods section, several problems belonging to the design of the study need to be considered when interpreting the results. Therefore, this paper is divided into two parts: the first discussing the results of the ROI-based FC analysis, and the second offering a critique of, and recommendations for, NF study design.

Materials and methods

The present data are an excerpt taken from an extensive study [10] comparing functional near- infrared spectroscopy (fNIRS) based neurofeedback with an electroencephalography (EEG) based neurofeedback (for details see [29] and electromyography (EMG) biofeedback training as semi-active control group (for details see [30]). The study (434/2010BO1) was approved by the local Ethics Committee of the Medical Faculty of the University and University Hospital of Tuebingen and conducted according to the ethical guidelines and principles of the international Declaration of Helsinki in its latest version.

Subjects

Out of the three groups originally comprised in the study design (see above), we only focused on the fNIRS group. 19 adults with ADHD completed 30 sessions of fNIRS–based

neurofeedback (age *M* = 30.37 years, *SD* = 9.25; 6 female). Out of the 19 participants, seven were prescribed methylphenidate. All subjects were of the combined ADHD subtype, with the following subscale breakdown on the HASE-Homburger ADHD scale for adults [31]: total symptoms, M = 34.18 (S.D. = 7.43); inattention, M = 17.05 (S.D. = 5.18); hyperactivity, M = 9 (S.D. = 3.69); impulsivity, M = 8.08 (S.D. = 2.37).

Study procedure

A complete fNIRS–based neurofeedback training was comprised of 33 sessions with one to three sessions per week. Sessions 31–33 were conducted six months after completion of the initial 30 sessions to check for long-term stability of regulation ability and outcome. After 15 sessions, subjects had a three-week intermission and were instructed to practice and implement their acquired feedback strategies in everyday life. In total, mean training duration was 28.61 weeks (*SD* = 9.00; Min/Max = 12.29–49.14). fNIRS measurements (changes in oxygenated hemoglobin concentration elicited by executive functioning tasks), EEG assessments (quantitative EEG, event-related potentials in cognitive tasks) and neuropsychological assessments (symptom ratings, concentration task) were conducted preceding the first session, after 15 sessions, after completion of 30 sessions, and after six months (for details on the complete design, see [10]). Here, we consider only the fNIRS data from the initial 30 training sessions.

fNIRS neurofeedback setting

Participants sat in front of a monitor in a dark and sound-attenuated room. During the active regulation phases, they received visual feedback reflecting changes in oxygenated hemoglobin (O_2Hb) in the left and right prefrontal cortex. fNIRS feedback was recorded by means of the ETG-4000, continuous wave system (Hitachi Medical Co., Japan) which was linked to the THERA PRAX® DC-EEG-neurofeedback- and biofeedback system using a DC-EEG- and bio-signal amplifier (neuroConn GmbH, Ilmenau, Germany) and a personal computer. To calculate the input signal for the THERA PRAX,® fNIRS data were fed from the ETG-4000 to the personal computer via TCP/IP protocol for online processing using MATLAB R2011.

To cover frontal sites on both hemispheres, we used two 3×5 optode probesets (consisting of seven photodetectors and eight light emitters) resulting in 22 channels per probeset, and a total amount of 44 channels (see Fig 1). The interoptode distance was 3 cm. Sampling rate was 10 Hz. Probesets were oriented based on the international 10–20 system of electrode placement [32]. Fpz was marked as mid-point, whileT3 and T4 were used as the positions to place the rearmost channel in the lowest line of the respective probeset. The fNIRS feedback signal was computed online using a common average reference (CAR) to deal with artifacts. The CAR is traditionally used in fNIRS experiments to remove global probeset artifacts such as head motion or arousal-related blood flow [33]. Furthermore, fNIRS and EEG NF experiments commonly employ the CAR to control such artifacts, as it is a computationally efficient method [8,9,34–37]. For each data point during the regulation phase, mean changes in O2Hb of four frontal channels per probeset were calculated. In a next step, the average activity (CAR) of all channels on the respective probeset was subtracted. Finally, the resulting O_2Hb (feedback) amplitudes for each probeset (four channels on the left and four on the right; see Fig 1) were averaged.

fNIRS neurofeedback trials

Every session consisted of three blocks of fNIRS-based neurofeedback. Each session lasted approximately one hour including preparation time and was comprised of 32 min of neurofeedback training. One training session included two feedback blocks of 12 regulation trials,

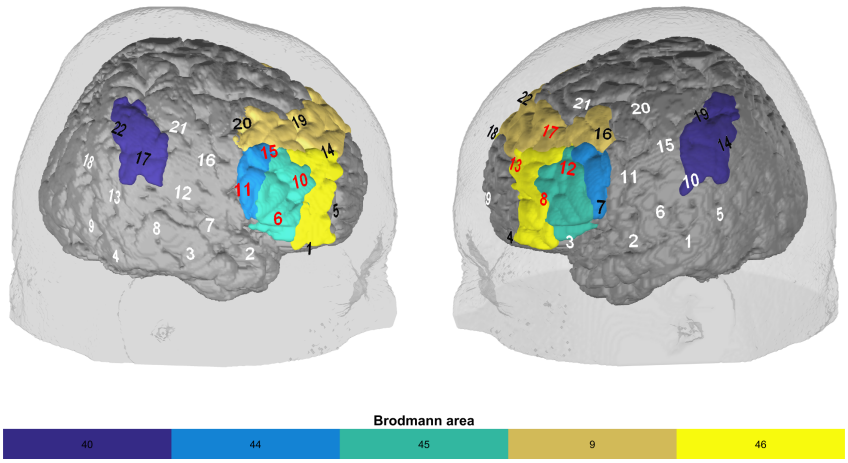


Fig 1. Probeset and regions of interest. The probeset used for the feedback training covered frontal, parietal, and temporal regions. Feedback channels are in red and covered the dlPFC (BA 9, 46) and IFG (BA 44, 45); channels in black are non-feedback channels that were also part of the ROI-based FC analysis. White channels were not included in the analysis. The supramarginal gyrus (BA 40) is also presented, as it is part of the FPCN that we included in the ROI-based FC analysis.

https://doi.org/10.1371/journal.pone.0200931.g001

each lasting 12 min, separated by an 8 min transfer block comprised of 8 regulation trials. At the beginning of each session, a 10 s baseline measurement was conducted. A feedback block comprised 12 regulation trials lasting 30 s preceded by roughly 25 s resting time and 5 s baseline measurement. The task was either to increase or decrease prefrontal O_2Hb concentration whereby up-regulation and down-regulation trials were equally likely. At the beginning of every regulation trial, a triangle was presented in the center of the computer screen oriented either upwards or downwards, indicating an activation or deactivation trial (i.e., a required increase or decrease in frontal O_2Hb concentration), respectively. Visual feedback of relative changes in O_2Hb was provided by means of an object on the screen which participants could select beforehand (e.g. a moon, a fish). Successful trials (at least 7 s of the last 15 s regulation in the desired direction) were visually reinforced by the symbol of a sun presented on the screen immediately following the trial. During transfer blocks, participants did not receive any continuous visual feedback about prefrontal oxygenation level but received reinforcement for successful trials directly after completion. The transfer condition served as the first step to transferring regulation strategies into daily life, where no direct feedback is given.

Data analyses

Calculation of pre-trial and trial fNIRS neurofeedback data. All subsequent data analysis is performed on data from all 30 training sessions. We calculated the average signal within the feedback channels for the 25 s resting period preceding feedback trials and for the 30 s trials for both activation and deactivation trials. The calculation of the average signal was as described in the 'fNIRS neurofeedback setting' section with the further averaging of all activation or deactivation trials across all subjects. In the analysis, we used all continuous feedback trials (as opposed to transfer trials) across sessions 1–30. This resulted in a grand average activation and deactivation pre-trial and trial feedback signal (see Fig 2).

Calculation of transition probabilities and adjustment of the analysis strategy. With a 50/50 activation/deactivation design, it is important that the probability of switching between an activation and deactivation trial is equal in order to ensure that no pre-baseline biases are introduced into the design. For example, if the likelihood is greater that the kind of feedback

trial switches in the next trial (i.e. activation to deactivation or deactivation to activation) then the subject may better prepare during the pre-trial, which might bias what should otherwise be a neutral preparatory phase. Indeed, during the course of analysis, we observed some anomalies in the data that led us believe that activation and deactivation trials had not been properly randomized by the neurofeedback software. Specifically, we observed a clear difference between activation and deactivation trials already in the baseline phase preceding the grand averaged feedback trials (i.e., a baseline bias effect). Namely, before activation trials it appeared that there was a decrease in feedback channel activity just before the beginning of the trial; for deactivation trials it was the opposite, an increase in feedback channel activity just before the trial (see Fig 2). That means, subjects were able to, either intentionally or unintentionally, predict the type of the next trial. With a perfectly randomized design this should not have been possible. Therefore, we decided to investigate the transitional probabilities, or the probability that the next trial type will be the same as the current one (congruent) or the opposite (incongruent).

To this end, we calculated the probability of transitioning from each type of trial to 1) the same type of trial, and 2) the opposite type of trial. To calculate statistical significance, we used a 2x2 repeated measures ANOVA with the within-subjects factors of trial type (activation vs. deactivation) and congruency (incongruent vs. congruent trials). Since this analysis confirmed a systematic bias introduced by improper randomization of trial presentation (see Results section)–and due to the effects such a bias can induce on studies of FC–we decided to analyze combined deactivation and activation trials and their preceding rests and compare success versus failure for differing patterns of FC for the remaining data analysis.

Learning rates. In order to calculate the influence of the transition probabilities on the rate of learning for subjects over time we divided the trials into two halves: the first 15 trainings

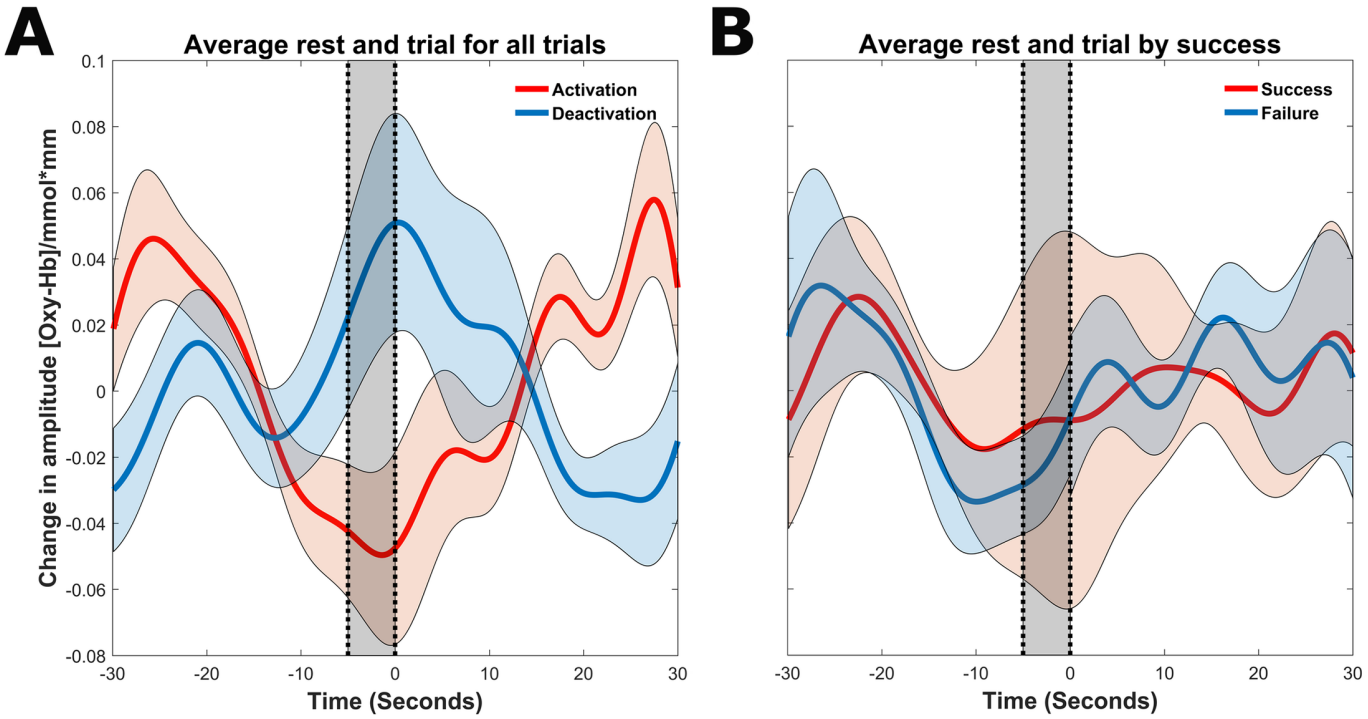


Fig 2. A. Grand average activation and deactivation trials. Trials are plotted with the standard error of the mean. The black shaded regions represents the baseline time calculated for the feedback trials. **B. Grand average success and failure trials**, averaged over all activation and deactivation trials.

https://doi.org/10.1371/journal.pone.0200931.g002

and the last 15 trainings. To calculate the learning rate for each subject, we grouped all activation and deactivation trials together and split them into congruent and incongruent trials. To test statistical significance of differences between halves and congruency, we applied a 2x2 repeated measures ANOVA with the within-subjects factors of time (first vs. second half performance) and congruency (incongruent vs. congruent). We also calculated the learning rates for all trials in the first and second halves and compared them with a paired t-test.

Functional connectivity differences between successful and failed feedback trials and the preceding rests. All preprocessing and analysis of FC was computed in MATLAB version 9.0 (The Mathworks Inc., Natick, Mass.) using routines created in our working group. The first step was to segment the trials into successful and unsuccessful trials (30 seconds) and their preceding rests (25 seconds) based on the presentation or lack of a reward. The next step was to recreate the signal that passed in the THERA PRAX® machine. To do this, a moving average was first applied to the raw O_2Hb data (five second moving window). Next, the common average from each probeset was subtracted from the signal (for more details see the above section fNIRS Neurofeedback trials). Next, the data was bandpass-filtered between .01-.1 Hz to remove potential influence from physiological artifacts. A single trial was created for each subject and each condition by concatenating all continuous feedback trials (across sessions 1–30) of said condition together. In a last preprocessing step before computing FC, a robust outlier detection algorithm was applied to each concatenated trial, removing outliers based on multivariate analysis of covariates and mean. Finally, for each subject a Pearson product-moment correlation was calculated between each pair of channels for the entirety of the trial. This resulted in a single value for each channel pair for each subject for each condition. These values were then normalized with a Fisher's r-to-Z transformation.

Statistical analysis of FC between regions of interest. Statistical analysis was performed in SPSS version 23.0 (IBM Corp, Armonk, NY). We mapped fNIRS channels to corresponding, underlying cortical areas based on a virtual registration method [38–40]. We chose regions of interest (ROIs) based on the composition of the feedback channels and the cognitive control network, which is ultimately responsible for regulating behavior during cognitively demanding tasks such as NF. We had six total ROIs: bilateral dlPFC, bilateral IFG, and bilateral parietal area. In order to test the statistical significance of FC between the regions, we computed an average in the FC between all channel pairs in the defined regions and then averaged these averages, giving one FC value for each regional pairing. We then computed 2x2 repeated measure ANOVAs for each regional pairing with the within-subjects factors of trial type (rest or trial) and success (success or failure). A Bonferroni-Holm correction was applied to resulting p-values to account for multiple comparisons.

Results

Grand average pre-trial and trial O_2Hb activation in neurofeedback channels

For a visual representation, please see Fig 2. When we investigated activation and deactivation trials separately, there was a clear bias introduced in the baseline, wherein for activation trials, the tendency of the feedback signal was to decrease drastically just before the start of the trial. In deactivation trials, the tendency was exactly the opposite, the feedback signal increased (in activation) just before the start of the trial. For both trial types, this affects the ease of achieving the feedback goal in the subsequent trial. Furthermore, it renders a study of FC virtually impossible, as this systematic effect on amplitude bleeds into the FC analysis [41–43]. When we combined all trials and separated based on success or failure, we observed no baseline–dependent effects, and thus a study of FC with combined trial types was possible.

Transition probabilities

Based on the above-reported findings of unexpected baseline differences in the fNIRS/feedback signal, we analyzed the transition probabilities between trials in detail. A 2x2 repeated measures ANOVA revealed that there was indeed an effect of transition type ($F(1,18) = 126.33$, $p < .001$, $\eta^2 = 0.875$). The likelihood of switching to the opposite trial type (incongruent trials) was significantly higher than staying with the same trial type (congruent trials) ($M_{switch} = .395$, $SD = .01$; $M_{non\text{-}switch} = .324$, $SD = .01$). There were no other significant main effects or interaction effects.

Learning rates

Learning rates showed no main effect of time or trial congruency (all $F(1,18) < 1.22$). There was a significant interaction effect of time*congruency ($F(1,18) = 9.33$, $p = .007$, $\eta^2 = 0.82$). Incongruent trials, which were significantly more likely to occur than congruent trials, were also statistically more successful in the first half of sessions than congruent trials ($M_{incongruent} = .630$, $SD = .132$; $M_{congruent} = .602$, $SD = .136$, $t(18) = 3.222$, $p = .005$). This effect disappeared in the second half of sessions ($M_{incongruent} = .627$, $SD = .121$; $M_{congruent} = .620$, $SD = .097$, $t(18) = .592$, $p = .561$). There were no differences in all trials from the first to the second halves ($M_{first} = .618$, $SD = .134$; $M_{second} = .624$, $SD = .103$, $t(18) = .154$., $p = .879$). These results are depicted in Fig 3.

ROI-based FC

After correcting for multiple comparisons, there was a significant main effect of success for connectivity between the left and the right dlPFC ($F(1,18) = 27.05$, $p < .001$, $\eta^2 = 0.600$). FC

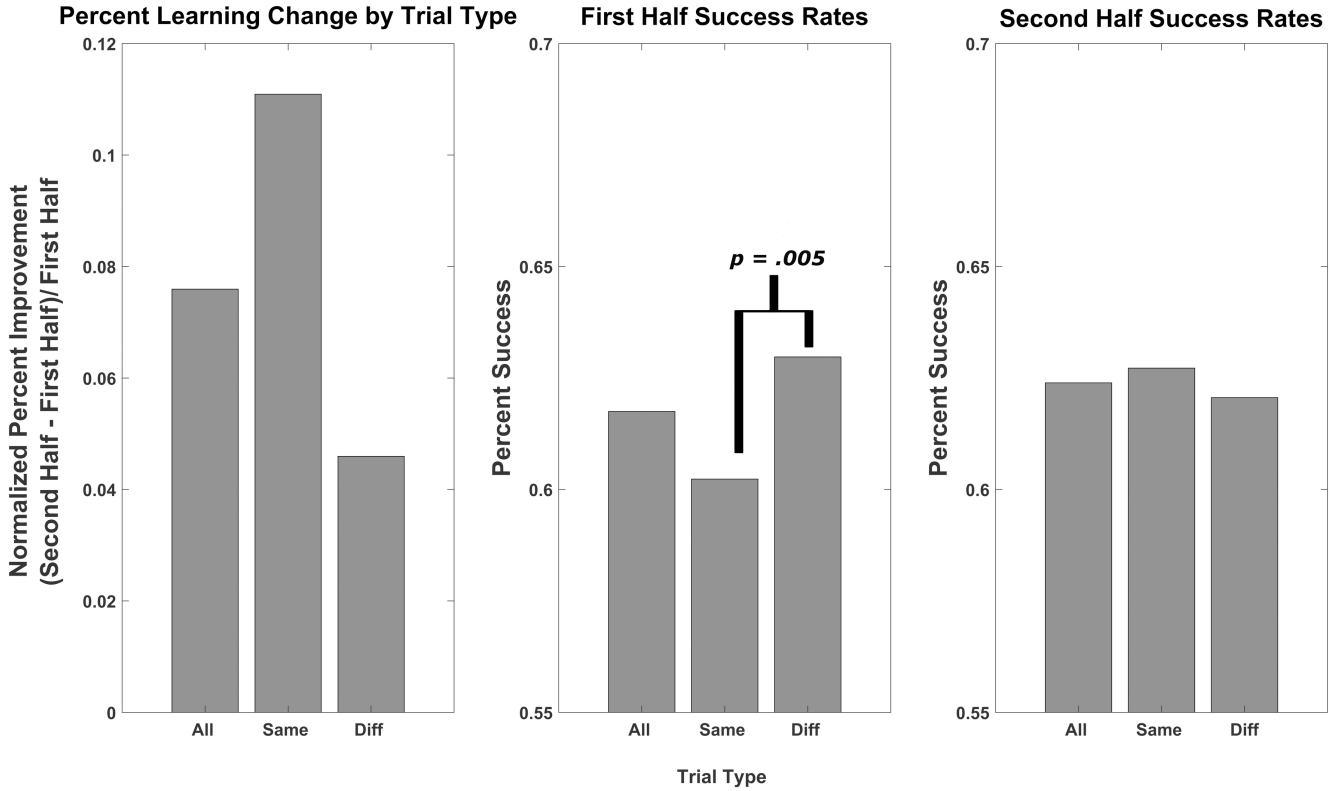


Fig 3. Learning rate between first and second half performance. The learning rates for all, congruent and incongruent trials are depicted in the left graph. The middle and right graphs depict the first and second half performances, respectively. All activation and deactivation trials were grouped together.

https://doi.org/10.1371/journal.pone.0200931.g003

was higher in the bilateral dlPFC during failed as compared to successful trials and rests ($M_{success}$ = .208, SD = .122; M_{fail} = .235, SD = .123). There was also a significant main effect of success for FC between the right IFG and the left parietal ROI ($F(1,18) = 15.08$, p = .014, η^2 = 0.456). FC was higher between the regions during successful trials and rests ($M_{success}$ = .128, SD = .100; M_{fail} = .116, SD = .096). There was also a marginally significant main effect of success for FC between the right dlPFC and the right parietal ROI ($F(1,18) = 9.63$, p = .091, η^2 = 0.342). FC between these two regions was higher during failed trials and rests than successful ones ($M_{success}$ = .185, SD = .129; M_{fail} = .207, SD = .128). There were no significant main effects for trial type or for interactions between trial type and success (see Fig 4).

Discussion

This paper–and particularly the subsequent discussion of our findings–is split into two parts: the first part addresses FC within a NF experiment for adults with ADHD, while the second part addresses the complex issue of designing a NF experiment.

In the FC analysis, we observed significant patterns of increased bilateral dlPFC connectivity and marginally significant patterns of increased FC between the right dlPFC and the right parietal ROIs during failed rests and trials. In contrast, we observed increased right IFG to left parietal connectivity in successful rests and trials. The "success network" involved significantly stronger connectivity between the right IFG and the left parietal regions. Both regions are involved in the FCPN [44], so a concurrent activation makes sense during a cognitively active task. In the "failure network(s)", the right dlPFC is centrally involved in both significantly stronger bilateral dlPFC FC and also marginally stronger FC with the right parietal ROI during failed trials and rests. Normally, the bilateral dlPFC is integrally involved in the FPCN [45–47]. Spreng et al. (2013) suggest that the dlPFC may actually be a common link between the Dorsal Attention Network and the FPCN. Furthermore, they found no FC between the dlPFC and the DMN, indicating once again its pivotal role in cognitively complex tasks. Sridharan et al.

