Aus dem Zentralinstitut für Seelische Gesundheit
Abteilung Klinische Psychologie
Leitung: Prof. Dr. Peter Kirsch

The regulation of large-scale brain networks via functional magnetic resonance imaging: Neurofeedback as treatment approach

Inauguraldissertation
zur Erlangung des Doktor scientiarum humanarum (Dr. sc. hum.)
der
Medizinischen Fakultät Mannheim
der Ruprecht-Karls-Universität
zu
Heidelberg

vorgelegt von
Franziska Weiß

aus
Moers
2022

Dekan: Prof. Dr. med. Sergij Goerdt
Referent: Prof. Dr. phil. Peter Kirsch

TABLE OF CONTENTS

Page

LIST OF ABBREVIATIONS

ACC	Anterior cingulate cortex
ADHD	Attention Deficit Hyperactivity Disorder
BOLD	Blood oxygen level dependent
fMRI	Functional magnetic resonance imaging
DCM	Dynamic causal modelling
DLPFC	Dorsolateral prefontal cortex
DMPFC	Dorsomedial prefrontal cortex
EEG	Electroencephalogram
FC	Functional connectivity
GLM	General linear model
GSR	Global signal regression
HRV	Heart rate variability
MDD	Major depressive disorder
MPFC	Medial prefrontal cortex
NeCoSchi	Network connectivity in schizophrenia
rt-fMRI NF(B)	Real-time fMRI neurofeedback
OCD	Obsessive-compulsive disorder
OFC	Orbitofrontal cortex
OSF	Open science foundation
PTSD	Posttraumatic stress disorder
PSTG	Posterior superior temporal gyrus
ROI	Region-of-interest
RVT	Respiratory volume per time
SCZ	Schizophrenia
VLPFC	Ventrolateral prefrontal cortex
VPFC	Ventral prefrontal cortex

1 INTRODUCTION

Functional magnetic resonance imaging (fMRI) measures the hemodynamic changes in the blood as a consequence of changed neural activity (Logothetis, 2008). Although severely reliant on the applied experimental protocol and statistical analyses (Bennett & Miller, 2013), it is one of the most optimal methods available for learning about brain function (Logothetis, 2008). fMRI allows for a visualization of brain activity during specific tasks or during the resting state to reveal baseline brain activity of a certain target group such as a specific patient population (Tian et al., 2008). For more than 30 years the method has been used with the aim of demonstrating typical activation patterns that occur with cognitive tasks (see for example (Booth et al., 1999; Gonzalez, (2005)) and to represent different physiological states in animals (Ogawa et al., 1990) and in humans (Haase et al., 2009).

Various fMRI studies with diverse study designs provided scientists and eventually the layperson with information about the brain on all imaginable aspects of cognitive functioning such as attention and memory tasks (Mainero et al., 2004; Sylvester et al., 2003). Along these lines, identifying structures involved in attention, it was for example found that a common cognitive process is responsible for the distribution of attention, which is governed by the superior parietal cortex (Sylvester et al., 2003).

As indicated by the previous example, fMRI allows to gain knowledge on functional aspects of the human brain that cannot be accessed via other neuroimaging means with comparable quality and it is the “main modality for clinical neuroimaging” (Mier & Mier, 2015, p. 5). In that respect, it enables encountering changed brain activation and as a matter of fact changed connectivity which hints at disruptions in the communication of brain regions (Mier & Mier, 2015).

Hereafter, a description of selected famous fMRI studies is provided to illustrate the widespread options fMRI offers in terms of examining how different mental phenomena and behavioral measures are represented in the brain.

Based on previous research it appears likely that changes in neural activity will also affect behavioral measures. A majority of studies have demonstrated the close link between the two (Bones et al., 2014; Critchley et al., 2001; Retter et al., 2021; Sugiura et al., 2000; Yoo & Jolesz, 2002; Zhao et al., 2016). As one of the most famous fMRI

studies, Eisenberger et al. (2003) investigated whether the neural basis of physical pain was similar to the neural basis of social pain. By means of a virtual ball-tossing game, the authors showed that congruent with the pattern in physical pain, activity in the right ventral prefrontal cortex (VPFC) and dorsal anterior cingulate cortex (ACC) were highly increased in a social exclusion condition. These brain activity measures were also associated with self-reported distress. The findings by Eisenberger et al. (2003) underline two very important aspects for subsequent fMRI studies: First, the association of physical measures and perceived mental phenomena and secondly, the close interplay between neural activity and behavioral measures.

As another example for the relationship between neural activity and behavioral measures, our current understanding of emotional engagement in personal situations from a brain perspective has been strongly advanced by valuable neuroimaging findings from Greene et al. (2001) who investigated the neural correlates of moral dilemmas. The authors demonstrated that moral decisions in a “up close and personal” scenario in contrast to a more intuitively impersonal scenario require a higher level of activity in brain areas known for their involvement in emotions such as medial frontal gyrus and posterior cingulate gyrus (Greene et al., 2001). This was interpreted as evidence for an enhanced emotional engagement if situations are more personal and it suggests that empathy is reduced in impersonal contexts (Greene et al., 2001).

The descriptions of the previous two studies underline that fMRI allows for performing strongly distinct and also very specific approaches and delivers formidable insights on the brain that might not be detected otherwise. These insights further can inform other related disciplines such as social psychology and can have a great impact. Similar to other neuroimaging modalities, fMRI can be seen as an interface that combines neuroscientific knowledge with insights from a variety of other disciplines and by that facilitates a more holistic understanding of specific (disturbed) processes e.g. in mental disorders.

The advent of fMRI as a technique enabled a visualization of brain activity (Bandettini et al., 1992; Heeger & Ress, 2002) and a consequent comparison of brain activity apparent in mental disorders and the pattern visible in healthy control participants which advanced the insights into the processes underlying mental disorders. Among others, this has the advantage of characterizing mental disorders based on their abnormalities in brain activity. As an example, Sheline et al. (2001) compared neural

activation to faces with neutral or fearful expression of depressive patients with those of healthy controls. It was found that depressive patients presented with hyperactivity in the left amygdala even when the processing happened outside of their awareness and when the faces were masked. After usage of antidepressant treatment, this pattern normalized. These striking findings suggest that patients with mental disorders and healthy controls can be distinguished from each other based on their neural signatures in day-to-day tasks. Furthermore, relevant hints on how pharmaceuticals enable a "reset" of the brain are provided. This information can be used to develop new treatments.

All of this very powerful knowledge and the constant progression of the method (Poldrack, 2012) has put the neuroscientific community in a state in which brain processes during the execution of a particular cognitive task can be well characterized and in which it is possible to relatively precisely predict how the brain of a person with a certain mental disorder will be affected functionally (Damaraju et al., 2014; Zhang et al., 2016). Furthermore, commonalities of activation patterns in different disorders exist and this knowledge is increasing constantly to eventually lead on to transdiagnostic treatment approaches (Yang et al., 2018). By now, after three decades of elaboration, the methodology of fMRI is sufficiently sound to finally be applied in a more treatment-related context: by means of real-time fMRI neurofeedback (rt-fMRI NF) aiming at inducing changes on the neural level that would affect behavioral measures.

In the following part an overview on general principles in rt-fMRI NF is given and insights on different factors in the practical application of it (e.g. target selection) are provided.

1.1 fMRI neurofeedback and its underlying principles

fMRI neurofeedback (NF) is an fMRI-based form of biofeedback (Munoz-Moldes & Cleeremans, 2020) that was first developed a quarter of a century ago (Cox et al., 1995). While it is often regarded as an individual field, fMRI NF builds on a longer tradition of biofeedback research, for example with electroencephalography (EEG), which was already in use for the conduct of NF at that time (Rockstroh et al., 1990). Generally, NF approaches aim at inducing changes in brain activity via volitional control by providing an otherwise non-accessible feedback signal associated with a target

brain process. rt-fMRI NF utilizes the high spatial resolution and whole brain coverage of the fMRI method (Weiskopf, 2012) in a treatment-related context. Usually, fMRI NF studies define a desirable brain state e.g. an increase in activation of a specific brain area. To achieve this desirable state and to keep it over a longer period of time, is the ultimate goal of rt-fMRI NF. To support in the process, feedback on the current brain state is represented on a screen in the MRI which is indicatable in different ways such as by means of a thermometer with an increased temperature.

In one of the first and very promising rt-fMRI NF studies, the dorsal and rostral-ventral ACC were supposed to be up-regulated in one session with different runs (Weiskopf et al., 2003). Participants did not receive instructions regarding specific strategies that would benefit the regulation, thus changes should be achieved by trial-and-error. Percent signal change as indicator of learning success was calculated across runs and significantly enhanced dorsal ACC activity was found in five of eight runs, while the rostral-ventral ACC was significantly increased in four runs (Weiskopf et al., 2003). It was concluded that the method might be an effective tool for investigation of behavioral, cognitive, and emotional function of certain brain areas (Weiskopf et al., 2003). In the following years it has been demonstrated that the application of rt-fMRI NF is successful in evoking changes on the brain level but also in the clinical picture (Kirsch et al., 2016; Ramot et al., 2017; Sukhodolsky et al., 2020; Young et al., 2014; Zilverstand et al., 2015; Zotev et al., 2011).