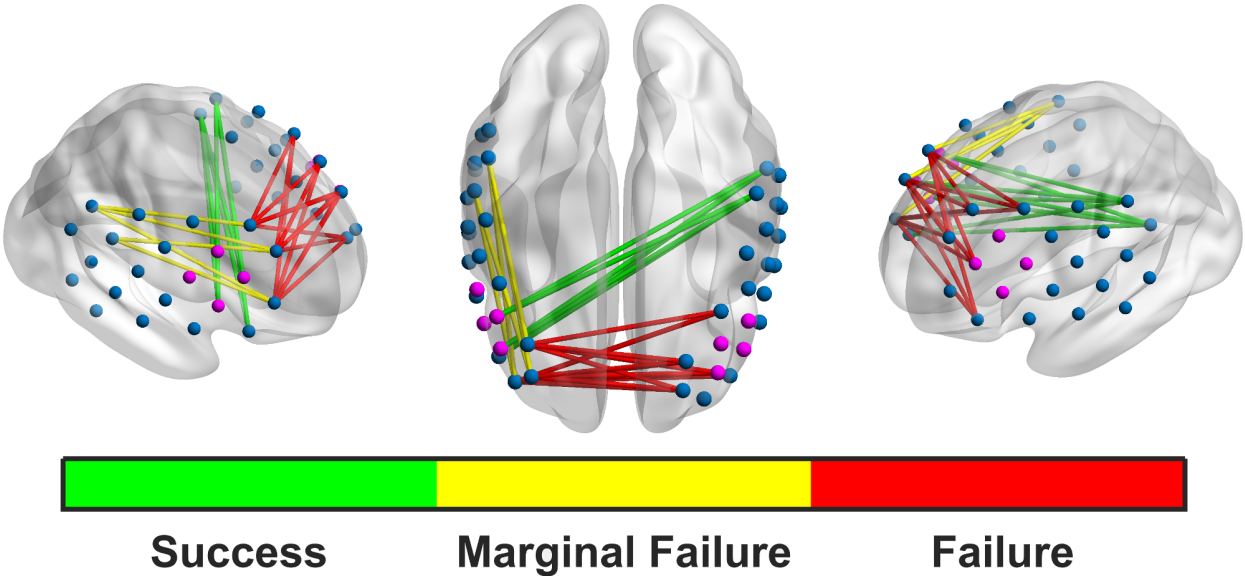


Fig 4. Significant ROI-based connectivity plots. The green Success FC occurs during successful rests and trials and is between the right IFG (making up a considerable part of the feedback channels) and left parietal lobe. The red Failure FC is between the left and right dlPFC, and is stronger on failed trials and rests. Connectivity between the right dlPFC and right parietal lobe is marginally stronger during failed rests and trials than successful ones.

https://doi.org/10.1371/journal.pone.0200931.g004

(2008) found that the right frontal insular cortex, an area that coactivates with the right dlPFC, possibly functions as a switching region, easing transition between the DMN and the FPCN. Of particular interest is that the right feedback channels are completely comprised of the right IFG. Because the right dlPFC was not involved in the feedback calculation, it may be that its activation was not encouraged, potentially hindering successful switching between rest and cognitively active trials. Furthermore, any strong activation (or deactivation in deactivation trials) in the right dlPFC (as well as other elements of the FPCN not involved in the feedback channels) would actually lower (or raise in deactivation trials) the average feedback activation due to the CAR being subtracted out. Therefore, when subtracting the CAR, the right IFG-left parietal iteration of the FPCN was the least affected. This shows us two things: first, using the CAR to isolate (training of) a particular region may not work as we intended. It appears that network activation is still influencing successful trials. Furthermore, the CAR may actually be punishing activation of other parts of the FPCN. Because of this, the right IFG-left parietal connectivity observed during successful trials emerged as the best option for activating the FPCN to achieve a successful trial given the design of the current experiment. Interestingly, in our FC analysis, there were no significant differences between rest and trials, rather only between success and failure. The successful NF trial, then, may be dependent on network interplay that begins first during the resting phase and continues into the actual feedback trial, particularly with the problems that we will soon discuss. This makes sense when we consider the resting phase to be more akin to a preparation phase for the upcoming trial. The lack of difference between rest and trial may also reflect the difficulty that subjects with ADHD have in switching between cognitively active and inactive states [18]. However, without a proper healthy control group, we cannot confirm this.

When we consider NF experimental design, we need to consider what we actually train when using a CAR in the algorithm: is the focus more on the desired training parameter or on the activity being subtracted out by the CAR? The debate surrounding the CAR stems initially from EEG research. In EEG, the CAR produces problems similar to what we observe in this NIRS experiment, but for different reasons. Nunez and Srinivasan [48] stress that the locally recorded EEG signal is always dependent on the distal reference. With a CAR, distal effects due to volume conduction will necessarily taint the true nature of the local signal, although the global artifacts will be reduced, resulting in a higher signal to noise ratio [49]. In the realm of NIRS, many studies have used or use a CAR to reduce system-wide influences, such as respiration, heartbeat and motion artifacts, on the brain signal [8,9,34,50]. Three of these studies are NIRS-based NF studies dealing with ADHD or impulsivity. These are all studies in which the CAR punished potentially helpful network activity. Nevertheless, Hudak et al. [34] and Marx et al. [9] realized beneficial results for highly impulsive and ADHD populations, respectively. It is therefore unlikely that the CAR diminished all network activity, but instead forced the network to operate differently. Still, optimization of the feedback algorithm might then allow for even better results.

The systematic bias in the pre-trial baseline of the feedback channels for both activation and deactivation trials is another important NF design consideration. An uneven distribution of transition probabilities between trial types likely caused this bias. It was significantly more likely for incongruent trials to occur than congruent ones. Therefore, the subjects were able to prepare themselves for the upcoming trial; whether this was conscious or unconscious cannot be determined from this study. Furthermore, this bias led to significant performance differential between congruent and incongruent trials in the first half. Subjects were more likely to succeed on the more probable incongruent trials than on their congruent counterparts. This differential disappeared in the second half, perhaps due to subjects having more experience with the less common congruent condition. As can be seen in Fig 3, the less common, and

therefore more difficult, congruent condition converges on the more common incongruent in the second half. However, from our dataset, it is not possible to conclude whether the 'failure network' that we identified in the FC analysis comes is linked to failed performance, or to a mismatch between what was expected and what actually happened because of the bias. Because of this bias, and its potential influence on FC, we decided to combine the activation and deactivation trials into one analysis. Interestingly, as a result of combining the trial types before the FC analysis, we show the role of the FPCN in both activation and deactivation trials. Just as the network must activate in unison to achieve success on activation trials, it must deactivate in unison to achieve deactivation success. The FC, then, should not differ between the two types.

When we combined activation and deactivation trials, we observed no pre-baseline differences in successful versus failed feedback trials, and so a comparison of FC patterns in success and failure was possible. In the present study, this bias likely affected FC nevertheless, as the preparation in the resting phase may have blended into the active trial state. While this pre-trial baseline bias may help subjects to achieve better rates of success on the actual feedback trials, it is troublesome for a few reasons. Firstly, when subjects are learning to regulate particular brain parameters, they are usually doing so implicitly through trial and error [51]. An experimental design that encourages deactivation before activation trials and activation before deactivation trials may only be training the timing of the natural neuronal signals and not actively encouraging increasing or decreasing of the intended parameter [52]. Furthermore, FC analysis is particularly influenced by amplitude changes in the signal [41,43]. When participants rapidly change the signal in anticipation of certain trial types, this can have a strong impact on connectivity, particularly when the window is short (i.e. 30 seconds or less). FC is based off of signal deviance from individual means; therefore, in a short window of calculation, these amplitude spikes will produce greater inflations in FC [42].

The reasons for the baseline bias are clear. There was a significantly greater probability of switching to a different trial type, and the baseline calculation for the coming trial was calculated as the average of the last five seconds of feedback channel activity. In isolation, either of these problems would not lead to drastic effects on the feedback trials themselves, but in combination, it produced the observed pre-baseline bias. A clear practical recommendation when moving forward is to always pseudo-randomize the trial presentation so that there is an equal chance of all trial types being next. This is a simple, but often overlooked, factor in study design. For example, the commercial NF machine used in our study is very regularly used in both scientific research and clinical practice. Very few studies analyze or report on this factor, though potentially all corresponding studies and treatments could benefit from controlling the randomization of trial type.

Otherwise, one could employ a non-local baseline calculation, or simply a much longer one that considers the entire trial. A universal baseline in the beginning of the experiment has the advantage of not being beholden to artifacts induced by local movements or spontaneous signal fluctuations, but it is more susceptible to delivering poor results in the experiment over time, as the signal is prone to drifts over time and also to displacement due to larger movements. One option that may be preferable to the local baseline calculation is to use a reference condition or a block built into the NF session. All subsequent NF trials are then compared to this reference trial instead of a baseline, thereby avoiding the pitfalls associated with local baseline bias [53,54].

One limitation of the current study was that 30 percent of the participants were taking Methylphenidate. Asking participants to cease their intake of Methylphenidate is not advisable for a study spanning such a long time-period. The effects of Methylphenidate on the BOLD response are not entirely conclusive, as depending on the brain area involved, it can both hinder and increase the response [55,56]. As this was a within-subjects analysis, the effects of Methylphenidate should have been constant throughout the entire 30 trainings.

Conclusions

In conclusion, the current study reflects several small but significant factors that have strong influences on NF design. Improper randomization or intentional unbalancing of NF trial type may cause unintended bias in the pre-feedback resting phase, particularly if paired with a local baseline. Even more potentially disruptive, the CAR needs to be carefully considered before introducing it into a feedback design that is dependent on network activity. One way forward would be to consider subtracting activity from a region of channels not connected to expected cognitive control networks. Of course, FC based-NF designs would also be a great way to target ADHD dysfunction.

Acknowledgments

We would like to thank all participants for their patience in taking part in such an extensive study. Additionally, we extend our gratitude to Ramona Täglich, Betti Schopp, Nina Gehrer, Lisa K. Schreiber, Stefanie Kettner and Sophie Grumbach for their help and invaluable contribution with the measurements. This research was funded by the LEAD Graduate School & Research Network [GSC1028], a project of the Excellence Initiative of the German federal and state governments. Justin Hudak is a doctoral candidate at the LEAD Graduate School & Research Network. The multicenter study, which provided the data of the NIRS feedback group, was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft [DFG] project numbers FA361/21-1 and STR 597/7-1). David Rosenbaum was partly supported by STORZ MEDICAL AG. Ann-Christine Ehlis was partly supported by IZKF Tübingen (Junior Research Group, grant 2115-0-0).

Author Contributions

Conceptualization: Justin Hudak, David Rosenbaum, Beatrix Barth, Ann-Christine Ehlis.

Data curation: Beatrix Barth.

Formal analysis: Justin Hudak, David Rosenbaum.

Funding acquisition: Andreas J. Fallgatter, Ann-Christine Ehlis.

Investigation: Justin Hudak, Beatrix Barth, Andreas J. Fallgatter.

Methodology: Justin Hudak, David Rosenbaum, Beatrix Barth.

Project administration: Andreas J. Fallgatter, Ann-Christine Ehlis.

Resources: Andreas J. Fallgatter, Ann-Christine Ehlis.

Software: David Rosenbaum.

Supervision: Andreas J. Fallgatter, Ann-Christine Ehlis.

Visualization: Justin Hudak.

Writing – original draft: Justin Hudak.

Writing – review & editing: Justin Hudak, David Rosenbaum, Beatrix Barth, Andreas J. Fallgatter, Ann-Christine Ehlis.

References

1. Gruzelier JH. EEG-neurofeedback for optimising performance. I: A review of cognitive and affective outcome in healthy participants. Neurosci Biobehav Rev. Elsevier Ltd; 2014; 44: 124–141. https://doi.org/10.1016/j.neubiorev.2013.09.015 PMID: 24125857

2. Haller S, Birbaumer N, Veit R. Real-time fMRI feedback training may improve chronic tinnitus. Eur Radiol. 2010; 20: 696–703. https://doi.org/10.1007/s00330-009-1595-z PMID: 19760238
3. Thibault RT, Lifshitz M, Raz A. Neurofeedback or neuroplacebo? Brain. 2017; 140: 862–864. https://doi.org/10.1093/brain/awx033 PMID: 28375458
4. Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ. Psychophysiological and modulatory interactions in neuroimaging. Neuroimage. 1997; 6: 218–229. https://doi.org/10.1006/nimg.1997.0291 PMID: 9344826
5. Kessler RC, Adler L, Barkley R, Biederman J, Conners CK, Demler O, et al. The Prevalence and Correlates of Adult ADHD in the United States: Results From the National Comorbidity Survey Replication. Am J Psychiatry. 2006; 163: 716–723. https://doi.org/10.1176/appi.ajp.163.4.716 PMID: 16585449
6. Polanczyk G, de Lima MS, Horta BL, Biederman J, Rohde LA. The Worldwide Prevalence of ADHD: a Systematic Review and Metaregression Analysis. Am J Psychiatry. 2007; 164: 942–948. https://doi.org/10.1176/ajp.2007.164.6.942 PMID: 17541055
7. Cortese S, Ferrin M, Brandeis D, Holtmann M, Aggensteiner P, Daley D, et al. Neurofeedback for Attention-Deficit/Hyperactivity Disorder: Meta-Analysis of Clinical and Neuropsychological Outcomes From Randomized Controlled Trials. J Am Acad Child Adolesc Psychiatry. 2016; 55: 444–455. https://doi.org/10.1016/j.jaac.2016.03.007 PMID: 27238063
8. Blume F, Hudak J, Dresler T, Ehlis A-C, Kühnhausen J, Renner TJ, et al. NIRS-based neurofeedback training in a virtual reality classroom for children with attention-deficit/hyperactivity disorder: study protocol for a randomized controlled trial. Trials. Trials; 2017; 18: 41. https://doi.org/10.1186/s13063-016-1769-3 PMID: 28118856
9. Marx A-M, Ehlis A-C, Furdea A, Holtmann M, Banaschewski T, Brandeis D, et al. Near-infrared spectroscopy (NIRS) neurofeedback as a treatment for children with attention deficit hyperactivity disorder (ADHD)â€"a pilot study. Front Hum Neurosci. 2015; 8: 1–13. https://doi.org/10.3389/fnhum.2014.01038 PMID: 25610390
10. Mayer K, Wyckoff SN, Fallgatter AJ, Ehlis A-C, Strehl U. Neurofeedback as a nonpharmacological treatment for adults with attention-deficit/hyperactivity disorder (ADHD): study protocol for a randomized controlled trial. Trials. ???; 2015; 16: 174. https://doi.org/10.1186/s13063-015-0683-4 PMID: 25928870
11. Alegria AA, Wulff M, Brinson H, Barker GJ, Norman LJ, Brandeis D, et al. Real-time fMRI neurofeedback in adolescents with attention deficit hyperactivity disorder. Hum Brain Mapp. 2017; 38: 3190–3209. https://doi.org/10.1002/hbm.23584 PMID: 28342214
12. Arns M, de Ridder S, Strehl U, Breteler M, Coenen A. Efficacy of Neurofeedback Treatment in ADHD: The Effects on Inattention, Impulsivity and Hyperactivity: A Meta-Analysis. Clin EEG Neurosci. 2009; 40: 180–189. https://doi.org/10.1177/155005940904000311 PMID: 19715181
13. Lofthouse N, Arnold LE, Hersch S, Hurt E, DeBeus R. A review of neurofeedback treatment for pediatric ADHD. J Atten Disord. 2012; 16: 351–72. https://doi.org/10.1177/1087054711427530 PMID: 22090396
14. Diaz Hernandez L, Rieger K, Koenig T. Low motivational incongruence predicts successful EEG resting-state neurofeedback in healthy adults. Neuroscience. 2016; 1–9. https://doi.org/10.1016/j.neuroscience.2016.12.005 PMID: 28007612
15. Holtmann M, Pniewski B, Wachtlin D, Wörz S, Strehl U. Neurofeedback in children with attention-deficit/hyperactivity disorder (ADHD)—a controlled multicenter study of a non-pharmacological treatment approach. BMC Pediatr. 2014; 14: 202. https://doi.org/10.1186/1471-2431-14-202 PMID: 25123917
16. Arns M, Heinrich H, Strehl U. Evaluation of neurofeedback in ADHD: The long and winding road. Biol Psychol. Elsevier B.V.; 2013; 1–8. https://doi.org/10.1016/j.biopsycho.2013.11.013 PMID: 24321363
17. Bluschke A, Roessner V, Beste C. Specific cognitive-neurophysiological processes predict impulsivity in the childhood attention-deficit/hyperactivity disorder combined subtype. Psychol Med. 2016; 46: 1–11. https://doi.org/10.1017/S0033291715001543
18. Posner J, Park C, Wang Z. Connecting the dots: A review of resting connectivity MRI studies in attention-deficit/hyperactivity disorder. Neuropsychol Rev. 2014; 24: 3–15. https://doi.org/10.1007/s11065-014-9251-z PMID: 24496902
19. Castellanos FX, Proal E. Large-scale brain systems in ADHD: beyond the prefrontal–striatal model. Trends Cogn Sci. 2012; 16: 17–26. https://doi.org/10.1016/j.tics.2011.11.007 PMID: 22169776
20. Arnsten AFT, Rubia K. Neurobiological circuits regulating attention, cognitive control, motivation, and emotion: Disruptions in neurodevelopmental psychiatric disorders. J Am Acad Child Adolesc Psychiatry. Elsevier Inc.; 2012; 51: 356–367. https://doi.org/10.1016/j.jaac.2012.01.008 PMID: 22449642
21. Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. A default mode of brain function. Proc Natl Acad Sci U S A. 2001; 98: 676–82. https://doi.org/10.1073/pnas.98.2.676 PMID: 11209064

22. Grady CL, Protzner AB, Kovacevic N, Strother SC, Afshin-pour B, Mcintosh AR. A Multivariate Analysis of Age-Related Differences in Default Mode and Task- Positive Networks across Multiple Cognitive Domains. 2017; https://doi.org/10.1093/cercor/bhp207

23. Raichle ME. The brain's default mode network. Annu Rev Neurosci. 2015; 38: 433–447. https://doi.org/10.1146/annurev-neuro-071013-014030 PMID: 25938726

24. Weissman DH, Roberts KC, Visscher KM, Woldorff MG. The neural bases of momentary lapses in attention. Nat Neurosci. 2006; 9: 971–8. https://doi.org/10.1038/nn1727 PMID: 16767087

25. Strangman G, Culver JP, Thompson JH, Boas DA. A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation. Neuroimage. 2002; 17: 719–731. https://doi.org/10.1016/S1053-8119(02)91227-9 PMID: 12377147

26. Mehta RK, Parasuraman R, Mckinley A, Neuroscience A. Neuroergonomics: a review of applications to physical and cognitive work. 2013; 7: 1–10. https://doi.org/10.3389/fnhum.2013.00889

27. Ehlis A-C, Schneider S, Dresler T, Fallgatter AJ. Application of functional near-infrared spectroscopy in psychiatry. Neuroimage. Elsevier Inc.; 2014; 85 Pt 1: 478–88. https://doi.org/10.1016/j.neuroimage.2013.03.067 PMID: 23578578

28. Radua J, Stoica T, Scheinost D, Pittenger C, Hampson M. Neural correlates of success and failure signals during neurofeedback learning. Neuroscience. IBRO; 2016; https://doi.org/10.1016/j.neuroscience.2016.04.003 PMID: 27063101

29. Mayer K, Blume F, Wyckoff SN, Brokmeier LL, Strehl U. Neurofeedback of slow cortical potentials as a treatment for adults with Attention Deficit-/Hyperactivity Disorder. Clin Neurophysiol. International Federation of Clinical Neurophysiology; 2016; 127: 1374–1386. https://doi.org/10.1016/j.clinph.2015.11.013 PMID: 26684900

30. Barth B, Strehl U, Fallgatter AJ, Ehlis A-C. Near-Infrared Spectroscopy based Neurofeedback of Prefrontal Cortex Activity: A Proof-of-Concept Study. Front Hum Neurosci. 2016; 10: 633. https://doi.org/10.3389/fnhum.2016.00633 PMID: 28018199

31. Rösler M, Retz-Junginger P, Retz WW, Stieglitz RD. Homburger ADHS Skalen für Erwachsene [Homburg's ADHD scales for adults] (HASE). Gottingen: Hogrefe; 2008.

32. Jasper HH. Report of the committee on methods of clinical examination in electroencephalography. Electroencephalogr Clin Neurophysiol Suppl. 1958; 10: 370–375. https://doi.org/10.1016/0013-4694(58)90053-1

33. Heinzel S, Metzger FG, Ehlis A-C, Korell R, Alboji A, Haeussinger FB, et al. Aging-related cortical reorganization of verbal fluency processing: a functional near-infrared spectroscopy study. Neurobiol Aging. Elsevier Inc.; 2013; 34: 439–450. https://doi.org/10.1016/j.neurobiolaging.2012.05.021 PMID: 22770542

34. Hudak J, Blume F, Dresler T, Haeussinger FB, Renner TJ, Fallgatter AJ, et al. Near-Infrared Spectroscopy-Based Frontal Lobe Neurofeedback Integrated in Virtual Reality Modulates Brain and Behavior in Highly Impulsive Adults. Front Hum Neurosci. 2017; 11: 425. https://doi.org/10.3389/fnhum.2017.00425 PMID: 28928644

35. Mayer K, Arns M. Electroencephalogram Neurofeedback: Application in ADHD and Epilepsy. Psychiatr Ann. 2016; 46: 594–600. https://doi.org/10.3928/00485713-20160906-01

36. Schönenberg M, Wiedemann E, Schneidt A, Scheeff J, Logemann A, Keune PM, et al. Neurofeedback, sham neurofeedback, and cognitive-behavioural group therapy in adults with attention-deficit hyperactivity disorder: a triple-blind, randomised, controlled trial. The Lancet Psychiatry. 2017; 366: 1–12. https://doi.org/10.1016/S2215-0366(17)30291-2

37. Egner T, Gruzelier JH. The Temporal Dynamics of Electroencephalographic Responses to Alpha / Theta Neurofeedback Training in Healthy Subjects. J Neurother. 2004; 8: 37–41. https://doi.org/10.1300/J184v08n01

38. Rorden C, Brett M. Stereotaxic Display of Brain Lesions. Behav Neurol. 2000; 12: 191–200. https://doi.org/10.1155/2000/421719 PMID: 11568431

39. Singh AK, Okamoto M, Dan H, Jurcak V, Dan I. Spatial registration of multichannel multi-subject fNIRS data to MNI space without MRI. Neuroimage. 2005; 27: 842–851. https://doi.org/10.1016/j.neuroimage.2005.05.019 PMID: 15979346

40. Tsuzuki D, Jurcak V, Singh AK, Okamoto M, Watanabe E, Dan I. Virtual spatial registration of standalone fNIRS data to MNI space. Neuroimage. 2007; 34: 1506–1518. https://doi.org/10.1016/j.neuroimage.2006.10.043 PMID: 17207638

41. Friston KJ. Functional and Effective Connectivity: A Review. Brain Connect. 2011; 1: 13–36. https://doi.org/10.1089/brain.2011.0008 PMID: 22432952

42. Hutchison RM, Womelsdorf T, Allen EA, Bandettini PA, Calhoun VD, Corbetta M, et al. Dynamic functional connectivity: Promise, issues, and interpretations. Neuroimage. Elsevier Inc.; 2013; 80: 360–378. https://doi.org/10.1016/j.neuroimage.2013.05.079 PMID: 23707587

43. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage. Elsevier Inc.; 2012; 59: 2142–2154. https://doi.org/10.1016/j.neuroimage.2011.10.018 PMID: 22019881

44. Godwin CA, Hunter MA, Bezdek MA, Lieberman G, Elkin-Frankston S, Romero VL, et al. Functional connectivity within and between intrinsic brain networks correlates with trait mind wandering. Neuropsychologia. Elsevier Ltd; 2017; 103: 140–153. https://doi.org/10.1016/j.neuropsychologia.2017.07.006 PMID: 28705691

45. Fassbender C, Simoes-Franklin C, Murphy K, Hester R, Meaney J, Robertson IH, et al. The Role of a Right Fronto-Parietal Network in Cognitive Control. J Psychophysiol. 2006; 20: 286–296. https://doi.org/10.1027/0269-8803.20.4.286

46. Spreng RN, Sepulcre J, Turner GR, W. Dale Stevens A, Schacter L. D. Intrinsic architecture underlying the relations among the default, dorsal attention, and frontoparietal control networks of the human brain. J Cogn Neurosci. 2013; 25: 1–19. https://doi.org/10.1162/jocn_a_00281 PMID: 22905821

47. Sridharan D, Levitin DJ, Menon V. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. Proc Natl Acad Sci. 2008; 105: 12569–12574. https://doi.org/10.1073/pnas.0800005105 PMID: 18723676

48. Nunez PL, Srinivasan R. Electric Fields of the Brain: The neurophysics of EEG. Electric Fields of the Brain: The neurophysics of EEG. 2009. https://doi.org/10.1093/acprof:oso/9780195050387.001.0001

49. Jaušovec N, Jaušovec K. Resting brain activity: Differences between genders. Neuropsychologia. 2010; 48: 3918–3925. https://doi.org/10.1016/j.neuropsychologia.2010.09.020 PMID: 20875436

50. Plichta MM, Heinzel S, Ehlis AC, Pauli P, Fallgatter AJ. Model-based analysis of rapid event-related functional near-infrared spectroscopy (NIRS) data: A parametric validation study. Neuroimage. Elsevier Inc.; 2007; 35: 625–634. https://doi.org/10.1016/j.neuroimage.2006.11.028 PMID: 17258472

51. Strehl U. What learning theories can teach us in designing neurofeedback treatments. Front Hum Neurosci. 2014; 8: 894. https://doi.org/10.3389/fnhum.2014.00894 PMID: 25414659

52. Abarbanel A. Gates, States, Rhythms, and Resonance: J Neurother. 1995; 1: 15–38. https://doi.org/10.1300/J184v01n02_02

53. Soekadar S, Witkowski M, Mellinger J, Ramos A, Birbaumer N, Cohen L. ERD-based online brain-computer-interface (BCI) in the context of neurorehabilitation: optimizing BCI learning and performance. 2011;

54. Aranyi G, Cavazza M, Charles F. Using fNIRS for Prefrontal-Asymmetry Neurofeedback: Methods and Challenges. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2015. pp. 7–20. https://doi.org/10.1007/978-3-319-24917-9_2

55. Rao SM, Salmeron BJ, Durgerian S, Janowiak JA, Fischer M, Risinger RC, et al. Effects of Methylphenidate on Functional MRI Blood-Oxygen-Level-Dependent Contrast. Am J Psychiatry. 2000; 157: 1697–1699. https://doi.org/10.1176/appi.ajp.157.10.1697 PMID: 11007731

56. Dodds CM, Muller U, Clark L, van Loon A, Cools R, Robbins TW. Methylphenidate Has Differential Effects on Blood Oxygenation Level-Dependent Signal Related to Cognitive Subprocesses of Reversal Learning. J Neurosci. 2008; 28: 5976–5982. https://doi.org/10.1523/JNEUROSCI.1153-08.2008 PMID: 18524902

5.2 Submitted Manuscripts

Combined near-infrared spectroscopy / EEG measurement elucidates functional deficits in post-error processing in adults with ADHD

Justin Hudak[1,2*], Saskia Deppermann[2], Andreas J. Fallgatter[1,2,3], Ann-Christine Ehlis[1,2]

[1]LEAD Graduate School & Research Network, University of Tübingen, Tübingen, Germany

[2]Department of Psychiatry and Psychotherapy, University Hospital Tübingen, Tübingen, Germany

[3]Centre for Integrative Neuroscience, University of Tübingen, Tübingen, Germany

[*] Corresponding author

E-mail: justin.hudak@med.uni-tuebingen.de

Abstract

The present study aims to underline prefrontal neural activation differences between healthy adults and adults diagnosed with attention-deficit/hyperactivity disorder (ADHD) following errors during a complex Eriksen flanker task embedded within a go-nogo paradigm. Simultaneously, EEG and functional near-infrared spectroscopy (NIRS) were applied, with the intent of using the combined information to enhance analysis. Reduced cortical activation in ADHD subjects was expected on post-error trials as there is behavioural evidence suggesting they respond with increased errors and higher reaction time variability following an error (Yordanova et al., 2011).

34 healthy controls (HC, 18–41 years, M = 25.3, 19 female) and 34 ADHD patients (ADHD, 18–55 years, M = 29.3, 14 female) participated in the study. NIRS was recorded via optodes centered over frontal-parietal regions of the cortex based on 10-20 EEG electrode positions. Simultaneous EEG was recorded, and P300 amplitude measured from Pz was used for an EEG-informed NIRS analysis. We measured cortical oxygenated haemoglobin corrected with the correlation-based signal improvement algorithm following errors, both on 'go' trials and 'nogo' trials. Activation differences between HC and ADHD were determined using two general linear model-based regression analyses, one in which the normal stick function representing triggers was uniformly set to amplitudes of 1, and one in which the stick function amplitudes were modulated by the amplitudes of the co-occuring P300, denoised from signal trials by wavelet-denoising. HC showed generally more activation in successful trials following errors than ADHD in both analyses. This group difference was further enhanced in the combined EEG-informed – as compared to standard – analysis; i.e. more channels and critical regions (bilateral prefrontal cortex; motor cortices) differed significantly between HC and ADHD patients (due to a stronger post-error vs. post-correct contrast in HC).

This study lends further evidence to marked neurophysiological differences in ADHD patients. In particular, we show one exciting new method for combining the refined temporal information from EEG with the spatially superior NIRS.

Keywords:

NIRS, EEG, ADHD, post-error behavior, P300

1. Introduction

1.1 Attention-deficit/hyperactivity disorder

Attention-deficit/hyperactivity disorder (ADHD) is a pervasive disorder that affects approximately 5% of children worldwide. While often studied in this juvenile domain, less is known about the disorder in adults, 2.5% of whom continue to suffer from ADHD symptoms after childhood (Faraone et al., 2015). The hyperactive, impulsive and inattentive symptoms associated with the disorder understandably affect several domains of life from school, to work, to personal relationships (Spencer et al., 2007). Recent research has focused on the neurobiological underpinnings of such symptoms in an effort to develop treatment options focused on specific brain processes. Thereby, previous research has focused extensively on the frontal, and specifically prefrontal lobe influence in the disorder (Konrad and Eickhoff, 2010; Rubia et al., 2014). Neuropsychologically, several studies confirm working memory, executive functioning and inhibitory deficits, that are often linked to either functional or structural alterations in frontal areas of both juvenile and adult patients (Cubillo et al., 2010; Ehlis et al., 2008; Rubia et al., 1999). Moreover, ADHD symptoms found in healthy control samples are also associated with impairments in frontal brain physiology (Fallgatter and Herrmann, 2001; Herrmann et al., 2009).

1.2 Error-monitoring / Post-error behavior

One of the neurocognitive alterations that have been repeatedly described in both children and adults with ADHD concerns the action monitoring domain, specifically the processing of response errors. Neurophysiologically, errors evoke specific event-related potentials (ERPs) in electroencephalography (EEG) recordings: The error-related negativity (ERN) – or error-negativity (Ne) – that is characterized by a sharp negative deflection within the first 100 ms following an incorrect button press in response time experiments (Falkenstein et al., 1991; Gehring et al., 1993); as well as the error positivity (Pe), which usually peaks around 250 ms following a response error. While both ERPs have been associated with slightly different aspects of the action monitoring process (Falkenstein et al., 2000; Nieuwenhuis et al., 2001), they seem to derive from a similar neural source within the medial prefrontal cortex (anterior cingulate cortex and surrounding structures; e.g., Herrmann et al., 2004; O'Connell et al., 2007). Previous studies repeatedly found diminished ERN/Ne and Pe amplitudes in both children/adolescents and adults with ADHD (e.g., Liotti et al., 2005; Groen et al., 2008; Geburek et al., 2013), even though some negative findings have been reported as well (see Groom et al., 2010; Shiels & Hawk, 2010) and age effects may partly play a role (Herrmann et al., 2010). In a recent study, we tried to replicate alterations in error-

monitoring potentials in an adult ADHD sample using EEG, while additionally considering post-error processes on trials subsequent to incorrect button presses (Ehlis et al., 2018) that had previously been found to elicit behavioral instability in juvenile ADHD (Yordanova et al., 2011). While we found no such behavioral alterations in post-error data, we did observe reduced ERN/Ne amplitudes on the error trials themselves. Furthermore, even in the absence of a significant behavioral effect, we found significantly reduced amplitudes of an ERP measure reflecting preparatory processes in the intertrial interval (contingent negative variation, CNV) specifically following errors, indicating subtle alterations in post-error adaptation also in adult patients, which may have been partly compensated but were still visible at the neurophysiological level (Ehlis et al., 2018).

1.3 Current Study

Here, we report the analysis of cortical activation patterns assessed within the same study using near-infrared spectroscopy (NIRS), in order to localize changes in post-error processes within fronto-temporo-parietal networks. Based on prominent models of action monitoring and cognitive control – as well as previous functional magnetic resonance imaging (fMRI) data in healthy subjects (e.g., King, Korb, von Cramon & Ullsperger, 2010; Kerns et al., 2004) – we hypothesize that adults with ADHD exhibit reduced lateral prefrontal activation on trials following errors as compared to a matched healthy control sample. Besides a unimodal analysis of the hemodynamic (i.e., NIRS) data, multimodal integration of the two highly complementary data sets (high temporal resolution for ERPs; high spatial resolution for functional NIRS data) was also attempted by performing an EEG-informed NIRS analysis (see Abreu et al. (2018) for a recent overview of similar analysis strategies with combined EEG-fMRI data).

While much of the research concerning ADHD and error-monitoring has been done with EEG as a modality (see above), NIRS can also be used in event-related designs, offering the chance to combine the imaging modalties (Ehlis et al., 2014). NIRS measures the brain's hemodynamic response to external stimuli and can do so at a much better temporal resolution than fMRI (though with an inferior spatial resolution). NIRS also has the advantage of being easily combined with EEG and lends itself well to measurements with ADHD due to its robust resistance to movement-based artifacts in comparison to EEG and fMRI (Cui et al., 2010; Ehlis et al., 2014; Ferrari and Quaresima, 2012). Additionally, NIRS is advantageous over fMRI due to its relatively cheap implementation and the ecological validity of the measurements, which can be done in a normal room in a comfortable chair (Ferrari and Quaresima, 2012). The combination of EEG and NIRS modalities allows for a more holistic understanding of neurobiological responses: NIRS affords superior spatial resolution, while EEG has exceptional temporal resolution (Mehta et al., 2013).

Furthermore, it is possible to combine the two modalities in a novel EEG-informed NIRS analysis. Developed recently for combined fMRI-EEG analysis, the combined analysis involves the use of certain EEG parameters (e.g., amplitudes of a certain ERP for each trial) to modulate the size of the triggers during the convolution of the hemodynamic response to better predict the BOLD response in the regression analysis (Debener et al., 2006). This method itself takes advantage of recent developments in automatic single-trial ERP denoising using wavelets (Ahmadi and Quian Quiroga, 2013; Quiroga and Garcia, 2003), and therefore, EEG-informed NIRS regression has rarely been used up to this point.

In the current study, we investigate the post-error behavior in ADHD subjects in comparison to healthy controls. To achieve this end, we developed a highly complex modified Eriksen Flanker task with an embedded go/no-go component. The task was designed to cause a sufficient number of errors to aid post-error analysis. Here, we expand on the EEG analysis done by Ehlis et al. (2018) by reporting an EEG-informed NIRS analysis, using single-trial wavelet-denoised P300 amplitudes to predict the hemodynamic response from concurrent trials. We then compare this novel analysis with a standard general linear model-based NIRS analysis with a normal stick function representing the triggers. The goal of the multimodal analysis was to better predict the inter- and intraindividual variability in hemodynamic responses among subjects and therefore provide a better picture of hemodynamic activity following error trials in subjects with ADHD versus healthy controls. Because Ehlis et al. (2018) observed no difference in the amplitudes of the P300 between the two groups – but a significant effect of preceding errors on P300 amplitudes evoked by the next flanker stimulus – we predicted that the contrast between the groups would remain unchanged, but that intragroup contrasts would reflect more specific differences between post-error and post-correct trials, specifically within the frontoparietal attention network.

2. Methods

2.1 Participants

Adult outpatients with ADHD (n = 34) were recruited chiefly through university-based mailing lists. ADHD diagnosis was in accordance with the latest version of the DSM-IV (Sass et al., 2003), as well as the ADHD self-rating scale (ADHD-SB) and the Wender Utah Rating Scale (WURS-K), both of which are subscales of the “Homburger ADHS-Skalen für Erwachsene” (Homburger ADHD-Scale for Adults; Rösler et al., 2008). Comorbid axis I disorders, antisocial and borderline personality disorders were used as exclusion criteria. Mild to moderate depression (Beck Depression Inventory II; Hautzinger, Keller, & Kühner, 2006) as well as some phobias were, however, still permitted. Healthy controls (n=34) were also screened and excluded

based on psychiatric and neurological illnesses. There were no significant differences between groups in age (HC: 27.62 ± 7.43, ADHD: 30.29 ± 9.47; t66 = 1.30, NS; d = 0.31), gender (HC: 18F/16M; ADHD: 13F/21M; $\chi 2 = 1.48$, NS; V = 0.15), handedness (HC: 30R/4L; ADHD: 28R/6L; $\chi 2 = 0.47$, NS; V = 0.08), IQ (HC: 118.6 ± 15.7; ADHD: 115.7 ± 14.6; t65 = −0.79, NS; d = −0.19), smoking status (HC: 10 No/24 Yes; ADHD: 11 No/ 22 Yes; $\chi 2 = 0.12$, NS; V = 0.04), or years of education (HC: 12.88 ± 1.21; ADHD: 12.75 ± 0.76; t54 = −0.51, NS; d = −0.13). ADHD patients were more hyperactive (ADHD-SB Hyperactivity; HC: 8.7 ± 4.3; ADHD: 20.5 ± 4.9; t66 = 10.60, $p < 0.001$; d = 2.56), inattentive (ADHD-SB Inattention; HC: 10.4 ± 4.3; ADHD: 24.6 ± 4.6; t66 = 13.15, $p < 0.001$; d = 3.19), and impulsive (I7 Impulsivity; HC: 4.9 ± 2.2; ADHD: 11.9 ± 2.9; t66 = 10.66, $p < 0.001$; d = 2.72) based on self-report. This sample characteristic information was adapted from Table 1 in Ehlis et al. (2018). All participants provided informed written consent. The study was carried out in accordance with the Declaration of Helsinki in its latest revision (64th World Medical Association General Assembly, Fortaleza, Brazil) and was approved by the Ethics Committee of the University Hospital of Tübingen.

2.2 Paradigm

The paradigm was a version of the Eriksen flanker task modified to include a go/no-go element. As the task is rather complex, please see Ehlis et al. (2018) for a more complete description. Briefly, subjects had to respond to the direction of an arrow or triangle, whereby the stimulus type (arrow vs. triangle) determined the response hand (left/right) while the direction the stimulus pointed in determined the response finger (index/middle finger). Additionally, based on the color of the stimulus (blue/red), subjects had to 'go', responding to the stimulus, or 'no-go', withholding their response. The color of 'go' and 'no-go' as well as the assignment of 'stimulus type' (arrow/triangle) to the response hand (left/right) was reversed for each of the two main blocks of the experiment and the stimulus order was counterbalanced across participants. Visual feedback of correct, incorrect or correct but slow responses was presented on the computer screen after each button press. In general, subjects were presented with n=400 trials over two blocks. In the case that subjects did not commit enough errors (<10), the maximum number of trials presented was increased up to n=600.

2.3 EEG recording and analysis

EEG was recorded using a 32-channel DC-amplifier (Brain Products, Germany) with 23 Ag/AgCl ring electrodes placed according to the International 10/20-System with three additional EOG electrodes. Recording reference was FCz and all impedances were kept below 5 kΩ. Data were recorded with Brain

Vision Recorder (Brain Products, Germany) (sampling rate: 1000 Hz; online filter: 0.1–100 Hz). Near-infrared spectroscopy (NIRS) was recorded simultaneously with EEG. Main EEG analysis is presented in Ehlis et al. (2018). Here, we used only the single trial P300 ERP located at Pz from correct trials following correct trials (post-correct trials) and correct trials following error trials (post-error trials), which was determined as the most positive peak between 300 and 630 ms after presentation of the flanker stimulus.

In order to achieve denoised single trial ERPs, we used an automatic wavelet based denoising technique developed by Ahmadi et al. (2013) that incorporates a method first developed by Quiroga and Garcia (2003). Briefly, in a first step, this method automatically selects wavelet coefficients that best fit to the desired ERP using the Non Zero Trees (NZT) denoising algorithm developed by Ahmadi et al. (2013). In the second step, single trial ERPs are denoised using the previously selected wavelet coefficients using the inverse wavelet transform. This allows for optimal estimation of peak amplitude of single trial ERPs. These peak amplitudes were then stored in separate vectors per subject for all correct trials post-error and all correct post-correct trials. Later these vectors were used in the NIRS analysis to weight the trigger function of the model-based analysis (see *NIRS analysis* below).

2.4 NIRS recording

We used the ETG-4000 Continuous Wave Optical Topography System (Hitachi Medical Co., Japan) to record NIRS signal with two 3 × 5 probesets containing eight emitter and seven detector optodes forming 22 distinct measurement channels. Emitter-detector distances were kept at 3 cm and the sampling rate was 10 Hz. For the orientation of the probesets over the skull, please see Figure 1. The probesets covered large areas of pre-frontal, frontal, temporal and parietal areas. The ETG-4000 measures the relative concentration changes in oxygenated (O_2Hb) and deoxygenated (HHb) hemoglobin relative to a pre-recorded baseline. Based on the concept of neurovascular coupling, we can infer the amount of neuronal activation at a certain channel via the relative concentration of O_2Hb and HHb recorded.

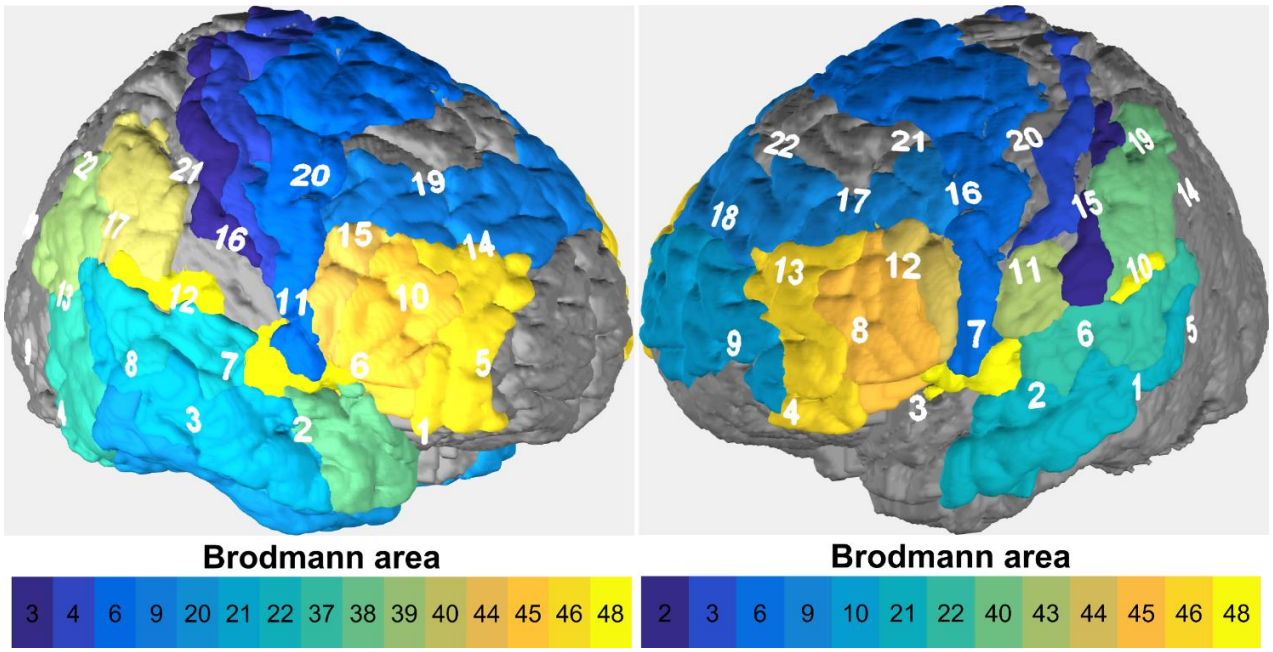


Figure 1. Channels of the probesets (white numbers) and corresponding Brodmann areas.

2.5 NIRS analysis

All data were analyzed using custom scripts programmed in MATLAB v2016b (The MathWorks, Inc.; Natick, Mass. USA). NIRS data were initially preprocessed using a bandpass filter (.01-.2 Hz). Next, the O_2Hb data were corrected for motion artifacts by implementing the correlation-based signal improvement method developed by Cui et al. (2010). Because of the large number of trials, we then used a semi-automated artifact rejection method, in which potential artifact-compromised trials were identified based off of amplitude thresholding, amplitude gradient thresholding, and correlation thresholding. When a trial contained either amplitudes exceeding .8 mmol*mm, gradients exceeding .8 mmol*mm changes over 1 second time, or a positive correlation exceeding .8 between the O_2Hb and HHb signals, it was visually inspected. In the event of local artifacts found during visual inspection, trials were then rejected. In a final step, all channels exceeding global variance greater than 3.5 times the global standard deviation were interpolated from surrounding channels. This resulted in fewer than 5% of channels being interpolated for all subjects.

In order to infer differences in cortical activation between the two groups, ADHD and control, and also between the two types of correct go response, go after-correct go trials and go after-error go trials, we analyzed NIRS data using a general linear model-based regression in which the canonical hemodynamic

response function (HRF) is convoluted with a stick function representing the triggers of each stimulus for each condition. In the basic analysis, the triggers are all weighted with an amplitude of 1. Activation changes were then calculated as the beta weights resulting from the regression. The method of least squares was used to determine the average deviance from the model HRF. To identify significant channels, we calculated one-sample t-tests contrasted against 0 for each channel and condition. The Armitage-Parmar correction for multiple comparisons was subsequently employed (Sankoh et al., 1997). Compound contrasts for correct after-error, correct after-correct, and [correct after-error – correct after-correct] were then plotted for the contrast of Healthy Controls – ADHD patients.