The potential of NF was already illustrated in animal studies that hint at self-regulation of neural activity (Schafer & Moore, 2011; Sitaram et al., 2017). In a study including monkeys, it was found that by use of auditory feedback and juice rewards, a willingly increase or decrease of neuronal activity in the frontal eye field could be evoked (Schafer & Moore, 2011). The authors suggested that these findings could pave the ground for NF training in the context of disorders of attention (Schafer & Moore, 2011).

In a nutshell, rt-fMRI NF training is a promising approach to enable changes in brain functioning. In the next section, more attention will be brought to understanding the fundamental basis of NF training.

1.1.1 Underlying theories

To facilitate and streamline NF learning, the general principles of learning should be known and integrated into NF protocols (Dudek & Dodell-Feder, 2021; Sitaram et al., 2017; Weiskopf, 2012). NF is based on the principles of operant conditioning which is characterized by an increased probability of performing a specific action if that action was rewarded previously (Martz et al., 2020; Skinner, 1938; Weiskopf et al., 2004). With regard to NF paradigms this implies that experience of the successfully achieved brain state is by nature rewarding. In support of this, an exemplary NF study has shown that regulation of target activity is in fact associated with reward value (Paret et al., 2018).

For successful implementation of NF, the presentation of rewards must be contingent with the execution of a specific behavior (Weiskopf et al., 2004). To evoke high contingency and thereby enable effective NF learning, fast and precise feedback is required which includes a minimal time delay (Weiskopf, 2012; Weiskopf et al., 2004).

Regarding this association of behavior and reward, one of the potentially highly relevant learning theories in the field of rt-fMRI NF is the “Two-Process-Theory” of the acquisition of autonomic control (Lacroix, 1981) which was originally developed for biofeedback. According to this model, two distinct processes elicit control of autonomic responses. In the first process, efferent processes are identified. The second process only proceeds if problems emerge in the first process. It includes an afferent process and is applicable when the participant is not aware of any appropriate behavior to control the target response or when they own behaviors which efficiently monitor the target response, but those cannot be recalled under the experimental conditions. In those cases, acquisition of control takes place as if a novel behavioral program is designed (Lacroix, 1981).

Of note, the aim of rt-fMRI NF should be that neuronal and behavioral changes as a consequence of rt-fMRI NF are due to the learning process of feedback from target brain regions and not due to general self-regulation with its potential therapeutic value (Tsuchiyagaito et al., 2021). In early NF studies it has been suggested that effective NF learning should only induce activity changes in the target area, but not in other brain areas (Weiskopf et al., 2004; Weiskopf et al., 2003). However, this view contradicts the current understanding of the brain as a network in which activation of a

brain structure is affected by other brain structures (Bullmore & Sporns, 2009, 2012; Power et al., 2011). The importance of brain networks is further highlighted by their clinical relevance (Seeley et al., 2009). Therefore, effective NF learning should not be characterized by disproportionate regional specificity.

Until now, research conducted on the underlying (learning-) theories of rt-fMRI NF is quite limited. It is important that this is investigated further (Sitaram et al., 2017) to improve future NF studies by factoring in knowledge on these (learning-) theories (Weiskopf, 2012). Based on elaborated learning theories, future studies could optimize their study designs and thereby allow for stronger effects of NF training. In a similar vein, moderators of NF learning will be discussed in a later part of this methods section.

Not only self-regulation of brain areas could be learned by NF, but feedback might also result in learning of manipulations of artefacts such as small movements (Weiskopf et al., 2004). Further examples for artefacts that can be learned are physiological sources (respiration, cardiac) which might be difficult to disentangle from the actual Blood oxygen level dependent (BOLD) signal. Furthermore, these manipulations might be easier to learn than the regulation of the targeted brain process itself suggesting that this could act as a trial-and-error strategy to achieve changes in the feedback signal. As opposed to some biofeedback approaches that in fact aim at training these sources, regulation of these is not desired in rt-fMRI NF. These artefacts could confound the NF training results and therefore it is important to account for those in subsequent analyses.

The following part of the dissertation is dedicated to the application of rt-fMRI NF.