In order to utilize the temporal information from the simultaneously recorded EEG, we also performed a separate EEG-informed regression, in which all of the parameters of the basic analysis remained the same, except that the amplitudes of the boxcars were set to the Fisher's z normalized amplitudes of the single-trial denoised P300 taken from Pz for each respective trial.

3. Results

Full behavioral and EEG results are presented in detail in a separate manuscript. For succinct presentation, we briefly describe the relevant behavioral and EEG findings of Ehlis et al. (2018) here before presenting the new analysis.

3.1 Behavioral Data

A significant post-error slowing (PES), for both groups, on correct trials following errors versus correct trials following correct trials was observed (PES; HC: 21.59 ± 58.53; ADHD: 43.94 ± 61.64; $t66 = 1.53$, NS; $d = 0.37$). Additionally, ADHD patients displayed elevated reaction time variability (SD-RT) for all trial types (HC: 110.97 ms ± 26.26; ADHD: 127.44 ms ± 29.50; $Z = 2.49$, $p < 0.05$; $d = 0.59$) and trend-level increased error rates (HC: 40.56 errors ± 23.36; ADHD: 53.00 errors ± 28.67; $Z = 1.77$, $p < 0.1$; $d = 0.48$). Interestingly, for healthy subjects, SD-RT in trials following errors showed a significant positive correlation with impulsivity symptomatology ($r = 0.447$, $p < 0.01$), lending evidence to a relationship between ADHD symptomatology and SD-RT following errors.

3.2 EEG Data

Participants in both groups showed a significant decrease in P300 amplitude on post-error trials versus post-correct trials (post-error: 7.98 ± 4.20 μV; post-correct: 8.88 ± 3.72 μV; $F_{1,50} = 10.98$, $p < 0.01$, $\eta p^2 = 0.18$). This difference was observed at both Pz and Cz. Furthermore, P300 amplitudes correlated significantly negatively with reaction time (RT) in the ADHD group following both slow (Cz: $r = -0.570$, $p < 0.01$; Pz: $r = -0.570$, $p < 0.01$) and incorrect responses (Pz: $r = -0.474$, $p = 0.013$). SD-RT also correlated negatively with P300 amplitudes in ADHD patients following slow (Cz: $r = -0.469$, $p = 0.014$; Pz: $r = -0.455$, $p = 0.017$) and incorrect responses (Cz: $r = -0.506$, $p < 0.01$).

3.3 NIRS analysis

The basic regression analysis for the contrasted condition of post-error versus post-correct trials showed a generally stronger activation across the brain, particularly in the frontal lobe, in healthy subjects as compared to ADHD patients. One channel in the right DLPFC was significantly more activated in this contrast after correcting for multiple comparisons (Channel 14: $t = 2.372$; $p = .021$; all other channels: $t < 2.148$; $p > .036$). See Figure 2 for all complex contrasts from the basic regression analysis.

In the EEG-informed regression analysis, in the contrasted condition of post-error versus post-correct trials, there was again a generally stronger activation across the whole brain in healthy subjects compared to ADHD patients, again with an emphasis on prefrontal areas, but this time also with more pronounced differences in the motor cortex. In total, there were 6 channels that were more significantly activated in post-error trials versus post-correct trials in healthy controls. The left and right DLPFC, the right IFG, pre-motor, motor, and somatosensory cortex were all significantly more activated (Channels 4, 16, 18 in left probeset and Channels 6, 16, and 21 in right probeset: all $t > 2.121$; $p < .038$; all other channels: $t < 2.066$; $p > .043$). See Figure 3 for all complex contrasts from the EEG-informed analysis.

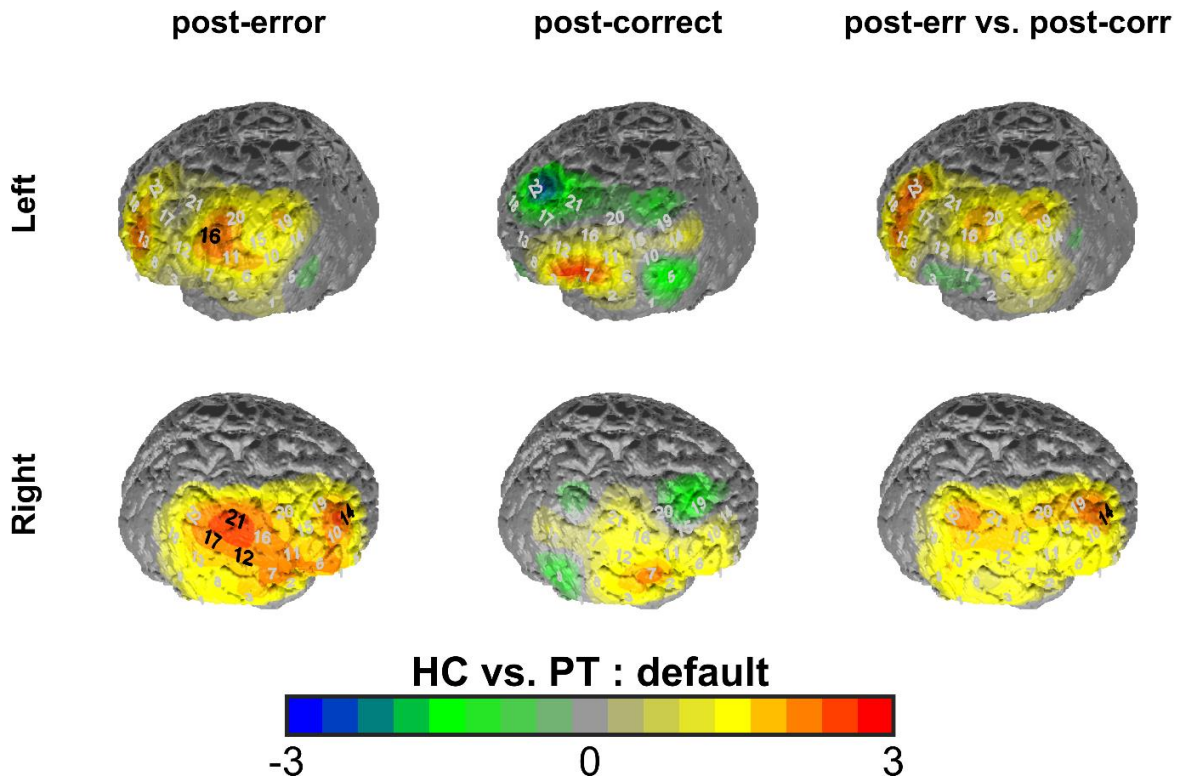


Figure 2. GLM-based regression analysis (NIRS data). The contrast healthy controls (HC) versus ADHD is plotted here for post-error, post-correct, and post-error versus post-correct trials. The plots are based off of correlation-based signal improvement ("cui") corrected oxygenated hemoglobin signals. Armitage-Parmar corrected t values of the beta weights are presented, where stronger positive t values represent stronger activation in the HC group. Significantly different channels are presented in bold.

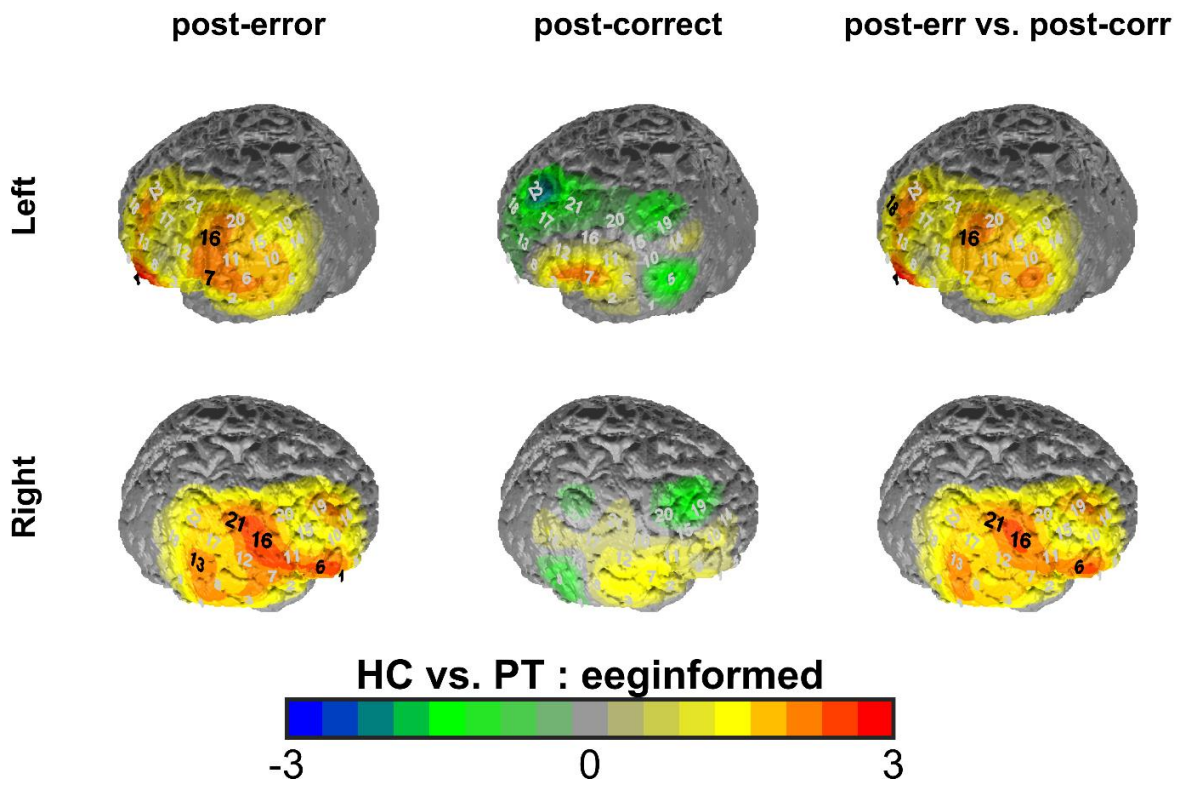
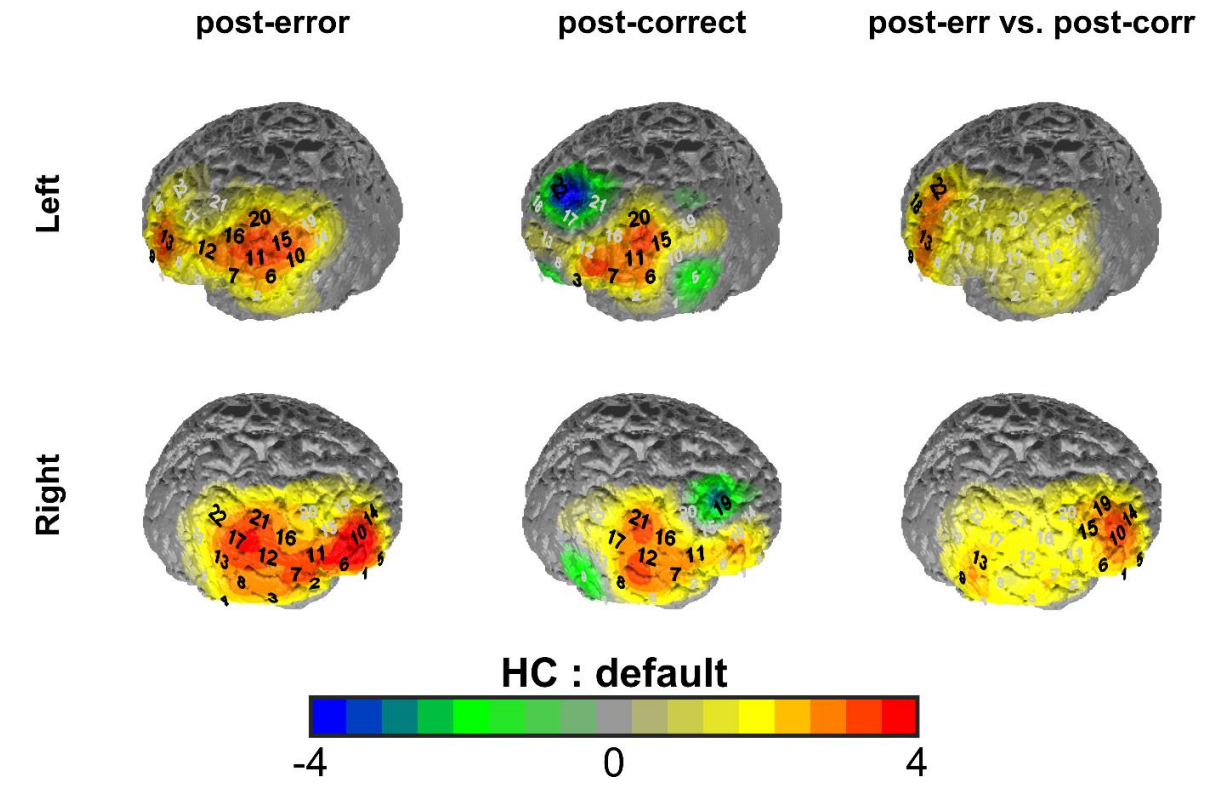


Figure 3. EEG-informed NIRS regression analysis. The contrast healthy controls (HC) versus ADHD is plotted here for post-error, post-correct, and post-error versus post-correct trials. The plots are based off of correlation-based signal improvement ("cui") corrected oxygenated hemoglobin signals. Armitage-Parmar corrected t values of the beta weights are presented, where stronger positive t values represent stronger activation in the HC group. Significantly different channels are presented in bold.

In healthy controls only, in the post-error versus post-correct contrast, 13 channels showed a more significant activation after EEG-informed analysis compared to the default analysis (EEG-informed: in total 26 channels; all $t > 2.391$; $p < .034$; all other channels: $t < 2.321$; $p > .034$; default: in total 13 channels; all $t > 2.383$; $p < .025$; all other channels: $t < 2.400$; $p > .025$; see Figure 4). These differences were most concentrated in parietal areas. In ADHD patients, there were no significant channels in the post-error versus post-correct contrast in either analysis (see Figure 5).


post-error
post-correct
post-err vs. post-corr
Left
Right
HC : default
-4
0
4

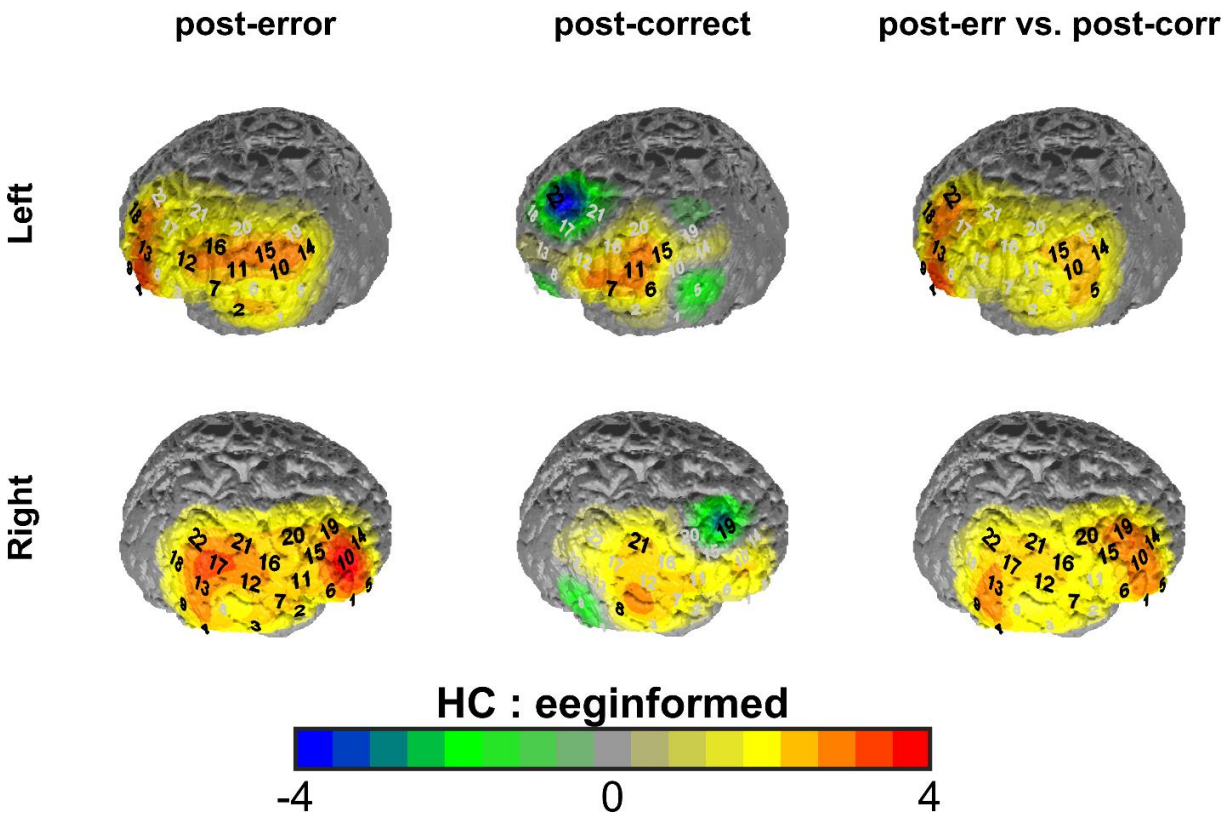

post-error
post-correct
post-err vs. post-corr
Left
Right
HC : eeginformed
-4
0
4

Figure 4. GLM-based and EEG-informed regression analyses. The individual HC group contrasts are plotted here for post-error, post-correct, and post-error versus post-correct trials. The plots are based off of correlation-based signal improvement corrected oxygenated hemoglobin signals. Armitage-Parmar corrected t values of the Beta weights are presented, where stronger positive t Values represent stronger activation in the HC group. Significantly different channels are presented in bold.

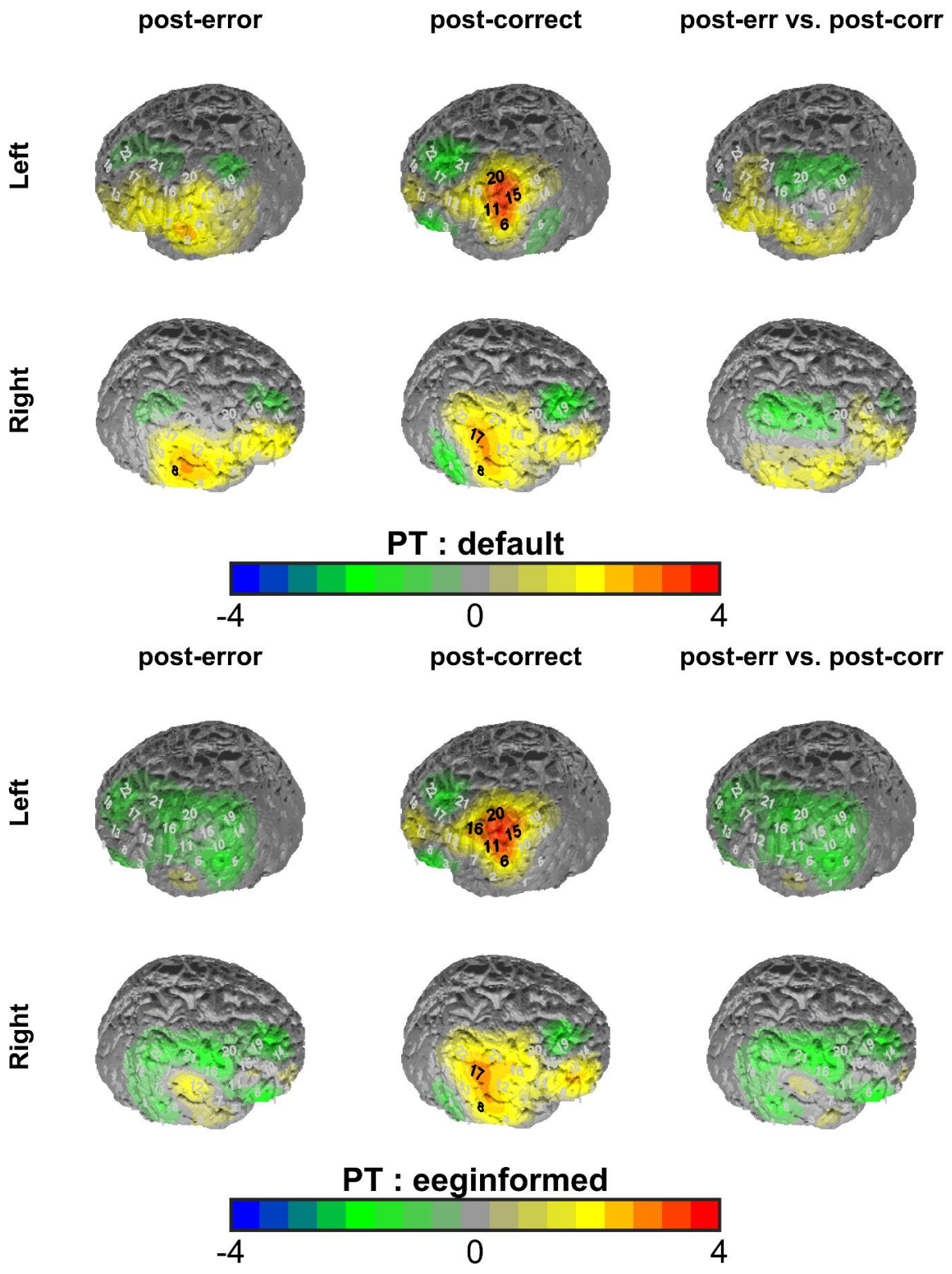


Figure 5. GLM-based and EEG-informed regression analyses. The individual ADHD group ("PT") contrasts are plotted here for post-error, post-correct, and post-error versus post-correct trials. The plots are based off of correlation-based signal improvement corrected oxygenated hemoglobin signals. Armitage-Parmar corrected t values of the Beta weights are presented, where stronger positive t values represent stronger activation in the ADHD group. Significantly different channels are presented in bold.

4. Discussion

In the current study, we investigated the neurophysiological differences between healthy controls (HC) and ADHD patients during post-error behavior, using a complex Eriksen Flanker task with a go/no-go element. The task was designed to force errors, and was adaptively elongated when not enough were committed. During the task, we recorded simultaneous EEG/NIRS, which has enabled us to utilize the best of both imaging modalities in a combined analysis of post-error, compared to post-correct hemodynamics. Our main finding was two-fold: firstly, in a standard general linear model-based regression analysis, we found that the HC group showed stronger post-error activation than the ADHD group, particularly in the right DLPFC. This alone is an interesting finding regarding post-error neurophysiology in ADHD that is in line with our a priori hypothesis, namely that post-error DLPFC activity is restrained in ADHD patients. However, using single-trial P300 amplitudes to inform the regression model proved to create an even more striking difference between the groups, elucidating ADHD processing deficits in bilateral DLPFC, right IFG, and premotor, motor and sensorimotor cortices, all of which are involved in cognitive control and response preparation.