1.1.2 Application with Region-of-Interest (ROI)-based approaches

Corresponding with its name, in ROI-based rt-fMRI NF feedback of a ROI is fed back to the participant. The method is characterized by its clear focus on single brain regions rather than the interaction of different brain regions and has been widely applied to research the ability to induce functional brain changes in specific brain areas of healthy controls and individuals with different mental health conditions.

1.1.2.1 Application in healthy controls

Below a summary of well-known rt-fMRI NF studies with healthy controls is provided to demonstrate the feasibility of this approach in a healthy population. It has been shown that rt-fMRI NF can be successfully applied in a variety of contexts such as emotion processing and pain perception (Emmert, Breimhorst, et al., 2017; Paret et al., 2014; Zotev et al., 2011).

rt-fMRI NF has been used to target a multitude of brain areas, but the target area that was focused in a large proportion of rt-fMRI NF studies is the amygdala which was applied in 23 of the included NF studies in a systematic review (Fede et al., 2020). In a well-known NF study Zotev et al. (2011) focused on improving emotion processing by targeting activity in this brain region. In this experiment, participants were instructed to recall positive autobiographical memories while up-regulating amygdala activity during NF training. It was found that amygdala activation significantly increased in the course of the training. Exemplary, a few years later, Paret et al. (2014) also aimed at regulating emotional dysregulation by targeting the amygdala. In this study with healthy controls amygdala activation was supposed to be down-regulated while watching aversive scenes and obtaining continuous visual feedback. It was found that successful down-regulation of the amygdala could be achieved, but this was not limited to the experimental group and instead also applied to the control group that received feedback of the basal ganglia. Per se this down-regulation of the amygdala could be seen as success of the rt-fMRI NF training, but the fact that the control group achieved the same improvements raises some issues about the specificity of the effect (Ochsner et al., 2004). Even more common NF target areas include the insula and medial prefrontal cortex / ACC / orbitofrontal cortex (OFC) which were applied 24 and 26 times respectively (Fede et al., 2020). A study by Emmert, Breimhorst, et al. (2017) addressed pain coping while experiencing heat-induced pain during the rt-fMRI NF experiment. Left anterior insula and ACC served as targets in the two NF groups and the training was performed across four runs. As a result, it was found that pain coping assessed with a Coping Strategies Questionnaire was positively associated with activation in the ACC, prefrontal cortex, hippocampus and visual cortex during NF. The results suggest that success of the training is related to individual behavioral measures (Emmert, Breimhorst, et al., 2017).

Alongside discussing the findings of individual rt-fMRI NF studies, it would have been desirable to also discuss meta-analyses on rt-fMRI NF efficacy in healthy controls. Unfortunately, excluding preprints, to date there is only one meta-analysis available which focuses on mediators of brain regulation rather than efficacy of the method per se (Emmert et al., 2016). This is a considerable gap in the literature that needs to be filled by future research.

1.1.2.2 Application in mental disorders

Next to the application in healthy controls, rt-fMRI NF training with divergent success has also been conducted in a broad variety of mental disorders such as anxiety disorders, trauma-and stressor-related disorders, mood disorders, substance use disorders, eating disorders, neurodevelopmental disorders and schizophrenia (SCZ), to mention a few. The following section aims at summarizing the evidence for NF effects in clinical populations. The order in which the clinical disorders are presented is for the most part based on their current prevalence worldwide, starting with the most prevalent one (Dattani et al., 2021).

In terms of anxiety disorders, a study by Zilverstand et al. (2015) investigated NF performance in individuals with spider phobia to address anxiety regulation. In a dual feedback design participants were instructed to down-regulate activity in the insula and up-regulate DLPFC activity during the presentation of anxiety provoking pictures. While the control group was told to learn from intuition, the experimental group should apply strategies of cognitive reappraisal. Anxiety levels were assessed as part of the NF paradigm. The study demonstrated that anxiety levels in the experimental group were reduced in comparison to the control group, although they still increased over sessions due to progressively demanding stimuli. While both groups obtained high prefrontal activity levels, only the experimental group was able to down-regulate insula activity. Down-regulation of insula activity also correlated with anxiety improvement (long-term). This evidence suggests that rt-fMRI NF is an encouraging method in patients with anxiety disorders.

Similar to anxiety disorders, convincing NF findings have also been demonstrated in the domain of trauma- and stressor-related disorders. In a feasibility study including three combat veterans with chronic posttraumatic stress disorder (PTSD) significant clinical improvements were found in two of the patients, while one improved slightly

(Gerin et al., 2016). In another study, including PTSD patients left amygdala activity was supposed to be up-regulated during the execution of a happy emotion induction task (Zotev et al., 2018). As a result of three rt-fMRI NF training sessions it was found that participants demonstrated a significant decrease on the Clinician-Administered PTSD Scale, particularly in avoidance and hyperarousal symptoms. The authors argue that veterans with combat-related PTSD could profit from this approach (Zotev et al., 2018).

Evidence from the field of mood disorders stems from NF studies including major depressive disorder (MDD) patients. In this subject group, it was found that participants were able to increase their amygdala responses while recalling positive autobiographic memory. Clinically, decreased anxiety ratings and enhanced happiness ratings were also found as a consequence of the NF training (Young et al., 2014). Mehler et al. (2018) tested whether the efficacy of rt-fMRI NF in depression is specific to feedback from emotion-regulating areas in contrast to a control region responsive to visual scenes. Although no significant group differences were detected, depressive symptoms diminished by 43% over groups. The authors propose that rt-fMRI NF training of higher visual areas might be a suitable substitute for targeted regulation as it can have therapeutic effects as well. However, the authors did not apply a waiting group which would allow such conclusions. It is proposed to conduct qualitatively high studies with a solid design to assure that the hypotheses are sufficiently testable. Based on the presented evidence, it is suggested that rt-fMRI NF is a promising therapeutic tool in depression treatment (Young et al., 2014).

Within the domain of substance use disorders, participants with high social drinking tendencies were found to be able to down-regulate ventral striatal cue-reactivity (Kirsch et al., 2016). Assessing non-treatment seeking, nicotine-dependent smokers, Hartwell et al. (2016) compared NF performance of participants receiving their own feedback values of a craving-related area in the medial prefrontal cortex (mPFC), other PFC regions and the ACC to the regulation performance of a control group which did not receive any feedback during the task in which craving-related pictures were shown. Over three scanning sessions it was found that the experimental group significantly decreased their neural activity to smoking-related cues and subjective craving ratings. However, it has to be noted that different ROIs were used on each training day, complicating comparisons of the time courses.

As another field of study, rt-fMRI NF has also been applied in obesity, which is a major global health issue (Wolfenden et al., 2019). In this subject group, Kohl et al. (2019) targeted activity of the DLPFC which plays a key role in inhibitory control of ingestive behavior. Food appraisal and snack intake were measured at different timepoints. As a result, participants were able to up-regulate activity in the target area and the control group which received feedback of the visual cortex additionally achieved an increase in left DLPFC activity while up-regulating. Actual food intake was not influenced by the training. The authors suggest the use of rt-fMRI NF as add-on to therapeutic strategies (Kohl et al., 2019).

Within the field of neurodevelopmental disorders, in children with diagnosed Attention Deficit Hyperactivity Disorder (ADHD) four sessions of rt-fMRI NF training were conducted (Alegria et al., 2017). Due to its' involvement in ADHD symptoms, the right inferior prefrontal cortex served as target in the experimental group. The active control group was presented with the signal of an unrelated brain area (left parahippocampal gyrus). As a result of the NF training, it was found that both groups presented with linear activation increases with rising number of runs in the target region and significant reduction in ADHD symptoms. However, it is questionable why the control group achieved the same changes in brain functioning as the experimental group. Promising results were also found in adolescents with Tourette syndrome (Sukhodolsky et al., 2020). These studies show that rt-fMRI NF might have potential for successful application in children and adolescents.

First NF studies have begun to target brain processes in SCZ and highly promising results were reported. A systematic review on rt-fMRI NF studies in SCZ included five studies with an average number of 7.2 sessions (Gandara et al., 2020). In an rt-fMRI NF study comparing SCZ patients to healthy controls, participants were instructed to up-regulate activity of the ACC which is important in social processing by means of a social feedback approach. Compared to baseline blocks, both SCZ and healthy control group obtained a significant increase in ACC activity during the regulation blocks (Cordes et al., 2015). Interestingly, SCZ patients up-regulated the dorsal part of the ACC, while up-regulation in healthy controls was focused on the rostral part. It is concluded that neural dysregulation in SCZ expressed in cognitive impairments can be addressed using rt-fMRI NF.

Voluntary control of the ACC in SCZ was also investigated by Dyck et al. (2016). As one of the smaller NF studies including only three patients over three training days, participants were instructed to up-regulate activity in the ACC aiming at a reduction of auditory verbal hallucinations (AVH). As a result, control over the ACC was achieved and improvements in AVH were reported (Dyck et al., 2016). Of note, the limited sample size minimizes the possibility to draw inferences on these results.