Activity within the right DLPFC following error trials would be expected as a result of healthy (symptomatic) error-monitoring behavior (Botvinick et al., 2001; King et al., 2010; Larson et al., 2016). While the ACC (i.e. the medial prefrontal cortex) is the area of the brain most often associated with error-monitoring performance, subsequent control processes are most likely implemented by the DLPFC and surrounding areas of the lateral prefrontal cortex (Kerns et al., 2004; King et al., 2010). Reduced DLPFC activation following errors in the present sample of ADHD patients is therefore in line with behavioral observations of altered post-error adaptation in juvenile ADHD (Yordanova et al., 2011) as well as EEG data of this project (Ehlis et al., 2018) showing 1) significantly reduced ERN/Ne amplitudes for ADHD patients in both correct and incorrect trials (signaling an insufficient action-monitoring process throughout this high-conflict flanker task); and 2) specific alterations in the inter-trial interval following errors in ADHD patients who exhibited reduced amplitudes of the CNV (signaling compromised preparatory processes before the next stimulus following incorrect button presses). Based on the present finding of reduced DLPFC activation following errors, this suboptimal attentional preparation for the next trial may be related to a reduced involvement of cognitive control structures that are usually involved in

post-error behavioral adjustments (e.g., post-error slowing, improvements in accuracy) by regulating response activity via the sensorimotor cortex and task-perception via sensory cortices (Amengual et al., 2013; King et al., 2010). In the current study, interestingly, no significant post-error behavioral differences were observed between ADHD and HC groups (see also Ehlis et al., 2018). One explanation could be that adult patients have learned throughout their lifetimes to better deal with error situations, despite dysfunctional cortical error-processing. This is consistent with the reduced, but still-present impulsivity symptomology measured in adults compared to in children (Faraone et al., 2015). Another explanation could be that post-error control in ADHD is regulated by subcortical structures that cannot be measured in the scope of the current NIRS/EEG study. For instance, Cohen et al. (2000) propose a model, and provide empirical evidence for the ACC having multiple pathways of conflict signaling, one that goes through the PFC, and another that is directed towards the brainstem structure Locus Coeruleus, which then mediates further motor control.

Furthermore, when we refined the analysis by using the P300 amplitudes of concurrent trials, the contrast of HC brain activation to that of the ADHD group was made much starker on post-error versus post-correct trials. In effect, this means that the HC group's P300 amplitudes, taken from Pz, are more predictive of the BOLD response than in the ADHD group, in particular in the bilateral DLPFC, the right IFG, and sensorimotor and motor cortices. This pattern of reduced activation fits very well with previous research on motor inhibition and selective attention in adult ADHD, categories under which our paradigm falls perfectly. Adults with ADHD show deactivation in IFG and DLPFC-frontal-striatal connections while showing increased activation in cerebellum and occipital cortex, compared to healthy controls on these tasks (Cubillo and Rubia, 2010; Rubia et al., 2014). This cerebellar-occipital compensation could explain why the ADHD group showed no egregious behavioral differences on this particular task. Furthermore, an analysis of the same types of tasks, differentiating between children and adults with ADHD, showed deficits in thalamus and basal ganglia for children, while their adult counterparts showed underactivated sensorimotor areas, again an area highlighted in our refined EEG-informed analysis (Hart et al., 2013). The P300 is an indicator of both attention to a target stimulus and of updating working memory to attend to task demands (Dinteren et al., 2014); therefore, the prefrontal regions should be involved for cognitive control, while the sensorimotor and motor regions update working memory and prepare and execute a response (D'Esposito and Postle, 2015; Woods et al., 2002).

When looking at the two groups individually, the ADHD group shows almost no difference between the standard and the EEG-informed regressions. In fact, there is slightly less activation overall in the EEG-

informed analysis. Interestingly, in contrast, the HC group's brain maps, particularly in the contrast of post-error trials versus post-correct trials, show significant activation increase in the EEG-informed analysis, indicating a stronger relationship for HC than ADHD between the generator of the P300 signal and the prefrontal, temporal, parietal and motor cortices, brain regions crucial for producing response readiness and efficiency. Many EEG studies find reduced amplitudes of the P300b in juvenile subjects with ADHD compared to healthy controls (for a review see Barry et al., 2003). Indeed, Johnstone and Barry (1996) found that ADHD subjects showed superior frontal P300 in conjunction with a reduced P300b, compared to healthy controls, suggesting perhaps some sort of frontal compensation. However, in adolescents and adults, differences in P300 amplitudes seem to disappear in frontal as well as parietal sources (Lazzaro et al., 2001, 1997; Prox et al., 2007). Perhaps one explanation for the lack of influence that the EEG-informed analysis had on the ADHD subjects could be that different compensatory mechanisms were learned by the ADHD group over the lifespan, and these mechanisms are not related to the P300. The P300 is believed to have diffuse cortical and subcortical origins, particularly deep within the medial temporal lobe, the temporo-parietal junction, and hippocampus and thalamus (for a review see: Fonarayova Key et al., 2005). As this P300 tends to normalize in adult ADHD, it may be that its generator lies for ADHD within these subcortical structures that also normalize in activation with age, explaining the disconnect between the P300 and super-cortical structures seen in our study (Cubillo et al., 2012, 2010; Rubia et al., 2014).

In conclusion, our study highlights two main findings: that the ADHD group has reduced right dlPFC activity on post-error trials when compared to healthy controls, and that the P300, an ERP reflective of stimulus response readiness, seems to better predict the BOLD response in the HC, rather than the ADHD groups. This EEG-informed analysis highlights several prefrontal, motor, and sensorimotor brain regions in HC that are more active than ADHD subjects following post-error versus post-correct trials. No specific post-error behavioral differences were found in the ADHD group, suggesting that they were nevertheless able to compensate somewhat for reduced activity. We speculated that the ADHD group may have a subcortical generator of the P300 which matures differently than in healthy controls and is thus less predictive of cortical activation. However, a major limitation of our study is that we cannot determine how exactly the ADHD subjects compensated. Perhaps future studies could explore a similar paradigm using the spatially more resolved fMRI in combination with EEG, which would provide insight into sub-cortical structures that are surely also implemented in the task. Still, the combined analysis we performed in the current study expands upon the information provided using only EEG or NIRS as measurement modalities and offers an interesting way of analyzing future combined measurement studies.

5. Acknowledgements

The authors have no acknowledgements.

6. Funding

This research was funded by the LEAD Graduate School & Research Network [GSC1028], a project of the Excellence Initiative of the German federal and state governments. Justin Hudak is a doctoral candidate at the LEAD Graduate School & Research Network. Ann-Christine Ehlis was partly supported by IZKF Tübingen (Junior Research Group, grant 2115-0-0).

7. References

Ahmadi, M., Quian Quiroga, R., 2013. Automatic denoising of single-trial evoked potentials. Neuroimage 66, 672–680. https://doi.org/10.1016/j.neuroimage.2012.10.062

Amengual, J.L., Oung, S., Schweikard, A., 2013. TRACKING POST-ERROR ADAPTATION IN THE MOTOR SYSTEM BY TRANSCRANIAL MAGNETIC STIMULATION. Neuroscience 250, 342–351. https://doi.org/10.1016/j.neuroscience.2013.07.024

Barry, R.J., Clarke, A.R., Johnstone, S.J., 2003. A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography. Clin. Neurophysiol. 114, 171–183. https://doi.org/10.1016/S1388-2457(02)00362-0

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D., 2001. Conflict Monitoring and Cognitive Control Despite the importance of these efforts to characterize the func- tion of cognitive control, most of them share an important limita- tion in scope. Most current theories focus nearly exclusively on the. Psychol. Rev. 108, 624–652. https://doi.org/10.1037//0033-295X.I08.3.624

Cohen, J.D., Botvinick, M., Carter, C.S., 2000. Anterior cingulate and prefrontal cortex: who's in control? Nat. Neurosci. 3, 421.

Cubillo, A., Halari, R., Ecker, C., Giampietro, V., Taylor, E., Rubia, K., 2010. Reduced activation and inter-regional functional connectivity of fronto-striatal networks in adults with childhood Attention-Deficit Hyperactivity Disorder (ADHD) and persisting symptoms during tasks of motor inhibition and cognitive switching. J. Psychiatr. Res. 44, 629–639. https://doi.org/10.1016/j.jpsychires.2009.11.016

Cubillo, A., Halari, R., Smith, A., Taylor, E., Rubia, K., 2012. A review of fronto-striatal and fronto-cortical brain abnormalities in children and adults with Attention Deficit Hyperactivity Disorder (ADHD) and new evidence for dysfunction in adults with ADHD during motivation and attention. Cortex 48, 194–215. https://doi.org/10.1016/j.cortex.2011.04.007

Cubillo, A.I., Rubia, K., 2010. Structural and functional brain imaging in adult attention-deficit/hyperactivity disorder. Expert Rev. Neurother. 10, 603–620. https://doi.org/10.1586/ern.10.4

Cui, X., Bray, S., Reiss, A.L., 2010. Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. Neuroimage 49, 3039–3046. https://doi.org/10.1016/j.neuroimage.2009.11.050

D'Esposito, M., Postle, B.R., 2015. The Cognitive Neuroscience of Working Memory. Annu. Rev. Psychol. 66, 115–142. https://doi.org/10.1146/annurev-psych-010814-015031

Debener, S., Ullsperger, M., Siegel, M., Engel, A.K., 2006. Single-trial EEG – fMRI reveals the dynamics of cognitive function 10. https://doi.org/10.1016/j.tics.2006.09.010

Dinteren, R. Van, Arns, M., Jongsma, M.L.A., Kessels, R.P.C., 2014. P300 Development across the Lifespan : A Systematic Review and Meta-Analysis 9. https://doi.org/10.1371/journal.pone.0087347

Ehlis, A.-C., Bähne, C.G., Jacob, C.P., Herrmann, M.J., Fallgatter, A.J., 2008. Reduced lateral prefrontal activation in adult patients with attention-deficit/hyperactivity disorder (ADHD) during a working memory task: a functional near-infrared spectroscopy (fNIRS) study. J. Psychiatr. Res. 42, 1060–7. https://doi.org/10.1016/j.jpsychires.2007.11.011

Ehlis, A.-C., Schneider, S., Dresler, T., Fallgatter, A.J., 2014. Application of functional near-infrared spectroscopy in psychiatry. Neuroimage 85 Pt 1, 478–88. https://doi.org/10.1016/j.neuroimage.2013.03.067

Fallgatter, A.J., Herrmann, M.J., 2001. Electrophysiological assessment of impulsive behavior in healthy subjects. Neuropsychologia 39, 328–333. https://doi.org/10.1016/S0028-3932(00)00115-9

Faraone, S. V., Asherson, P., Banaschewski, T., Biederman, J., Buitelaar, J.K., Ramos-Quiroga, J.A., Rohde, L.A., Sonuga-Barke, E.J.S., Tannock, R., Franke, B., 2015. Attention-deficit/hyperactivity disorder.

Nat. Rev. Dis. Prim. 1, 15020. https://doi.org/10.1038/nrdp.2015.20

Ferrari, M., Quaresima, V., 2012. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. Neuroimage 63, 921–935. https://doi.org/10.1016/j.neuroimage.2012.03.049

Fonarayova Key, A.P., Dove, G.O., Maguire, M.J., 2005. Linking barin waves to the brain: an ERP primer. Dev. Neuropsychol. 27, 183–215. https://doi.org/10.1207/s15326942dn2702

Hart, H., Radua, J., Nakao, T., Mataix-Cols, D., Rubia, K., 2013. Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: Exploring task-specific, stimulant medication, and age effects. JAMA Psychiatry 70, 185–198. https://doi.org/10.1001/jamapsychiatry.2013.277

Hautzinger, M., Keller, F., Kühner, C., 2006. Beck-Depressions-Inventar: Revision.

Herrmann, M.J., Saathoff, C., Schreppel, T.J., Ehlis, A.C., Scheuerpflug, P., Pauli, P., Fallgatter, A.J., 2009. The effect of ADHD symptoms on performance monitoring in a non-clinical population. Psychiatry Res. 169, 144–148. https://doi.org/10.1016/j.psychres.2008.06.015

Johnstone, S.J., Barry, R.J., 1996. Auditory event-related potentials to a two-tone discrimination paradigm in attention deficit hyperactivity disorder. Psychiatry Res. 64, 179–192.

Kerns, J.G., Cohen, J.D., MacDonald, A.W., Cho, R.Y., Stenger, V.A., Carter, C.S., 2004. Anterior Cingulate Conflict Monitoring and Adjustments. Science (80-.). 303, 1023–1027. https://doi.org/10.1126/science.1089910

King, J.A., Korb, F.M., Cramon, D.Y. Von, Ullsperger, M., 2010. Post-Error Behavioral Adjustments Are Facilitated by Activation and Suppression of Task-Relevant and Task-Irrelevant Information Processing 30, 12759–12769. https://doi.org/10.1523/JNEUROSCI.3274-10.2010

Konrad, K., Eickhoff, S.B., 2010. Is the ADHD brain wired differently? A review on structural and functional connectivity in attention deficit hyperactivity disorder. Hum. Brain Mapp. 31, 904–916. https://doi.org/10.1002/hbm.21058

Larson, M.J., Clayson, P.E., Keith, C.M., Hunt, I.J., Hedges, D.W., Nielsen, B.L., Call, V.R.A., 2016. Cognitive control adjustments in healthy older and younger adults : Conflict adaptation , the error-related negativity (ERN), and evidence of generalized decline with age. Biol. Psychol. 115, 50–63.

https://doi.org/10.1016/j.biopsycho.2016.01.008

Lazzaro, I., Anderson, J., Gordon, E., Clarke, S., Leong, J., Meares, R., 1997. Single trial variability within the P300 (250-500 ms) processing window in adolescents with attention deficit hyperactivity disorder. Psychiatry Res. 73, 91–101. https://doi.org/10.1016/S0165-1781(97)00107-8

Lazzaro, I., Gordon, E., Whitmont, S., Meares, R., Clarke, S., 2001. The Modulation of Late Component Event Related Potentials by Pre-Stimulus EEG Theta Activity in ADHD. Int. J. Neurosci. 107, 247–264. https://doi.org/10.3109/00207450109150688

Mehta, R.K., Parasuraman, R., Mckinley, A., Neuroscience, A., 2013. Neuroergonomics : a review of applications to physical and cognitive work 7, 1–10. https://doi.org/10.3389/fnhum.2013.00889

Prox, V., Dietrich, D.E., Zhang, Y., Emrich, H.M., Ohlmeier, M.D., 2007. Attentional processing in adults with ADHD as reflected by event-related potentials. Neurosci. Lett. 419, 236–241. https://doi.org/10.1016/j.neulet.2007.04.011

Quiroga, R.Q., Garcia, H., 2003. Single-trial event-related potentials with wavelet denoising. Clin. Neurophysiol. 114, 376–390. https://doi.org/10.1016/S1388-2457(02)00365-6

Rösler, M., Retz-Junginger, P., Retz, W.W., Stieglitz, R.D., 2008. Homburger ADHS Skalen für Erwachsene [Homburg's ADHD scales for adults] (HASE). Hogrefe, Gottingen.

Rubia, K., Alegria, A., Brinson, H., 2014. Imaging the ADHD brain: disorder-specificity, medication effects and clinical translation. Expert Rev. Neurother. 14, 519–538. https://doi.org/10.1586/14737175.2014.907526

Rubia, K., Overmeyer, S., Taylor, E., Brammer, M., Williams, S.C., Simmons, A., Bullmore, E.T., 1999. Hypofrontality in attention deficit hyperactivity disorder during higher-order motor control: a study with functional MRI. Am. J. Psychiatry 156, 891–896. https://doi.org/10.1176/ajp.156.6.891

Sankoh, A.J., Huque, M.F., Dubey, S.D., 1997. Some comments on frequently used multiple endpoint adjustment methods in clinical trials. Stat. Med. 16, 2529–2542. https://doi.org/10.1002/(SICI)1097-0258(19971130)16:22<2529::AID-SIM692>3.0.CO;2-J

Sass, H., Wittchen, H.U., Zaudig, M., Houben, I., 2003. Diagnostisches und Statistisches Manual Psychischer Störungen DSM-IV-TR: Textrevision, 1st ed. Hogrefe Verlag, Bern.

Spencer, T.J., Biederman, J., Mick, E., 2007. Attention-Deficit / Hyperactivity Disorder : Diagnosis ,

Lifespan , Comorbidities , and Neurobiology 32, 631–642.

Woods, S.P., Lovejoy, D.W., Ball, J.D., 2002. Neuropsychological Characteristics of Adults with ADHD: A Comprehensive Review of Initial Studies. Clin. Neuropsychol. 16, 12–34. https://doi.org/10.1076/clin.16.1.12.8336

Yordanova, J., Kolev, V., Albrecht, B., Uebel, H., Banaschewski, T., Rothenberger, A., 2011. May posterror performance be a critical factor for behavioral deficits in attention-deficit/hyperactivity disorder? Biol. Psychiatry 70, 246–254. https://doi.org/10.1016/j.biopsych.2011.02.026

5.3 Supplementary Manuscripts

Blume et al. Trials (2017) 18:41
DOI 10.1186/s13063-016-1769-3

Trials

STUDY PROTOCOL Open Access

NIRS-based neurofeedback training in a virtual reality classroom for children with attention-deficit/hyperactivity disorder: study protocol for a randomized controlled trial

Friederike Blume[1*], Justin Hudak[1], Thomas Dresler[1,2], Ann-Christine Ehlis[2], Jan Kühnhausen[1,3], Tobias J. Renner[1,4] and Caterina Gawrilow[1,3,5]

Abstract

Background: Children with attention-deficit/hyperactivity disorder (ADHD) suffer from attention deficits, motor hyperactivity, and impulsive behaviour. These impairments are experienced at home, at school, and with friends. Functional imaging studies show that ADHD behaviour and impairments in executive functions (EFs) are mirrored by aberrant neurophysiological functioning. Moreover, several studies show that ADHD behaviour, impairments in EFs, and a lack of self-control contribute to poor school performance. Non-pharmacological interventions such as neurofeedback training (NFT), for instance, aim at improving neurophysiological and neuropsychological functioning as well as behaviour. Consequently, NFT is expected to improve school performance, EFs, and self-control in children with ADHD. Generalization of acquired self-regulation skills from laboratory to real life is crucial for a transfer to everyday situations and is hypothesized to be facilitated via training using virtual reality (VR) environments. Consequently, experiencing NFT in VR is expected to yield greater effects than training in two dimensions (2D).

Methods/design: Ninety children with a clinical diagnosis of ADHD will be included in the study. Participants may be medicated or unmedicated. After random assignation to one of three conditions, all participants receive 15 training sessions of either near-infrared spectroscopy (NIRS)-based NFT in VR, NIRS-based NFT in 2D, or electromyogram-based biofeedback training in VR. ADHD symptoms, self-control, EF, health-related quality of life, school performance, and motor activity measured via parent, teacher, and child reports or objectively will be assessed before and after the intervention and at a 6 months follow-up. Furthermore, we are interested in parents' expectations about the training's effects.

Discussion: This is, to our knowledge, the first study investigating the efficacy of NFT for children with ADHD in a VR compared to a 2D environment. Furthermore, this study will contribute to the discussion about the efficacy and specific and unspecific effects of NFTs in children with ADHD. In addition to commonly assessed variables such as ADHD symptoms, NIRS and behavioural data obtained in EF measures, health-related quality of life, and parents' expectations about the intervention's effects, this study will investigate the effects on self-control, school performance, and motor activity.

(Continued on next page)

* Correspondence: friederike.blume@uni-tuebingen.de
[1]LEAD Graduate School & Research Network, University of Tübingen, Gartenstrasse 29, 72074 Tübingen, Germany
Full list of author information is available at the end of the article

(Continued from previous page)

Trial registration: ClinicalTrials.gov, NCT02572180. Registered on 19 November 2015.

Keywords: Attention-deficit/hyperactivity disorder, Neurofeedback, Biofeedback, Near-infrared spectroscopy, Electromyography, Virtual reality, Randomized controlled trial, School performance

Background

Children with attention-deficit/hyperactivity disorder (ADHD) are inattentive, hyperactive, and impulsive [1]. They also experience difficulties in waiting for rewards, planning actions, and self-controlling in situations characterized by delay [2–4]. Affecting 5% of all children worldwide, ADHD is one of the most prevalent mental disorders in children [5].

School performance in children with ADHD

The core symptoms of ADHD, namely inattention, hyperactivity, and impulsivity, are present in various settings, for instance, when working on tasks that require sustained attention or while doing homework. Hence, ADHD affects performance levels at home and at school [1]. Children with ADHD demonstrate lower school achievement as a consequence of ADHD symptoms and concomitant impairments in executive functioning (EF) when compared to children without ADHD [6]. In addition, children with ADHD are four to five times more likely to be in need of special educational services compared to children without ADHD [7]. Several studies support the notion that the ADHD symptomatology acts as a primary reason for educational underachievement [7–9]. Several studies also provide evidence that deficits in EF such as, for instance, working memory and processing speed, might be crucially involved in impaired school performance of children with ADHD (see, e.g. [10–13]). Consequently, treatment of ADHD in schoolchildren should aim at improving behaviour as well as self-control and EF to eventually improve school performance.

Neurophysiological findings in children with ADHD

Behavioural characteristics of children with ADHD are mirrored by altered cortical and subcortical activity patterns that can be measured with brain imaging techniques such as electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS) [14–16]. In EEG studies, children with ADHD show not only an increased theta/beta ratio, but also a reduced contingent negative variation (CNV) (see, e.g. [14, 17–19]). With fNIRS, Ehlis and colleagues [16] were able to provide evidence for a reduced concentration of oxygenated haemoglobin (oxy-Hb) in the ventrolateral prefrontal cortex of adults with ADHD, compared to a healthy control group, during a working memory task. This finding was replicated in children with ADHD for the inferior prefrontal cortex during a Stroop colour-word task [15]. These deviations from normal brain activity constitute neurophysiological correlates of behavioural problems and impaired EF in patients with ADHD (see, e.g. [14, 16, 20, 21]). Consequently, we assume a treatment aiming at normalizing these deviant neurophysiological patterns to improve behaviour and EF in children with ADHD.

Neurofeedback training (NFT) in children with ADHD

Neurofeedback training (NFT) sessions are interventions based on the above-mentioned neurophysiological findings. They aim at improving self-regulation on two levels: on a neurophysiological as well as on a cognitive behavioural level [22]. In NFT, brain activity is translated into simple visual or acoustic signals which are immediately fed back to the patient [23]. Depicting learning as a controlled, effortful, and explicit as well as implicit, automatic process that is influenced by cognitive-attributional variables such as motivation, allows patients to acquire techniques that allow them to self-regulate their brain activity [22]. Hence, NFT aims at facilitating phasic changes of brain activity and enhancing neurophysiological functioning [22]. In addition, NFT aims to improve self-regulation on cognitive behavioural levels; i.e. participants are required to concentrate, to sit still, to endure boredom, and not to react on impulse during the training sessions.

Studies examining the effects of EEG-based NFT show inconsistent results. For instance, Holtmann and Cortese and colleagues [24, 25] could not find evidence for an improvement of ADHD symptoms that was specifically related to the NFT itself. However, Arns and colleagues [26] found significant effects when comparing 15 studies in a meta-analysis. Furthermore, Marx and colleagues [27] showed in a pilot study that NIRS-based NFT in children with ADHD significantly reduced ADHD symptomatology after 12 training sessions. Extending beyond a mere influence of NFT on ADHD symptomatology, Meisel and colleagues [28] demonstrated that NFT significantly improved academic performance in children with ADHD. In contrast, stimulant medication could not be shown to effectively help schoolchildren in overcoming poor school performance, although it has a significant effect on improving behaviour [29]. However, further research is required to clarify the effects of NFT,

especially NIRS-based NFT. Besides effects on ADHD symptoms, school performance, EF, EF-related frontal lobe activation, health-related quality of life (HRQoL), and self-control, the present study will investigate potential moderating influences of baseline ADHD symptoms, self-control, and IQ as well as training motivation. The effects of two NIRS-based NFT types (see below) will be compared to effects from an active control condition receiving an electromyogram (EMG)-based biofeedback training (BFT).