Also aiming at a reduction in AVHs, a methodologically very sound rt-fMRI NF study including patients with SCZ spectrum disorder and healthy controls employed a double-blind, randomized, crossover intervention design (Zweerings et al., 2019). Participants were instructed to up-regulate activity of the inferior frontal gyrus (IFG) and posterior superior temporal gyrus (pSTG) ROIs on one training day and down-regulate activity of these ROIs on the second day. Strategies such as recalling positive autobiographic memories were suggested to the participants. Promising significant modifications in resting state functional connectivity (FC) expressed as increased connections of the default mode network (DMN) and language areas as a consequence of rt-fMRI NF training were found (Zweerings et al., 2019). It is proposed that rt-fMRI NF approaches like this are highly valuable, not least because changes in FC are considered an important indicator for clinical measures.

Another interesting target region in SCZ is the insula. Instructed to increase activity in the anterior insula over four training sessions SCZ patients successfully up-regulated the target and this effect improved over training sessions (Ruiz et al., 2014). Of clinical importance, negative symptoms and duration of the illness negatively correlated with regulation success. Based on such findings, it is suggested that the application of rt-fMRI NF can induce behavioral and neural changes in SCZ.

A frequently asked question is how rt-fMRI NF performance could be improved. Meditation and mindfulness-based training to aid NF performance are investigated increasingly (Bauer et al., 2020; Weiss, Aslan, et al., 2020). As a methodological advancement, a recent rt-fMRI NF study with SCZ and schizoaffective disorder patients conducted meditation training before the NF training to aid NF performance. In an attempt to reduce AVH, default mode network and central executive network were targeted for modification. It was found that decreased DMN connectivity correlated with decreased AVHs and increased DMN and central executive network anticorrelations as a consequence of NF (Bauer et al., 2020). These promising results

underline the high significance of rt-fMRI NF FC application clinically and on a neural level.

Adding to the previously discussed evidence from individual clinical studies, a limited number of meta-analyses have also been conducted. In a meta-analysis of rt-fMRI NF efficacy in chronic pain patients, the authors found highly diverse results with no evident trends (Patel et al., 2020). However, it should be noted that for fMRI NF only three studies were included which themselves suffer from low sample sizes (Patel et al., 2020). A meta-analysis on rt-fMRI NF in various clinical disorders found a moderate impact on the targeted brain region(s) over training which is even higher when no NF signal is present (Dudek & Dodell-Feder, 2021). It is concluded that voluntary control over brain activity is feasible (Dudek & Dodell-Feder, 2021).

In sum, previous rt-fMRI NF studies with clinical populations suggest that the method possesses a strong potential to enable changes in brain functioning and clinical symptoms (Dudek & Dodell-Feder, 2021). Despite improvable comparability of various studies, existing results are highly promising. However, for each clinical population there is just a limited number of studies available and therefore the ability to draw meaningful conclusions on the clinical efficacy is reduced (Tursic et al., 2020). It will be a future task to gain additional insights on potential clinical effects of the method.

1.1.3 Methodological aspects

The previously described rt-fMRI NF studies have shown that the method has a great potential for use as a therapeutic treatment tool, however, the evidence for its efficacy in a specific disorder is still rather thin. Results and quality of fMRI research strongly depend on the applied design (Heunis et al., 2020; Poldrack, 2012). Therefore, it is very important to invest a significant amount of time and rigor in designing the study in advance and in pondering which design might be best suited for answering the research question at hand (Weiskopf et al., 2004). Although there is a lot of ongoing discussions regarding the optimal application of the method, this section will introduce those important topics which have been discussed most (Dudek & Dodell-Feder, 2021; Fede et al., 2020; Sorger et al., 2019).

1.1.3.1 Procedures of target selection

In rt-fMRI NF the target of regulation is of critical importance for the execution and consequent findings of the study. On a general level, in the process of designing the study, one must define the target brain area that is supposed to be regulated. The target is usually defined based on deviant neural activity within a specific mental disorder which is associated with behavioral abnormalities or clinical symptoms (Tursic et al., 2020). Importantly, the target process should be distinct and reliable for direct visualization and there should be sufficient evidence for its involvement in a given mental disorder or task. Furthermore, the potential to be estimated in real-time must exist.