Control conditions for NFT studies

In prior NFT research, different control conditions have been used to investigate the efficacy of NFTs. For instance, sham feedback has been implemented, but is criticized due to strong ethical concerns and participants' poor compliance to treatment [30–33]. In the present study, an active control condition receiving an EMG-based BFT will be used to illustrate specific as well as unspecific effects of NIRS-based NFT. Looking at the effects of NFT and BFT, it is important to recall that NFT aims at improving self-regulation on two levels, neurophysiological as well as cognitive behavioural [22]. As illustrated below, the latter level is also targeted in BFT. In both NIRS-based NFT and EMG-based BFT, participants are expected to acquire self-regulation skills that allow the exertion of control over a specific endogenous parameter, for instance, prefrontal activity in the NFT and activity in the musculi supraspinatus in the EMG condition. In addition, participants learn to self-regulate behaviour such as being attentive, sitting still, not reacting on impulse, and enduring boredom. Consequently, we expect both NIRS-based NFT and EMG-based BFT to yield similar behavioural effects, as participants learn to self-regulate behaviour in both conditions. However, effects related to the acquisition of self-regulation skills related to the respective endogenous parameter are uniquely attributable to the parameter itself. As only NIRS-based NFT aims at normalizing aberrant brain activity, which is assumed to constitute a neurophysiological correlate of behavioural problems in children with ADHD [15], we consequently expect larger total effects from the NIRS-based NFT than from the EMG-based BFT.

NFT in a virtual reality (VR) environment

To our knowledge, until now, no NFT study in children or adults with ADHD employing a virtual reality (VR) environment as a training setting has been conducted. However, from our perspective, there are several reasons suggesting that patients with ADHD can profit from training in a VR environment.

First, it is hypothesized that both the acquisition of self-regulation skills in the laboratory and the transfer to everyday life situations (e.g. a classroom setting) will be facilitated by training in a naturalistic VR environment. VR environments are often used in the treatment of mental disorders such as anxiety disorders and post-traumatic stress disorder, and were shown to be equally effective compared to therapies employing exposures to real-life situations [34]. Strong effects of therapies using naturalistic VR environments are attributed to the fact that various naturalistic stimuli, i.e. sounds, visual impressions, and haptic experiences, stimulate different sensory channels at once, thereby eliciting realistic psychological and behavioural responses [34]. Consequently, children with ADHD are expected to behave similarly inattentively, hyperactively, and impulsively in VR as well as in real-life classrooms. In NFT and BFT sessions, therapists may use these responses to work towards changes in behaviour by correcting inadequate, and by reinforcing appropriate behaviour, i.e. by training successful self-regulation of behaviour. Furthermore, aberrant psychological responses occurring in specific situations, for instance, an underactivation of prefrontal cortical areas in children with ADHD, are elicited by a naturalistic VR environment such as a VR classroom [34]. In NFT, but not in an EMG-based BFT, these inadequate responses are corrected as participants acquire self-regulation strategies that allow them to normalize their brain activity. Additionally, the effects of therapies employing naturalistic VR environments can be attributed to the high degree of realism that supports the transfer of skills acquired in the therapy or training to real-life situations, i.e. from a VR to a real-life classroom [34]. Consequently, we expect larger effects from NFT taking place in naturalistic VR environments than from training taking place in two-dimensional (2D) settings, as the acquisition and transfer of behavioural and psychological self-regulation skills are facilitated.

Second, after reviewing predictors and moderators of the efficacy of cognitive training, Keshavan and colleagues suggest that training motivation plays a major role [35]. This is in line with results presented by Käthner and colleagues, who provide evidence for a significant influence of motivation on task performance in the brain-computer interface [36]. The crucial role of training motivation in making cognitive training effective can be explained by findings that support the assumption that motivational state and positive mood facilitate prefrontal activation and consequently cognitive control, that is "the ability to select thoughts or actions in relation to internal goals" [37, 38]. In NFT for children with ADHD, both cognitive control and variability in prefrontal activity are essential, as participants are instructed to select thoughts that allow for a self-regulated increase or decrease of prefrontal activity. Consequently, NFT should aim at creating training

settings that foster training motivation and positive mood. According to Keshavan and colleagues, intrinsic motivation in cognitive training can best be fostered by providing a "personalized context that links cognitive training to goals of everyday life" [35]. With the naturalistic VR classroom of the present study, a personalized context of everyday life is provided and should consequently foster cognitive control and prefrontal activation of the participants. Consequently, we expect NFT and BFT taking place in a naturalistic VR environment to yield larger effects than training in 2D. Furthermore, the effects of the training are expected to be moderated by the training motivation.

Hypotheses

First, we hypothesize that NIRS-based NFT of the frontal lobe (dorsolateral prefrontal cortex (dlPFC) and EMG-based BFT improve ADHD symptoms, self-control, EF, HRQoL, school performance, and motor activity in children with ADHD independent of whether the training is conducted in 2D or VR. Second, we expect larger positive effects for NIRS-based NFT in 2D and VR than for EMG-based BFT in VR both at a post-test time point and at 6 months follow-up. Third, we expect the effects of NIRS-based NFT to be larger in the VR condition. Fourth, for NIRS-based NFT in 2D and VR, we expect an increase prefrontal in cortical activation during EF tasks at post-test and at 6 months follow-up.

Methods/design

This manuscript as well as the trial it describes are in accordance with the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidelines [39, 40]. See Additional file 1 for the SPIRIT checklist.

Participants and recruitment

We will recruit approximately 90 participants with a clinical diagnosis of ADHD (any presentation) that is given based on the 5th edition of the *Diagnostic and Statistical Manual for Mental Disorders* [41] via medical offices of paediatricians, child and youth psychologists, and psychiatrists as well as offices of occupational therapists. Furthermore, we will recruit participants via the outpatient department of the Department of Child and Adolescent Psychiatry and Psychotherapy, University Hospital Tübingen, and local school psychologists. In addition, circular emails sent to members of the University of Tübingen, websites of the authors' departments, local newspapers, and radio stations will announce the study. Information gained using the long version of the Conners 3 questionnaire for teachers and parents [42], the Strengths and Difficulties Questionnaire (SDQ-Deu) [43], and an interview with the parents are used to confirm diagnoses of ADHD. The training sessions will take place in the Department of Psychiatry and Psychotherapy at the University of Tübingen. See Table 1 for an overview of the eligibility criteria.

Randomization

The design involves three conditions ($n = 30$ per condition; $N = 90$) to which recruited children will be assigned randomly upon confirmation of all inclusion criteria. One of the principal investigators of this study who is only occasionally involved in training and testing participants executes the randomization. A block randomization procedure is applied, and balancing the conditions for age, gender, and ADHD medication stratifies the randomization.

Interventions

Two conditions involve 15 sessions of a NIRS-based NFT, one in a VR classroom setting and one in a 2D classroom setting to control for specific effects of trainings in VR and 2D. The third condition involves an EMG-based BFT in VR and constitutes a control condition that allows the evaluation of effects that are uniquely attributable to the NFT itself.

Every training session lasts approximately 60–70 min including a preparation phase at the beginning (20 min), in which the NIRS cap and optodes are fitted to the head, or the EMG electrodes are placed on both musculi supraspinatus and both mastoids. For the participants in the VR classroom setting, the head-mounted display (HMD) is mounted. The training sessions also include the NFT or BFT (45 min) and a training phase with stimulus cards at the end of the training sessions 6–15

Table 1 Eligibility criteria

Inclusion criteria	In school Grades 1–4 (age 6–10).
	Clinical diagnosis of ADHD combined, predominantly inattentive or predominantly hyperactive-impulsive presentation according to the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5)
	Written informed consent from parents/legal guardian
Exclusion criteria	IQ <70 as assessed with the Culture Fair Test (CFT) 1-R or the CFT 20-R [44, 45]
	Parent-reported diagnosis of the following: serious physical illness or chronic diseases such as pulmonary diseases, heart diseases, diabetes, hypertension, and rheumatic diseases; neurological disorders including stroke, multiple sclerosis, and epilepsy; indicated psychiatric disorders including obsessive-compulsive disorder, chronic tic disorders, Tourette's syndrome, and suicidal behaviour
	Prior or current participation in neurofeedback training (NFT)/biofeedback training (BFT)
	Other psychotherapeutic treatment or any kind of attention training, also in the course of an ergotherapeutic treatment, while participating in the study

(5 min). The stimulus cards present a screenshot of the 2D screen. Within the laboratory setting, these cards are introduced as cue stimuli associated with brain activation or increased muscular activity of the musculi supraspinatus as learned during the NFT or BFT. When employed at home, they are thought to facilitate activation and to establish an association between, for instance, doing homework and brain or muscle activation [44].

Every training session with NIRS-based NFT or EMG-based BFT consists of three blocks, of which the first and the second are with continuous performance feedback (feedback condition). In the third block, no contingent performance feedback is provided (transfer condition), which is thought to foster generalization of acquired self-regulation skills to real-life situations [45]. For the NIRS-based NFT, the first and the second blocks consist of 12 trials, the third of 8. Each trial starts with an active phase of 30 s in which the respective endogenous parameter, that is oxy-Hb in the bilateral dlPFC for NIRS-based NFT, should be regulated and is followed by a resting phase of 20 s at the end. For the EMG-based BFT, the first and the second block consist of 24 trials; the third block consists of 16 trials. Each trial starts with an active phase of 15 s in which muscular activity of the musculi supraspinatus should be regulated and is followed by a resting phase of 10 s. Durations of active and resting phases in NIRS-based NFT and EMG-based BFT conditions vary due to different response times of the respective endogenous parameters towards the beginning of self-regulation processes employed by the participants [46, 47]. Among the conditions, the amounts of trials are varied in order to obtain an equal total training duration for all conditions.

Lighting in the VR and 2D classroom provides the feedback. For the NIRS-based NFT, lighting increases with increasing activity in the bilateral dlPFC, i.e. increasing oxy-Hb, and decreases with decreasing activity, i.e. decreasing oxy-Hb. Lighting for the EMG-based BFT increases with increasing muscular activity in the right musculus supraspinatus compared to the left and decreases with increasing activity in the left musculus supraspinatus compared to the right.

During the first eight training sessions, the training follows a protocol with 50% activation and 50% deactivation trials. For the second half of the training sessions, the protocol changes to 80% activation and 20% deactivation trials. At the beginning of each trial, an arrow appearing on the blackboard of the VR or 2D classroom pointing upwards indicates an activation trial, while an arrow pointing downwards indicates a deactivation trial.

After eight training sessions, participants have a break of 2 to 3 weeks that should further support transfer to real-life settings by using stimulus cards with screenshots of the training setting. The stimulus cards show the image of the classroom that participants see during the training with the arrow pointing upwards and will be introduced in the laboratory setting at the end of training session numbers 6, 7, and 8. Participants are instructed to look at the cards, employ activation strategies they learned during the training, and imagine increasing the lighting five to six times. After this activation task, they solve a riddle appropriate for their age and knowledge. For the break, participants are instructed to practice activation at home at least once per day prior to a situation that requires sustained attention, e.g. doing homework. For the rest of the training sessions, participants are asked to continue practicing activation at home. Furthermore, they still practice with the cards at the end of every training session. The 2D and VR classroom is shown in Fig. 1. An overview of the study course is presented in Fig. 2.

Positive reinforcement

In both NFT and BFT, an animated teacher in the VR or 2D classroom provides reinforcement via positive auditory feedback if the participant performed successfully in the past trial. In addition, smileys appear on the blackboard of the classroom to provide positive reinforcement at the end of every successful trial. Success is calculated as follows. For the NIRS-based NFT, a baseline is calculated as the average oxy-Hb signal from the eight dlPFC channels (four on the left and four on the right hemisphere) in the last 5 s before the start of each trial. For the EMG-based BFT, the baseline is calculated as the average normalized EMG output (right musculus supraspinatus EMG output minus left musculus supraspinatus EMG output) in the last second before the start of each trial. Reinforcement is provided with one smiley when the participant has spent 60–69% of the time of the second half of the trial on the required side of the baseline. For the NIRS-based NFT, below the baseline is a decrease in the oxy-Hb signal from the eight dlPFC channels, and above is an increase in the oxy-Hb signal from

Fig. 1 2D and VR classroom

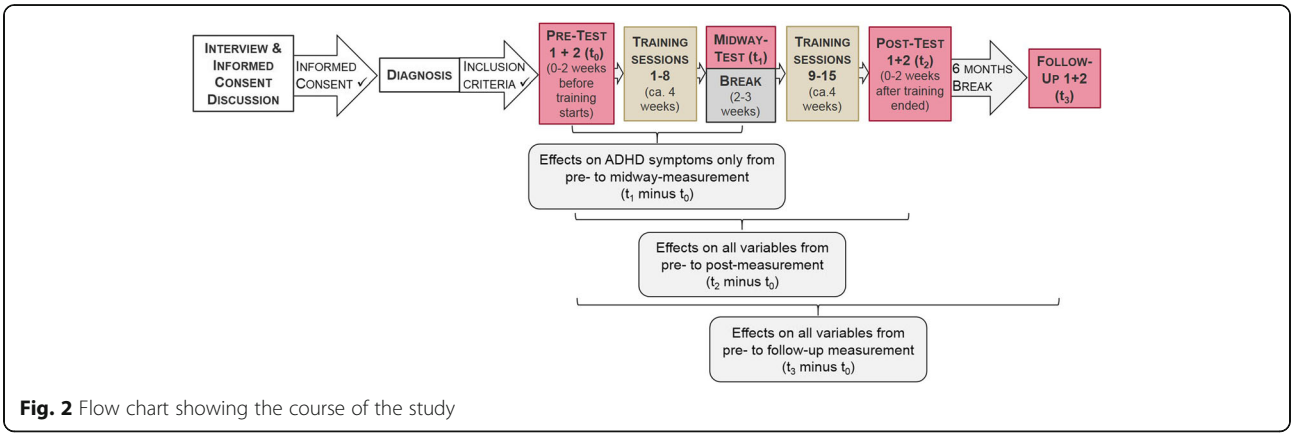


Fig. 2 Flow chart showing the course of the study

the eight dlPFC channels. For the EMG-based BFT, below the baseline is an increase in the activity in the left musculus supraspinatus compared to the right, and above is an increase in the activity in the right musculus supraspinatus compared to the left. Analogously, the participant receives two smileys with 70–79% and three smileys with at least 80% of the second half of the trial spent in the required direction. Furthermore, reinforcement for the second block changes adaptively with the performance in the first block. If the participant scored between 40 and 60% success rate in the first block, the second block will remain exactly like the first. If the participant achieves lower than a 40% success rate, the threshold will decrease to .8 standard deviations (*SD*) in either direction relative to the baseline, so that fluctuations in the light are more sensitive to performance. In addition, the threshold for receiving smileys would decrease to 50%, 60%, and 70% of the time that must be spent on the required side of the baseline, for one, two, or three smileys, respectively. If the participant scores higher than a 60% success rate in the first block, the threshold increases to 1.2 *SD* above and below the baseline, hence making changes in the lighting, requiring more relative activation or deactivation. In addition, the baseline is artificially augmented to be .1 *SD* above or below the calculated baseline. Consequently, in order to receive smileys, the participant has to maintain a stronger activation or deactivation than before. The third block will be calculated in the same way as the second.

The VR and 2D classroom

In the VR and 2D classroom, every participant is seated at a virtual table in the second row of a primary school classroom (see Fig. 1.). Other pupils surround him/her, and a teacher sits in the front of the classroom at a desk. Visual, auditory, and mixed distractors such as, for instance, paper planes flying through the room, fellow students whispering, or people knocking on the door will be randomized to occur in 50% of all trials. Their appearance is balanced between trials and sessions, and the distractors appear with a distance of at least 60 s between two distractors.

Token system

Children are rewarded for their participation. At the beginning of the study, they receive a sticker album and one sticker per test or training session in the course of the study. If participants report to have trained regularly with the stimulus cards during the break, they can earn two more stickers.

Adverse events

Reported adverse events and other unintended effects of the interventions employed in this study or the trial conduct are recorded and discussed with psychologists as well as child and youth psychiatrists.

Assessments

Culture Fair Intelligence Test 1-R and 20-R

The Culture Fair Intelligence Test 1-R (CFT 1-R) [48] is a non-verbal intelligence test that can be used for children aged 5 years and 3 months to 9 years and 11 months. It consists of five subtests on substitutions, mazes, classifications, similarities, and matrices. The test can be applied in a short and a long form that differ in testing time only, but not in the amount of subtests applied. The short form will be used in this study. Reliability scores for the subtests vary between $r = .75$ and $r = .90$, and reliability for the short form is reported to be $r = .94$. Retest reliability with a second measurement after 2.5 months is $r = .90$ [48]. For children aged 8 years and 5 months and older, the Culture Fair Intelligence Test 20-R (CFT 20-R) [49] is used. The CFT 20-R is a non-verbal intelligence test consisting of two parts, each containing four identically constructed subtests on completing series, classifications, matrices, and topological reasoning. In this study, the test is applied in its short

version, which consists of only the first of the two parts. The reliability of the CFT 20-R is $r = .92$ for the short version of the test. The retest reliability is adequate, with $r = .85$ after 2 months [49]. Hence, both tests show adequate psychometric properties to measure intelligence in the study sample.

Conners 3rd Edition (Conners 3) - German translation

The German long versions of the Conners 3 for parents (C3-P) and teachers (C3-T) consist of 108 items for parents and 112 items for teachers. The Conners 3 tools assess ADHD symptoms but also learning problems, EF, peer relationships, and aggression/defiance [42]. Furthermore, the questionnaires contain screener items for anxiety and depression. Both versions of the Conners 3 have adequate psychometric properties for teachers and parents with good internal consistency for most of the scales (Cronbach's $\alpha > .85$) and acceptable values for the remaining scales (Cronbach's $\alpha > .70$). Test-retest reliabilities are also good, with average values of about $r = .85$. Consequently, the Conners 3 in its German version seems to be appropriate to assess the effects of the intervention administered in this study.

Strengths and Difficulties Questionnaire

The Strengths and Difficulties Questionnaire (SDQ) consists of 25 items and assesses behavioural strengths and difficulties of children on five scales: prosocial behaviour, hyperactivity, emotional problems, peer relationships, and conduct problems [43]. Different versions for teachers and parents are available and are used in this study. The factorial structure of the original English questionnaires was also found for the German translations (SDQ-Deu) [50]. Psychometric properties for the German versions are good, with high internal consistency for the whole questionnaire (Cronbach's $\alpha = .82$) and at least acceptable scores for the subscales (Cronbach's α between .58 and .76). The retest reliability is specified with $r = .62$ [51]. Consequently, the SDQ-Deu is an appropriate measure to assess effects of the therapeutic intervention of this study.

KINDL-R

The KINDL-R questionnaires assess health-related quality of life (HRQoL) via self-report and parent rating on six scales: physical and mental well-being, self-esteem, family, friends, and functional capability in daily life at school. In this study, the Kid-KINDL-R for children aged 7–13, as well as the Kiddo-Kindl-R for parents of children aged 7–13, are applied. Psychometric quality and overall consistency of the parent questionnaire are good, with Cronbach's $\alpha = .85$ for the total scales and values ranging between $\alpha = .63$ and $\alpha = .71$ for the subscales [52]. Likewise, psychometric quality of the self-report questionnaire for children is good, with Cronbach's $\alpha = .82$ for the total scales and values between $\alpha = .54$ and $\alpha = .73$ for the subscales. Hence, the KINDL-R questionnaires constitute a suitable instrument to measure HRQoL in this study.

Brief Self-Control Scale (SCS-K-D)

The German brief version of the Self-Control Scale (SCS-K-D) assesses self-control using 13 items [53]. The SCS-K-D in the version presented by Rauch and colleagues [53] assesses self-control via parent report. With a retest reliability of $r = .82$, the psychometric quality is good. Adding to the parent report, we adapted the questionnaire to a self-report measure that can be used with children. Piloting the adapted version of the self-report questionnaire for children, we confirmed its psychometric quality, as internal consistency was high, with Cronbach's $\alpha = .80$. Consequently, the SCS-K-D is a suitable instrument to assess self-control capacity in the study sample.

Questionnaire on academic self-efficacy

Academic self-efficacy is a concept describing expectations about competences that will be exhibited when confronted with academic demands. These expectations are often described from the students' own perspectives. The self-report used in this study questionnaire consists of seven items, and the internal consistency varies between Cronbach's $\alpha = .70$ and .73 due to different measurements [54]. We reworded the items and piloted them in 34 children aged 8–10. Internal consistency of the adapted scale was similar to the original scale with Cronbach's $\alpha = .71$. Although psychometric quality is only acceptable, this measure is regularly and successfully used to assess self-efficacy in children (see, e.g. [55]).

Behaviour Rating Inventory of Executive Function (BRIEF)

The Behaviour Rating Inventory of Executive Function (BRIEF) is a set of questionnaires that assess executive functions of children aged 6–16 (parent and teacherreports) and 11–16 (self-report) [56]. For this study, only parent and teacher reports are applied. These questionnaires contain 86 items that load on eight subscales of two main indices. The index 'behaviour regulation' subsumes the subscales *inhibition, shifting,* and *emotional control*. The index 'cognitive regulation' comprises the subscales *initiate, working memory, plan/organize, organization of materials,* and *task-monitoring*. The internal consistency of the teacher and parent questionnaires is very good, with values between $\alpha = .79$ and $\alpha = .98$ [56]. The retest reliability for the parent questionnaires is adequate, with values higher than $r = .80$ for most of the scales. The values are more than $r = .90$ for the teacher questionnaires.

Parents' expectations about the training's effects

Parents' expectations about the training's effects are assessed using the scale *expectations of changes* of the Fragebogen zur Erfassung relevanter Therapiebedigungen (FERT), a questionnaire that assesses relevant therapy conditions [57], in an adapted format. The scale consists of eight items and has been adapted from reporting about a person's own experiences of his or her therapy to reporting about an intervention that is experienced by the child of the reporting person. The factorial reliability of the original scale was ρc = .94 [57].

Neuropsychological tests assessing executive functions (EFs), general cognitive abilities, verbal fluency, and sustained attention

Stop-Signal Task We use the Stop-Signal Task by Verbruggen, Logan, and Stevens [58] to assess response inhibition. Participants are instructed to react as fast as possible to a primary stimulus in this paradigm. However, a stop signal occurs as a secondary stimulus in 25% of the trials, indicating that the reaction should be inhibited. If the reaction is inhibited correctly, the time between the presentation of primary and secondary stimulus is delayed by 50 ms for the next trial. If the reaction is not inhibited, the presentation of the secondary stimulus decreases by 50 ms. The range of delay between presentation of primary and secondary stimulus is 150–550 ms. The Stop-Signal Task has been shown to reliably measure response inhibition in children with ADHD [59].