Technically, the most common method to define the target is to use a predefined mask of the area in which activity has to be regulated (Tursic et al., 2020). The activity of areas within the mask is then extracted from the brain activity of each subject and a feedback value is formed, see for example Weiss et al. (2020). Alternatively, a frequently used method is to select the target based on a preceding functional localizer scan presenting a paradigm that is relevant to the disorder at hand or activates the targeted brain region (Mehler et al., 2018; Tursic et al., 2020). On that front, as an example, Hartwell et al. (2016) selected an area in the prefrontal cortex or alternatively ACC that was activated by the presentation of craving-related cues. While the authors argue that this way of target selection provides more meaningful feedback, it is questionable that a different target was selected for each visit as it reduces comparability between visits. Moreover, the comparability among participants is also reduced.

While the first method assures high comparability across participants, the second method ensures clinical relevance of the target area. Depending on the specific focus of the planned study, the optimal method for target selection is chosen i.e. studies with a strong clinical focus might rather opt for localizer approaches than predefined masks.

1.1.3.2 Definition of NF success: The dependent variable

Quantification of successful NF regulation can be conducted via different options. Broadly speaking, NF success can be based on changes in activation, changes in

symptoms or changes in task performance as a consequence of NF. Because the first two possibilities are most common, they will be outlined here.

Activation

The most logical option to define NF success is to focus on the brain activation level itself. If the instructed change in activation in a predefined direction was significantly achieved, it can be defined as successful regulation. Among many other options, the comparison can either be based on the within-subject level compared to a previous baseline or on the between-subject level e.g. by comparing brain activation of the experimental group to a control group that either received feedback of a different brain area, a different participant (yoke) or computer-generated feedback. Different control groups are further discussed in chapter 1.1.3.4. So far, most rt-fMRI NF studies have focused on changes in activation as a dependent variable (Cordes et al., 2015; Dyck et al., 2016; Young et al., 2014), because it is a clear biological marker.

Symptoms

A more pragmatic view in clinical studies is to investigate symptom changes as a consequence of NF training. For exploration of symptom changes it is recommended to include diagnosed patients instead of healthy controls as participants, as a certain degree of impairment before the start of the training is required to be able to detect changes afterwards. Symptom improvements as a consequence of NF training were e.g. found in depression (Mehler et al., 2018) and PTSD (Zotev et al., 2018). However, rare evidence exists that symptom improvements after rtfMRI NF can also be found in subclinical populations (Scheinost et al., 2013). Whereby symptom improvements can be seen as distinct and solid indications that the training was effective, they are accompanied by the shortcoming that they can hardly be uncoupled from the influence of other sources. Accordingly, it is possible that the achieved effect can be ascribed to the additional attention patients received instead of the actual NF training. This has been shown by the findings of deCharms et al. (2005) in which both experimental and control NF group reported an improvement in symptoms as a result of the training. For that reason, it seems preferable to first demonstrate changes in brain functioning and then as a secondary outcome, show that these changes are correlated with changes in symptoms. By this, the unique contribution of NF itself can be illustrated better and could not be misclassified for having a moderating function.

1.1.3.3 Feedback presentation

Feedback presentation has a strong effect on the outcome of an NF study. A variety of options exists on how regulation performance of brain activity can be fed back in rt-fMRI NF. In most rt-fMRI NF studies a simple thermometer for displaying the activity is used (Hartwell et al., 2016; Kirsch et al., 2016; Kohl et al., 2019). Alternatively, a dual-feedback approach in which the activity of two ROIs is displayed for regulation can be applied (Zilverstand et al., 2015). While this increases complexity of the task and might be more difficult for the participants to realize, it expands the possibilities of rt-fMRI NF.

To increase involvement in the task, feedback can be presented by means of a line graph or a video animation in which e.g. a rocket has to be controlled (Alegria et al., 2017; Orlov et al., 2018). In a social feedback design, Cordes et al. (2015) presented the face of a male person on the screen and the task was to increase the intensity of the smile of the person displayed. Although it might have been attractive for the participants to participate in such a task which could have been accompanied by an increased motivation, activation of other brain areas that were not actively targeted such as the fusiform-face area is very likely. Dudek & Dodell-Feder (2021) compared the frequencies of the distinct approaches and found that the thermometer was applied in 76% of the studies, whereas line graph and video were both represented with 12%.

Regarding exceptional ways of feedback presentation, in a virtual environment brain computer interface and rt-fMRI NF study feedback was served in the form of adjusted colors of the virtual environment (Lorenzetti et al., 2018). In another unique rt-fMRI NF study, haptic feedback in addition to visual feedback was provided (Young et al., 2020). These approaches seem very interesting, but it still remains open how feasible they are for future implementation due to the inherent complexity. Although each method has its' advantages and disadvantages, when selecting the optimal feedback presentation for the study at hand, the testability of the hypotheses should be of highest relevance. Therefore, it should be guaranteed that more complex approaches, while potentially being highly motivating, also allow for testability of the hypotheses and do not induce additional confounding factors.