Corsi Block Tapping Task The Corsi Block Tapping Task [60] is used in a computerized version from PEBL [61, 62] in both its forward and backward versions to assess visuo-spatial working memory capacity. Participants are asked to remember a series of locations that are presented on a computer screen. At the beginning of each trial, the participant sees nine blue blocks on the screen. Then one block after another lights up in yellow for 1000 ms until the sequence length is reached. Starting with a sequence length of 2, the task consists of two trials with the same sequence length presented to the participant. If at least one sequence of the two is replicated correctly by clicking on the blocks on the screen with a mouse, the sequence length increases by 1 for the next block. In the backward task, the subject must click the blocks in the reverse order of presentation. If both tasks are not replicated correctly, the test ends. Interstimulus intervals (ISIs) and intertrial intervals are set to 1000 ms. Data on the psychometric quality of the test are available for a version using three items for each sequence length. The reliability of this version is high, with $r = .95$ [63].

Digit span task (WISC-IV) The digit span task from the Wechsler Intelligence Scale for Children (fourth edition) (WISC-IV) [64], in both its forward and backward versions, is used to assess verbal working memory. Reliability of the digit span task is reported to be $r = .76$ for the backward and $r = .84$ for the forward version [64]. Hence, the digit span task from the WISC-IV is an appropriate instrument to measure verbal working memory in the study sample.

Verbal fluency task (VFT) The verbal fluency task (VFT) used in this study was developed in the research group Psychophysiology and Optical Imaging at the Department of Psychiatry and Psychotherapy of the University of Tübingen and is based on the Regensburger Wortflüssigkeits-Test (RWT) [65]. Data from a NIRS measurement are recorded while the participant completes this task to assess differences in cortical brain activation resulting from the therapy. The VFT assesses semantic and phonetic fluency as well as semantic memory and consists of three blocks with three different tasks in every block. Every task is 30 s long and is followed by a resting phase of 30 s. In the first task (phonetic fluency), participants are instructed to name nouns beginning with a given letter. They are instructed not to name proper names and they are not allowed to name a series of compound words in which one of two words always remains the same such as in bird bone, bird bath, bird call, for instance. One of the following sets of letters is randomly assigned to each of the three measurements: E, P, G, A, F, M, and K, H, R. Furthermore, the sequence of the letters is randomized to prevent sequence effects. The difficulty of finding nouns beginning with a specific letter is balanced between the groups. For the second task (semantic memory), that is, the control task, participants are instructed to name the days of the week starting with Monday, and to name approximately one day per second. In the third task (semantic fluency), participants are instructed to name nouns belonging to a given category. To each measurement, one group of categories, either 'animals, professions, drinks', 'colours, clothes, hobbies', or 'fruits, sports, toys', is assigned randomly. The sequence of the categories is randomized, and the difficulty of the categories is balanced between the sets of words. On the behavioural level, reproducibility of the VFT is good, with $r = .70$ within a 3-week time interval [66]. Reproducibility of brain activity as measured with fNIRS was acceptable, with $r = .50$ at a single subject level [66]. Hence, the VFT, as it is used in this study, can be expected to be an appropriate instrument to measure semantic and phonetic fluency as well as semantic memory and corresponding task-related brain activity in the study sample.

n-back task The *n*-back task used in this study was developed in the research group Psychophysiology and Optical Imaging at the Department of Psychiatry and Psychotherapy of the University of Tübingen. The task assesses working memory and consists of three different conditions: a 0-back, a 1-back, and a 2-back task. The tasks are presented to the participants in nine blocks, i.e. three blocks per condition containing 15 trials each, with a 20-s resting phase between active blocks. The stimulus duration is set to 300 ms, and the ISI to 1700 ms. In the 0-back task, participants are instructed to press the space bar as quickly as possible whenever they see a certain stimulus. In the 1-back task, they should respond with the space bar when any stimulus appears twice in a row. In the 2-back task, participants are instructed to press the space bar as quickly as possible whenever the current stimulus and the second last are the same. Target stimuli always constitute 4 out of the 15 presented stimuli in each block, and blocks are presented in a randomized order. In order to construct an age-appropriate version of the *n*-back task, stimuli are in image form, i.e. a moon, ball, or house. Before the actual test begins, participants practice every condition. The conditions used for testing contain different symbols than the ones used in the actual test. Data from a NIRS and an EEG measurement are recorded while the participant accomplishes this task to assess differences in cortical brain activation resulting from the therapy.

Go/NoGo task The Go/NoGo task used in this study was adapted to pictorial form from a version developed by the research group of Psychophysiology and Optical Imaging at the Department of Psychiatry and Psychotherapy of the University of Tübingen. The task assesses response inhibition and consists of eight blocks with 16 trials each. The ISI is fixed to 1150 ms, and all stimuli are presented for 350 ms. Four of the eight blocks consist of go-trials only; hence, participants are instructed to press the space bar as quickly as possible whenever they see a stimulus, i.e. randomly one of three different pictures, on the computer screen. The other four blocks are designed with 12 go- as well as 4 no-go-trials. Participants are instructed to press the space bar as quickly as possible whenever they see a go-stimulus, but to inhibit the reaction when a no-go-stimulus, i.e. a fork, appears on the screen. Blocks with only go-trials and blocks consisting of mixed trials follow each other in an alternating order, separated by a resting block of 30 s. Data from a NIRS and an EEG measurement are recorded while the participant accomplishes this task to assess differences in cortical brain activation resulting from the therapy.

Matrix span task (WISC-IV) The matrix span task, taken from the WISC-IV [64], assesses general cognitive abilities. The reliability of the matrix span task is reported to be $r = .89$ [64]; hence, it is an appropriate instrument to measure general cognitive abilities in the study sample.

Sustained attention The Conner's Continuous Performance Test (CPT) from PEBL [61, 67] is used to assess sustained attention and response inhibition. This test presents 360 letters with a size of one inch to the participant one at a time on a computer screen. The letters are presented in 18 blocks with 20 letters each, and the blocks follow each other consecutively. The duration of the presentation of a letter is approximately 250 ms, while the ISI varies between 1.0, 2.0, and 4.0 s. Within every triplet of blocks, the length of the ISIs is randomly distributed. Participants are instructed to always press the space bar as quickly as possible as soon as a letter appears. However, when the letter X appears, the space bar must not be pressed. The relative occurrence of an X, which remains constant across all blocks and triplets, is fixed at 10%; hence, in 90% of all letters presented, it is any letter but an X. The test-retest reliability of the Conner's CPT is good, with values ranging between $r = .55$ and $r = .84$ [67]. Consequently, the Conner's CPT is an appropriate test to measure sustained attention as well as response inhibition in the study sample.

Academic performance

Mathematics The Lernverlaufsdiagnostik Mathematik für zweite bis vierte Klassen (LVD-M 2–4) assesses math performance in German primary schoolchildren from Grades 2–4 [68]. Every participant receives a math test consisting of 24 tasks randomly selected at every measurement. Reliability has been estimated and ranges between $r = .79$ and $r = .92$ [68]. In correlation analyses with other German math tests such as the DEMAT [69–71], validity has been demonstrated. Hence, this test can reliably assess math performance in the study sample.

Reading and writing The Lese- und Rechtschreibtest (SLRT-II), an advanced version of the Salzburger Lese- und Rechtschreibtest (SLRT), is used to assess reading and writing skills in schoolchildren from Grades 1 to 5 (1–6 for the subtest for reading). Two parallel versions are available. The reliability coefficients for the parallel tests for reading skills range between $r = .90$ and $r = .98$. For the tests of writing skills, the interrater reliability is very high, with $r = .998$. The test-retest reliability for the writing test is between $r = .80$ and $r = .97$ with the second measurement taken 5 weeks after the first. Parallel test reliabilities range between $r = .69$ and $r = .85$ for Grades 1–4. Hence, both tests show good quality criteria and can be applied in this study.

Neurophysiological and other physiological measures

Electroencephalogram (EEG) EEG data are collected using 22 EEG channels positioned according to the international 10–20 system. Two channels of the actiCap system (Brain Products GmbH, Germany) are used to detect horizontal eye movements and are attached 1.5 cm lateral to the outer canthus of both eyes. One additional electrode is used to detect vertical eye movement and is attached 1.5 cm below the middle of the right lower eyelid.

Near-infrared spectroscopy (NIRS) NIRS is an optical imaging technique examining the blood oxygenation level-dependent (BOLD) response of brain tissue. Light from the near-infrared spectrum (700–1000 nm wavelength) can penetrate the skull and is mainly absorbed by the two chromophores oxygenated haemoglobin (oxy-Hb) and deoxygenated haemoglobin (deoxy-Hb). As the two chromophores differ in their absorption maxima, variations of the concentration of both types in the brain tissue can be derived [72]. Due to neurovascular coupling, changes in concentration of oxy- and deoxy-Hb occur in response to cortical activation [72–74]. Hence, oxy- and deoxy-Hb provide information about brain activity in respective areas [72–74]. In the present study, data are acquired with the ETG-4000 Optical Topography System (Hitachi Medical Co., Japan), which is a continuous wave system working with two different wavelengths (695 ± 20 and 830 ± 20 nm) and a temporal resolution of 10 Hz, using a 44-channel array. Relative changes of absorbed near-infrared light are transformed into concentration changes of oxy-Hb and deoxy-Hb by means of a modified Beer-Lambert law.

The 28 NIRS optodes (14 light sources (emitters), 14 detectors) are arranged in a combined NIRS/EEG cap designed to function with the Oculus Rift HMD Development Kit 2. The caps are individually localized by the EEG channels FCz and Cz according to the 10–20 system [75]. In order to assign NIRS channels that are situated in between adjacent pairs of emitters and detectors to their corresponding cortical regions, a spatial registration method of NIRS channels is applied [76]. In order to normalize the combined EEG/NIRS caps for children aged 6–10, we used the neuronavigation data of a 9-year-old girl normalized with the average brain from this age range, taken from the Template-O-Matic project [77]. A cap for a combined EEG/NIRS measurement from a previous study was placed on the girl's head. Using neuronavigation [78], optode and channel positions together with their corresponding cortical projection points on the head were obtained. The resulting coordinates from the neuronavigation were transferred to the standard Montreal Neurological Institute (MNI) space. Mapped on a virtual brain template, caps for combined EEG/NIRS measurements were customized with maximum coverage of the bilateral dlPFC (Brodmann areas 9, 46) that are used as feedback channels in the NIRS-based NFT of this study. See Fig. 3 for the alignment of the NIRS channels on the cortex surface. Hence, seven emitter and seven detector optodes are spread over prefrontal, central, temporal, and parietal areas of each hemisphere. The emitter-detector distance is 3 cm; we also employ one temporal channel on each hemisphere with a short-optode distance of 1 cm that can be used in later analysis for artefact removal (muscle artefacts as related to biting, for instance, as well as skin perfusion artefacts or other extra-cerebral signal components).

Electromyogram (EMG) EMG data are collected using the BrainAmp EEG system by Brain Products. Two electrodes placed on the surface of the skin, bilaterally on the subjects' supraspinatus muscles, measure the ratio of muscle tension between the right and left supraspinatus muscle. Reference electrodes placed on both mastoids complete the setup for the measurement. The value calculated by subtracting the normalized muscle tension of the left supraspinatus muscle from the normalized muscle tension of the right provides the feedback. Higher tension on the right will be equated to 'activation', higher tension on the left to 'deactivation'.

Accelerometer In this study, accelerometers are used to objectively measure motor hyperactivity at the non-dominant wrist, hip, and ankle at pre-, post-, and follow-up measurement as well as during every training session. The accelerometer used in this study is the wGT3X+ by the company ActiGraph. This device measures acceleration on the vertical, horizontal, and perpendicular axes with a range of −6 to +6 g (g = gravitational force). This small and very light sensor (5.6 cm × 3.3 cm × 1.5 cm; 19 g) is fixed to the waistband using a light belt or a clip. Furthermore, the ECGMove 3 (see below for a description of the device) measures acceleration of the torso.

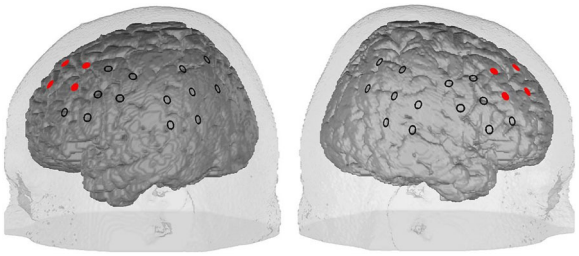


Fig. 3 Alignment of the NIRS channels on the cortex surface. The eight channels from which the feedback signal is computed are marked in red

Heart rate variability Data from the electrocardiogram (ECG) are collected during every training session using the very light sensor ECGMove 3 (from Movisens) with a size of 62.3 mm × 38.6 mm × 11.5 mm. The sensor is fixed with two Ag/AgCl cup electrodes to the skin below the sternum. Heart rate variability can then be calculated from the ECG data.

Motivation

Before the start of every training session, the participants are asked to complete a questionnaire consisting of six items that assess motivation. The questionnaire was developed at the LEAD Graduate School & Research Network of the University of Tübingen. Motivation for the training session is operationalized in four dimensions: effort (i.e. "I will make an effort to do well in the training today"), joy (i.e. "I am looking forward to today's training session", "I only came to training because I had to"), value attributed to the training session (i.e. "I am convinced that this training session is important for me"), and importance of showing a good training performance (i.e. "It is important for me to show a good training performance", "I am disappointed when I do not succeed in switching the lighting in the classroom on and off").

Time points of assessments

The first assessment (t_{-1}) takes place in order to check for all relevant inclusion and exclusion criteria. Zero to two weeks before the first training session, two baseline measurements of all relevant variables take place (t_0). Medication washout is required for one of the two test sessions in which neuropsychological and neurophysiological measurements are applied. For an overview of the variables assessed under medication washout, see Fig. 4. After eight training sessions, ADHD symptoms are assessed in a midway test (t_1). Zero to two weeks after the last training session, a post-test measurement of all relevant variables, again under medication washout for one of the two sessions, takes place (t_2). Six months after the last training session, all relevant variables are again assessed in a follow-up test (t_3).

For an overview of the tests, questionnaires, and methods employed at different time points in the study, see Fig. 4, which was designed in accordance with the standard protocol items for clinical trials [39, 40].

BRIEF Questionnaire assessing executive functions, *C3-P* Conners 3 parent questionnaire (long form), *C3-T* Conners 3 teacher questionnaire (long form), *CFT* Culture Fair Test, *CPT* Conner's Continuous Performance Test, *FERT* Questionnaire to assess relevant therapy circumstances, *HRV* heart rate variability, *KINDL-R* Questionnaire to assess HRQoL, *LVD-M 2–4* Curriculum-based assessment of mathematics skills for Grades 2–4, *SCS-K-D* Brief Self-Control Scale, *SDQ-Deu* Strengths and Difficulties Questionnaire, German version, *SLRT-II* Comprehensive assessment of reading and writing skills of children, *VFT* Verbal fluency task

Primary outcome measures

Mean group scores of every condition will be calculated for all primary outcome measures. The C3-P and C3-T [42] assess ADHD symptoms as rated by parents and teachers at pre- (t_0), midway (t_1), post- (t_2), and follow-up test (t_3), hence allowing us to assess changes within and between conditions from pre- to midway, from pre- to post-, and from pre- to follow-up test. Furthermore, at pre- (t_0), post- (t_2), and follow-up test (t_3), brain activity, i.e. the mean levels of oxy-Hb and deoxy-Hb at various channels across different cortical areas, is assessed using fNIRS data as well as EEG data from the *n*-back task assessing working memory, the Go/NoGo task assessing response inhibition, and the VFT assessing general brain activity. Consequently, changes within and between conditions in brain activity from pre- to post-, and from pre- to follow-up test can be assessed. In addition, behavioural data (i.e. mean reaction times, mean reaction time variability (*SD*), and the mean total number of commission and omission errors) are obtained from the *n*-back task and the Go/NoGo task at pre- (t_0), post- (t_2), and follow-up test (t_3), hence allowing us to assess changes from pre- to post-, and from pre- to follow-up test within and between conditions.

Secondary outcome measures

Mean group scores of every condition will be calculated for all secondary measures. Secondary outcome measures assess diverse constructs at pre-test (t_0), post-test (t_2), and follow-up test (t_3), hence allowing us to compare changes within and between conditions from pre- to post-test and from pre- to follow-up test. Children's HRQoL is assessed using the KINDL-R questionnaires for parents and children. Parents' satisfaction with as well as their expectations about the intervention's effects are assessed using the FERT questionnaire [57]. Children's mathematics, reading, and writing skills are assessed using the LVD-M 2–4 [68] and the SLRT-II [79]. Children's self-control and academic self-efficacy are assessed using the SCS-K-D and a scale by Schwarzer and colleagues [54] in both a version for parents and one for their children. Executive functioning is furthermore assessed using a digit span task [64] (verbal working memory), the Corsi Block Tapping Task [62] (visuo-spatial working memory), and the BRIEF [56], a questionnaire handed out to parents and teachers. Sustained attention is assessed using the Conner's Continuous Performance Test [61, 67], and response inhibition is assessed using the Stop-Signal Task [80]. General cognitive ability is assessed using the

	STUDY PERIOD					
	Enrolment	Allocation	Post-allocation			
	Diagnostic assessment	Pre-test	Midway-Test	Post-test	Every training session	Follow-Up
TIMEPOINT	t_{-1}	t_0	t_1	t_2		t_3
CFT 1-R or 20-R	x					
Conners 3	x	x	x	x		x
SDQ-Deu	x	x		x		x
KINDL(R)	x	x		x		x
SCS-K-D		x		x		x
Academic self-efficacy		x		x		x
BRIEF		x		x		x
FERT		x		x		x
Stop Signal Task		x*		x*		x*
Corsi Block Tapping Task		x*		x*		x*
Digit Span Task		x*		x*		x*
VFT[3]		x*		x*		x*
N-back Task[3]		x*		x*		x*
Go/NoGo Task[3]		x*		x*		x*
Matrices Span Task		x*		x*		x*
Conner's CPT		x		x		x
LVD-M 2-4		x		x		x
SLRT-II		x		x		x
EEG		x*		x*		x*
NIRS		x*		x*	x[1]	x*
Accelerometer		x*		x*	x	x*
ECG/HRV					x	
Motivation for the training session[3]					x	
EMG					x[2]	

Fig. 4 SPIRIT figure presenting an overview of the tests, questionnaires, and other methods employed at different time points in the study. [1]If participants are assigned to one of the conditions receiving a NIRS-based NF training. [2]If participants are assigned to the condition receiving an EMG-based BF training. [3]These tests/questionnaires were developed in the departments of the authors of this study. *Data collection requires medication washout

matrix span task [64]. Moreover, activity data are collected with actigraphs measuring acceleration on the vertical, horizontal, and perpendicular axes with a range of −6 to +6 g (g = gravitational force). Heart rate variability, as calculated from the ECG, as well as the motivation for every training session, as assessed with a self-report questionnaire for the children, serve as secondary outcome measures.

Statistics

Calculation of the sample size The sample sizes for the two analytical approaches were calculated using G Power version 3.1.9.2. Firstly, we calculated the sample size that is required in order to yield a significant effect of treatment within conditions. We expect appropriate

effect sizes to range between those known for within and between designs, hence expecting an effect size of *ES* = .69 [26] with a predefined α of .05 and a power of at least .80. Using a one-tailed *t* test due to directed hypotheses, the study requires at least 15 subjects per group, assuming a post- versus pre-effect, or at least 27 subjects, assuming treatment versus passive waiting control group effect. Secondly, we calculated the sample size that is required for a repeated measures analysis of variance (ANOVA) with three groups and two measurement dates in order to be able to detect effects of at least small to medium effect sizes. Hence, assuming an *ES* of .35, a predefined α of .05, a power of at least .80, and a correlation of .5 between repeated measures results in a total sample size of 84, that is, 28 per group. Consequently, taking into consideration the results of our first and second analyses, we aim for 30 participants per group.

Statistical evaluation of the results For all outcome variables, we will conduct repeated measures ANOVA as well as post hoc tests. Accelerometer data will be analysed using support vector machines (Kühnhausen J, Brefeld U, Reinelt T, Gawrilow C: Using accelerometers to predict ADHD diagnoses in children, submitted) to monitor the presence of symptoms of hyperactivity. In the case that data will not be normally distributed, adequate non-parametric tests will be applied.

All data from questionnaires completed by participants, parents, and teachers who adhered to the study protocol will be included in the analyses; this also includes data from participants or informants who left the study at a certain point of time during the course of the study, i.e. after the midway test (t_1) or after the post-test (t_2). If data from (neuro-) psychological tests are missing, respective data from all following measurements will also be excluded from the analyses, as learning effects are expected due to participation in the respective assessments. Furthermore, data from each participant will be analysed in the participant's respective condition (i.e. as randomized). If data from certain items of the questionnaires are missing, we will apply appropriate procedures to deal with missing values as suggested in the manual of the respective questionnaire.

Data security and storage All data are acquired and stored using anonymous codes. Codes and corresponding real names are noted on a code list stored in a lockable cupboard that can only be accessed by staff members of the project. The code list will be destroyed after the data collection, including follow-up tests, is finished. All data collected will be deleted after ten years from their first publication. No data monitoring committee is required for this study, as this is not a multicentre study.

Discussion

We presented an innovative study design and protocol of a randomized controlled trial (RCT) with NIRS-based neurofeedback training in children with ADHD. First, this study aims to investigate the specific effects of NIRS-based NFT compared to effects of EMG-based BFT on children with ADHD. Both variants of the training are conducted in a VR classroom environment. Second, we aim to compare differential effects of NIRS-based NFT in a 2D and a VR environment. Third, this study examines effects of NIRS-based NFT and EMG-based BFT on self-control as well as on school performance of children with ADHD.

There are already promising findings providing evidence for the efficacy of NIRS-based NFT in children with ADHD in the scope of a pilot study [27]. The study presented here now aims to further examine the findings in a comprehensive design. An active control condition receiving an EMG-based BFT will serve to differentiate specific as well as unspecific effects of the interventions. In addition to strong ethical concerns and poor compliance to treatment in NFTs using sham feedback as a control condition [30, 31], a control condition receiving a sham feedback is not adequate to approach this question. NFT and BFT for ADHD treatment generally train self-regulation in different domains [22, 81] in the fashion of an operant conditioning paradigm. On the one hand, participants acquire self-regulation skills that allow control of a specific endogenous parameter, namely brain activation. On the other hand, they learn to self-regulate behavioural conditions such as being attentive, sitting still, and enduring boredom. Therefore, we expect NFT as well as BFT to yield the same degree of effects in the latter domains, while only the acquisition of self-regulation skills related to the specific endogenous parameter will yield unique effects on ADHD symptomatology. Hence, comparing the effects of a NIRS-based NFT in the VR setting and those of an EMG-based BFT in VR in this study will illustrate the proportion of specific effects as well as effects common to both interventions.

With the study design presented, we furthermore aim at examining whether an NFT in a naturalistic VR setting might yield greater effects than an NFT in a 2D setting. From a theoretical point of view, both the acquisition of self-regulation skills in the laboratory and their transfer to everyday life situations (e.g. a classroom setting) might be facilitated by training in a naturalistic VR environment [34]. The VR environment elicits psychological and behavioural responses that would similarly occur in real life [34]. As these responses occur within a therapeutic setting, they provide the starting point for behavioural and psychological interventions [34]. Transfer of skills acquired in the training is furthermore facilitated due to the high degree of realism of the

training setting [35]. In addition, training motivation has been identified as an important moderator of the efficacy of cognitive training, as it fosters cognitive control and prefrontal activity [35]. Training motivation may be increased by a personalized context that links the goals of the training to everyday life [35]. Hence, as a naturalistic VR environment, such as a VR classroom, links goals of the training to a real-life situation, we should expect NFT and BFT taking place in a naturalistic VR environment to yield larger effects than training in 2D. The present study will investigate whether effects of a NIRS-based NFT are larger when the training is conducted in a naturalistic VR environment compared to a 2D setting.

Children with ADHD experience poor school performance [6–8] as well as a core deficit in self-control [3], that is, "the deliberate, conscious, effortful subset of self-regulation" [82]. However, aspects often neglected in prior studies include the effects of NFT and BFT on self-control and school performance. Depicting NFT and BFT as interventions that train the exertion of self-regulation in two domains, namely self-regulating an endogenous parameter and self-regulation of behaviour, it seems plausible to expect effects on and to assess self-control while investigating effects on the self-regulation of brain activity and behaviour. Furthermore, as poor school performance is related to difficulties in behaviour, EF, and self-control [8], and both NFT and BFT have been shown to improve behaviour and EF (e.g. [83, 84]), we should expect all three interventions administered in this study to improve school performance. Hence, it is vital to assess the effects of a NIRS-based NFT and an EMG-based BFT on school performance.

The present study is limited in that participants, parents, and the trainers administering the intervention are not blinded — even though an official debriefing of parents and children as well as communication of individual results will only take place after finishing the study, hence, after the follow-up test is completed. First, participants, parents, and trainers are not blinded due to time constraints that make it impossible to use both the NIRS machine and the EMG equipment simultaneously. Second, participants either wear or do not wear the HMD, and hence they will know whether it is the 2D or VR condition they belong to. Third, as measurements during training sessions require constant observation of the data being recorded, it is impossible to blind trainers for the kind of intervention administered. However, participants, parents, and trainers are informed that every participant receives a potent intervention. Hence, we hope that bias is reduced to a minimum.

We have presented the design and protocol for a randomized controlled trial on a NIRS-based NFT in a VR classroom for children with ADHD. In addition to assessing the effects of an NFT using this relatively new technology, and besides the fact that this is, to our knowledge, the first study examining differential effects of an NFT in children with ADHD in a 2D and a VR setting, we add the assessment of concepts that have rarely been considered in prior NFT studies to established measures.

Trial status

The trial is ongoing.

Additional file

Additional file 1: SPIRIT checklist. (PDF 480 kb)

Abbreviations

ADHD: Attention-deficit/hyperactivity disorder; BFT: Biofeedback training; BRIEF: Behaviour Rating Inventory of Executive Function; C3-P: Conners 3 parent questionnaire (long form); C3-T: Conners 3 teacher questionnaire (long form); CFT: Culture Fair Test; CPT: Conner's Continuous Performance Test; deoxy-Hb: Deoxygenated haemoglobin; dlPFC: Dorsolateral prefrontal cortex; DSM-5: Diagnostic and Statistical Manual of Mental Disorders, fifth edition; EEG: Electroencephalogram; EF: Executive function; EMG: Electromyogram; FERT: Questionnaire to assess relevant therapy circumstances; fMRI: Functional magnet resonance imaging; fNIRS: Functional near-infrared spectroscopy; HMD: Head-mounted display; HRQoL: Health-related quality of life; HRV: Heart rate variability; ISI: Interstimulus interval; KINDL-R: Questionnaire to assess HRQoL; LVD-M 2–4: Curriculum-based assessment of mathematics skills for Grades 2–4; MNI: Montreal Neurological Institute; NFT: Neurofeedback training; NIRS: Near-infrared spectroscopy; oxy-Hb: Oxygenated haemoglobin; RTC: Randomized controlled trial; SCP: Slow cortical potential; SCS-K-D: Brief Self-Control Scale; SD: Standard deviation; SDQ-Deu: Strengths and Difficulties Questionnaire, German version; SLRT-II: Lese- und Rechtschreibtest, an advanced version of the Salzburger Lese- und Rechtschreibtest (SLRT); VFT: Verbal fluency task; VR: Virtual reality

Acknowledgements

We acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Tübingen.

Funding

This research was funded by the LEAD Graduate School & Research Network [GSC1028], a project of the Excellence Initiative of the German federal and state governments.

Availability of data and materials

Data will be made available upon request.

Authors' contributions

FB acquired funding for this research project via the Intramural Research Funds of the LEAD Graduate School & Research Network and, together with all other authors, designed the study as well as the study protocol. FB furthermore drafted the manuscript with the help and critical revision of all other authors. JH acquired funding for this research project via the Intramural Research Funds of the LEAD Graduate School & Research Network and, together with all other authors, designed the study as well as the study protocol. TD acquired funding for this research project via the Intramural Research Funds of the LEAD Graduate School & Research Network and, together with all other authors, designed the study as well as the study protocol. ACE acquired funding for this research project via the Intramural Research Funds of the LEAD Graduate School & Research Network and, together with all other authors, designed the study as well as the study protocol. JK, together with all other authors, designed the study as well as the study protocol. TJR, together with all other authors, designed the study as well as the study protocol. CG acquired funding for this research project via the Intramural Research Funds of the LEAD Graduate School & Research Network and, together with all other authors, designed the study as well as the study protocol. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
The study was approved by the local ethics committee of the faculty of medicine at the University of Tübingen (093/2015BO1) and was written in accordance with the latest version of the Declaration of Helsinki. Before starting with the diagnostic procedure to check inclusion and exclusion criteria, primary caretakers are invited for an informed consent discussion in which the goals of the study, the study design, possible negative side effects, and expected benefits are explained. The study is also explained to the participants. Participants and parents are informed that they are free to withdraw consent to participation and use of their data at any time, and that a withdrawal will not imply any negative consequences. Participants are only invited to the diagnostic procedure and included in the study if written informed consent of both participants and primary caretakers is given.

Author details
[1]LEAD Graduate School & Research Network, University of Tübingen, Gartenstrasse 29, 72074 Tübingen, Germany. [2]Department of Psychiatry and Psychotherapy, University of Tübingen, Calwerstrasse 14, 72076 Tübingen, Germany. [3]Department of Psychology, University of Tübingen, Schleichstrasse 4, 72076 Tübingen, Germany. [4]Department of Child and Adolescence Psychiatry and Psychotherapy, University of Tübingen, Osianderstrasse 14-16, 72076 Tübingen, Germany. [5]Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Schlossstrasse 29, 60486 Frankfurt/Main, Germany.

Received: 30 June 2016 Accepted: 21 December 2016
Published online: 24 January 2017

References

1. American Psychiatric Association. DSM-5 Attention Deficit/Hyperactivity Disorder Fact Sheet. Arlington: American Psychiatric Association; 2013.
2. Barkley RA. Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. Psychol Bull. 1997;121:65–94.
3. Barkley RA. ADHD and the nature of self-control. New York: Guildford Press; 1997.
4. Sonuga-Barke EJS. Psychological heterogeneity in AD/HD—a dual pathway model of behaviour and cognition. Behav Brain Res. 2002;130:29–36.
5. Faraone SV, Asherson P, Banaschewski T, Biederman J, Buitelaar JK, Ramos-Quiroga JA, Rohde LA, Sonuga-Barke EJS, Tannock R, Franke B. Attention-deficit/hyperactivity disorder. Nat Rev. 2015;1:1–23.
6. Daley D, Birchwood J. ADHD and academic performance: why does ADHD impact on academic performance and what can be done to support ADHD children in the classroom? Child Care Health Dev. 2010;36:455–64.
7. Loe IM, Feldman HM. Academic and educational outcomes of children with ADHD. J Pediatr Psychol. 2007;32:643–54.
8. Frazier TW, Youngstrom EA, Glutting JJ, Watkins MW. ADHD and achievement: meta-analysis of the child, adolescent, and adult literatures and a concomitant study with college students. J Learn Disabil. 2007;40:49–65.
9. Rodriguez A, Jarvelin M-R, Obel C, Taanila A, Miettunen J, Moilanen I, Henriksen T, Pietilainen K, Ebeling H, Kotimaa A, Linnet K, Olsen J. Do inattention and hyperactivity symptoms equal scholastic impairment? Evidence from three European cohorts. BMC Public Health. 2007;7:327.
10. Miller AC, Keenan JM, Betjemann RS, Willcutt EG, Pennington BF, Olson RK. Reading comprehension in children with ADHD: cognitive underpinnings of the centrality deficit. J Abnorm Child Psychol. 2013;41:473–83.
11. Sciberras E, Mueller KL, Efron D, Bisset M, Anderson V, Schilpzand EJ, Jongeling B, Nicholson JM. Language problems in children with ADHD: a community-based study. Pediatrics. 2014;133:793–800.
12. Sowerby P, Seal S, Tripp G. Working memory deficits in ADHD: the contribution of age, learning/language difficulties, and task parameters. J Atten Disord. 2011;15:461–72.
13. Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF. Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. Biol Psychiatry. 2005;57:1336–46.
14. Barry RJ, Johnstone SJ, Clarke AR. A review of electrophysiology in attention-deficit/hyperactivity disorder: II. Event-related potentials. Clin Neurophysiol. 2003;114:184–98.
15. Negoro H, Sawada M, Iida J, Ota T, Tanaka S, Kishimoto T. Prefrontal dysfunction in attention-deficit/hyperactivity disorder as measured by near-infrared spectroscopy. Child Psychiatry Hum Dev. 2010;41:193–203.
16. Ehlis A-C, Bähne CG, Jacob CP, Herrmann MJ, Fallgatter AJ. Reduced lateral prefrontal activation in adult patients with attention-deficit/hyperactivity disorder (ADHD) during a working memory task: a functional near-infrared spectroscopy (fNIRS) study. J Psychiatr Res. 2008;42:1060–7.
17. Barry RJ, Clarke AR, Johnstone SJ. A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography. Clin Neurophysiol. 2003;114:171–83.
18. Hinterberger T, Veit R, Strehl U, Trevorrow T, Erb M, Kotchoubey B, Flor H, Birbaumer N. Brain areas activated in fMRI during self-regulation of slow cortical potentials (SCPs). Exp Brain Res. 2003;152:113–22.
19. Hennighausen K, Schulte-Körne G, Warnke A, Remschmidt H. Contingent negative variation (CNV) in children with hyperkinetic syndrome—an experimental study using the Continuous Performance Test (CPT). Z Kinder Jugendpsychiatr Psychother. 2000;28:239–46.
20. Chabot RJ, Serfontein G. Quantitative electroencephalographic profiles of children with attention deficit disorder. Biol Psychiatry. 1996;40:951–63.
21. Loo SK, Barkley RA. Clinical utility of EEG in attention deficit hyperactivity disorder. Appl Neuropsychol. 2005;12:64–76.
22. Gevensleben H, Moll GH, Rothenberger A, Heinrich H. Neurofeedback in attention-deficit/hyperactivity disorder — Different models, different ways of application. Front Hum Neurosci. 2014;8:1–10.
23. Drechsler R, Straub M, Doehnert M, Heinrich H, Steinhausen H-C, Brandeis D. Controlled evaluation of a neurofeedback training of slow cortical potentials in children with Attention deficit/hyperactivity disorder (ADHD). Behav Brain Funct. 2007;3:35.
24. Holtmann M, Sonuga-Barke E, Cortese S, Brandeis D. Neurofeedback for attention-deficit/hyperactivity disorder. Child Adolesc Psychiatr Clin N Am. 2014;23:789–806.
25. Cortese S, Ferrin M, Brandeis D, Holtmann M, Aggensteiner P, Daley D, Santosh P, Simonoff E, Stevenson J, Stringaris A, Sonuga-Barke EJS. Neurofeedback for attention-deficit/hyperactivity disorder: meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. J Am Acad Child Adolesc Psychiatry. 2016;55:444–55.
26. Arns M, de Ridder S, Strehl U, Breteler M, Coenen A. Efficacy of neurofeedback treatment in ADHD: the effects on inattention, impulsivity and hyperactivity: a meta-analysis. Clin EEG Neurosci. 2009; 40:180–9.
27. Marx A-M, Ehlis A-C, Furdea A, Holtmann M, Banaschewski T, Brandeis D, Rothenberger A, Gevensleben H, Freitag CM, Fuchsenberger Y, Fallgatter AJ, Strehl U. Near-infrared spectroscopy (NIRS) neurofeedback as a treatment for children with attention deficit hyperactivity disorder (ADHD) — a pilot study. Front Hum Neurosci. 2015. https://doi.org/10.3389/fnhum.2014.01038.
28. Meisel V, Servera M, Garcia-Banda G, Cardo E, Moreno I. Neurofeedback and standard pharmacological intervention in ADHD: a randomized controlled trial with six-month follow-up. Biol Psychol. 2013;94:12–21.
29. Raggi V, Chronis A. Interventions to address the academic impairment of children and adolescents with ADHD. Clin Child Fam Psychol Rev. 2006;9: 85–111.
30. Holtmann M, Pniewski B, Wachtlin D, Wörz S, Strehl U. Neurofeedback in children with attention-deficit/hyperactivity disorder (ADHD)—a controlled multicenter study of a non-pharmacological treatment approach. BMC Pediatr. 2014;14:202.
31. Birbaumer N, Elbert T, Brigitte R, Daum I, Wolf P, Canavan A. Clinical-psychological treatment of epileptic seizures: a controlled study, Perspectives and promises in clinical psychology. New York: Plenum Press; 1991. p. 81–96.
32. Albrecht B, Uebel-von Sandersleben H, Gevensleben H, Rothenberger A. Pathophysiology of ADHD and associated problems — starting points for neurofeedback interventions? Front Hum Neurosci. 2015;9:359.
33. Gevensleben H, Rothenberger A, Moll GH, Heinrich H. Neurofeedback in children with ADHD: validation and challenges. Expert Rev Neurother. 2012;12:447–60.

34. Bohil CJ, Alicea B, Biocca FA. Virtual reality in neuroscience research and therapy. Nat Rev Neurosci. 2011;12:752–62.
35. Keshavan MS, Vinogradov S, Rumsey J, Sherrill J, Wagner A. Cognitive training in mental disorders: update and future directions. Am J Psychiatry. 2014;171:510–22.
36. Käthner I, Ruf CA, Pasqualotto E, Braun C, Birbaumer N, Halder S. A portable auditory P300 brain-computer interface with directional cues. Clin Neurophysiol. 2013;124:327–38.
37. Kouneiher F, Charron S, Koechlin E. Motivation and cognitive control in the human prefrontal cortex. Nat Neurosci. 2009;12:939–45.
38. Subramaniam K, Kounios J, Parrish TB, Jung-Beeman M. Brain mechanism for facilitation of insight by positive affect. J Cogn Neurosci. 2009;21:415–32.
39. Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, Hróbjartsson A, Mann H, Dickersin K, Berlin JA, Doré CJ, Parulekar WR, Summerskill WSM, Groves T, Schulz KF, Sox HC, Rockhold FW, Rennie D, Moher D. SPIRIT 2013 Statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;2013(158):200–7.
40. Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, Dickersin K, Hróbjartsson A, Schulz KF, Parulekar WR, Krleža-Jerić K, Laupacis A, Moher D. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ. 2013;346.
41. Falkai P, Wittchen H-U. Diagnostisches und Statistisches Manual Psychischer Störungen DSM-5. 1st ed. Göttingen: Hogrefe; 2014.
42. Lidzba K, Christiansen H, Drechsler R. Conners Skalen zu Aufmerksamkeit und Verhalten-3. Deutschsprachige Adaption der Conners 3rd Edition von Keith Conners. Bern: Huber; 2013.
43. Goodman R. The Strengths and Difficulties Questionnaire. A research note. J Child Psychol Psychiatry. 1997;38:581–6.
44. Strehl U. Lerntheoretische Grundlagen und Überlegungen zum Neurofeedback. In: Strehl U, editor. Neurofeedback. 1st ed. Stuttgart: Verlag W. Kohlhammer; 2013. p. 13–30.
45. Strehl U, Leins U, Goth G, Klinger C, Hinterberger T, Birbaumer N. Self-regulation of slow cortical potentials: a new treatment for children with attention-deficit/hyperactivity disorder. Pediatrics. 2006;118:e1530–40.
46. Villringer A, Planck J, Hock C, Schleinkofer L, Dirnagl U. Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults. Neurosci Lett. 1993;154:101–4.
47. Inman VT, Ralston HJ, De CM Saunders JB, Bertram Feinstein MB, Wright EW. Relation of human electromyogram to muscular tension. Electroencephalogr Clin Neurophysiol. 1952;4:187–94.
48. Weiß RH, Osterland J. CFT 1-R. Grundintelligenztest Skala 1 - Revision. Göttingen: Hogrefe Verlag; 2013.
49. Weiß RH, Weiß B. CFT 20-R. Grundintelligenztest Skala 2 - Revision. Göttingen: Hogrefe Verlag; 2006.
50. Woerner W, Becker A, Friedrich C, Rothenberger A, Klasen H, Goodman R. Normierung und Evaluation der deutschen Elternversion des Strengths and Difficulties Questionnaire (SDQ): Ergebnisse einer repräsentativen Felderhebung. Z Kinder Jugendpsychiatr Psychother. 2002;30:105–12.
51. Goodman R. Psychometric properties of the Strengths and Difficulties Questionnaire. J Am Acad Child Adolesc Psychiatry. 2001;40:1337–45.
52. Ravens-Sieberer U, Ellert U, Erhart M. Health-related quality of life of children and adolescents in Germany. Norm Data from the German Health Interview and Examination Survey (KiGGS). Bundesgesundheitsblatt - Gesundheitsforsch - Gesundheitsschutz. 2007;50:810–8.
53. Rauch WA, Gawrilow C, Schermelleh-Engel K, Schmitt K. Dispositionelle Selbstkontrollkapazität bei Kindern. Diagnostica. 2014;60:61–72.
54. Schwarzer R, Jerusalem M. Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Berlin: Dokumentation der psychometrischen Verfahren im Rahmen der wissenschaftlichen Begleitung des Modellversuchs selbstwirksame Schulen. Berlin: Freie Universität Berlin; 1999.
55. Juang LP, Silbereisen RK. The relationship between adolescent academic capability beliefs, parenting and school grades. J Adolesc. 2002;25:3–18.
56. Drechsler R, Steinhausen H-C. BRIEF. Verhaltensinventar zur Beurteilung exekutiver Funktionen. Göttingen: Hogrefe Verlag; 2013.
57. Vollmann K, Hautzinger M, Strehl U. Entwicklung und Überprüfung eines Fragebogens zur Erfassung relevanter Therapiebedingungen (FERT) [dissertation]. 2009.
58. Verbruggen F, Logan GD, Stevens MA. STOP-IT. 2013. http://www.psy.vanderbilt.edu/faculty/logan/#stopit.
59. Soreni N, Crosbie J, Ickowicz A, Schachar R. Stop Signal and Conners' Continuous Performance Tasks Test—Retest reliability of two inhibition measures in ADHD children. J Atten Disord. 2009;13:137–43.
60. Corsi PM. Human memory and the medial temporal region of the brain. Diss Abstr Int. 1972;34:891.
61. Mueller ST, Piper BJ. The Psychology Experiment Building Language (PEBL) and PEBL test battery. J Neurosci Methods. 2014;222:250–9.
62. Mueller ST. The PEBL The PEBL. 2011.
63. Schellig D. Block-Tapping-Test (BTT). Manual. Frankfurt: Pearson Assessment; 1997.
64. Petermann F, Petermann U. Wechsler Intelligence Scale for Children - Fourth Edition *Deutsche Bearbeitung*. Frankfurt: Pearson; 2011.
65. Aschenbrenner S, Tucha O, Lange KW. Regensburger Wortflüssigkeits-Test (RWT). Göttingen: Hogrefe; 2001.
66. Schecklmann M, Ehlis A-C, Plichta MM, Fallgatter AJ. Functional near-infrared spectroscopy: a long-term reliable tool for measuring brain activity during verbal fluency. Neuroimage. 2008;43:147–55.
67. Conners KC, Epstein JN, Angold A, Klaric J. Continuous performance test performance in a normative epidemiological sample. J Abnorm Child Psychol. 2003;31:555–62.
68. Strathmann AM, Klauer KJ. LVD-M 2–4 Lernverlaufsdiagnostik für zweite bis vierte Klassen. Göttingen: Hogrefe; 2012.
69. Krajewski K, Küspert P, Schneider W, Visé M. DEMAT 1+. Deutscher Mathematiktest für erste Klassen. Göttingen: Beltz; 2004.
70. Krajewski K, Liehm S, Schneider W. DEMAT 2+. Deutscher Mathematiktest für zweite Klassen. Göttingen: Beltz; 2004.
71. Gölitz D, Marx H, Schneider W. DEMAT 4. Deutscher Mathematiktest für vierte Klassen. Göttingen: Beltz; 2004.
72. Obrig H, Wenzel R, Kohl M, Horst S, Wobst P, Steinbrink J, Thomas F, Villringer A. Near-infrared spectroscopy: does it function in functional activation studies of the adult brain? Int J Psychophysiol. 2000;35:125–42.
73. Alderliesten T, De Vis JB, Lemmers PMA, van Bel F, Benders MJNL, Hendrikse J, Petersen ET. Simultaneous quantitative assessment of cerebral physiology using respiratory-calibrated MRI and near-infrared spectroscopy in healthy adults. Neuroimage. 2014;85:255–63.
74. Steinbrink J, Villringer A, Kempf F, Haux D, Boden S, Obrig H. Illuminating the BOLD signal: combined fMRI–fNIRS studies. Magn Reson Imaging. 2006;24:495–505.
75. Jasper H. Report of the committee on methods of clinical examination in electroencephalography. Electroencephalogr Clin Neurophysiol. 1958;10:370–5.
76. Cutini S, Scatturin P, Zorzi M. A new method based on ICBM152 head surface for probe placement in multichannel fNIRS. Neuroimage. 2011;54:919–27.
77. Wilke M, Holland SK, Altaye M, Gaser C. Template-O-Matic: a toolbox for creating customized pediatric templates. Neuroimage. 2008;41:903–13.
78. NeuroConn. BrainsightTM. 2012.
79. Moll K, Landerl K. Lese- und Rechtschreibtest (SLRT-II). Weiterentwicklung des Salzburger Lese- und Rechtschreibtests (SLRT). 2, korrig. Bern: Verlag Hans Huber; 2014.
80. Verbruggen F, Logan GD, Stevens MA. STOP-IT: Windows executable software for the stop-signal paradigm. Behav Res Methods. 2008;40:479–83.
81. Gevensleben H, Holl B, Albrecht B, Schlamp D, Kratz O, Studer P, Wangler S, Rothernberger A, Moll GH, Heinrich H. Distinct EEG effects related to neurofeedback training in children with ADHD: a randomized controlled trial. Int J Psychophysiol. 2009;74:149–57.
82. Baumeister RF, Vohs KD, Tice DM. The strength model of self-control. Curr Dir Psychol Sci. 2007;16:351–5.
83. Bakhshayesh AR, Hänsch S, Wyschkon A, Rezai MJ, Esser G. Neurofeedback in ADHD: a single-blind randomized controlled trial. Eur Child Adolesc Psychiatry. 2011;20:481–91.
84. Arns M, Conners CK, Kraemer HC. A decade of EEG theta/beta ratio research in ADHD: a meta-analysis. J Atten Disord. 2013;17:374–83.