1.1.3.3.1 Timing of feedback presentation

Another important decision aside to the decision of the feedback target is which time interval to use for the presentation of the feedback. Generally, feedback can be provided continuously or intermittently (Dudek & Dodell-Feder, 2021; Koush et al., 2013). In continuous feedback, an updated feedback value is presented after every new acquired brain volume (Watanabe et al., 2017). In intermittent feedback, feedback is usually shown at the end of a block only and presents the average value that was achieved during this particular block (Watanabe et al., 2017). A review compared type of feedback timing among NF studies (Fede et al., 2020). Over all conducted NF studies continuous NF, was more frequently applied (124 times) in comparison to intermittent NF (17 times). Whereas most NF studies opted for continuous NF, there are arguments for both of these perspectives. Intermittent feedback has the advantage of a decreased attentional and cognitive load (Scheinost et al., 2020). Furthermore, more scans are used for averaging which potentially results in an increased signal quality (Koush et al., 2013). Regarding artefacts, based on findings in resting-state fMRI with sliding window technique, the risk of noise might be higher in sliding window connectivity NF in which FC values of a specific timeframe are collected and summarized to one feedback value for each sliding window as opposed to intermittent feedback (Lindquist et al., 2014; Scheinost et al., 2020). On the other hand, the inherent delay of the feedback is smaller in continuous NF and therefore, it might be easier to associate it with the brain activity (Watanabe et al., 2017). The authors of the previously mentioned review (Fede et al., 2020) recommend application of intermittent feedback, because many studies found that this is more efficient than continuous feedback (Hellrung et al., 2018; Johnson et al., 2012), but for some regions such as the auditory cortex continuous feedback is advised. The latter conclusion is based on findings from an NF study by Emmert, Kopel, et al. (2017) who compared auditory cortex down-regulation by means of continuous and intermittent feedback and found that continuous feedback resulted in greater deactivation. Furthermore, using continuous feedback the effect increased over sessions which was absent with intermittent feedback. In sum, as with most other domains, it clearly depends on the existing research question which method is more suitable.

1.1.3.4 Implicit vs. explicit instructions

A common distinction among rt-fMRI NF instructions is whether these are implicit or explicit (Dudek & Dodell-Feder, 2021; Fede et al., 2020; Munoz-Moldes & Cleeremans, 2020). In explicit instructions, participants are presented with any kind of instruction regarding control of the feedback signal (Dudek & Dodell-Feder, 2021). This is lacking in implicit paradigms and therefore they potentially require more trial-and-error (Dudek & Dodell-Feder, 2021; Fede et al., 2020).

The previously mentioned review paper compared implicit and explicit feedback display. It was found that most of the previously conducted NF studies made use of explicit NF (129), while only 14 studies applied implicit NF (Fede et al., 2020). The authors recommend the use of explicit feedback for clinical applications (Fede et al., 2020). However, for specific cases e.g. decreasing fear response, implicit feedback seems more suitable (Fede et al., 2020; Koizumi et al., 2016; Taschereau-Dumouchel et al., 2018). A meta-analysis on the efficacy of rt-fMRI NF by Dudek & Dodell-Feder (2021) found no differences of explicit versus implicit instructions in efficacy of rt-fMRI NF on symptoms.

Adequacy of implicit and explicit tasks has also been investigated by means of EEG NF. Regarding this, a study by Kober et al. (2013) found that participants who did not use specific mental strategies during NF training were more successful in controlling the sensorimotor rhythm activity suggesting that no explicit strategies might be necessary. However, a study with rt-fMRI NF reported that performance improved as a consequence of the application of explicit regulation strategies associated with the functions of the brain areas (Scharnowski et al., 2015). According to Sepulveda et al. (2016) who found no improvement in regulation ability after explicit instructions, the role of explicit instructions remains unclear. Similarly, Sitaram et al. (2017) concludes that further research is required for demonstrating whether implicit/covert feedback is sufficient for learning.

Whilst appearing unambiguous at first glance, the differentiation between implicit and explicit NF tasks was criticized for its' lack of clarity, as "implicit" tasks exist that still include an explicit component and thus do not match the criterion in the proper sense (Munoz-Moldes & Cleeremans, 2020). Therefore, Munoz-Moldes and Cleeremans (2020) suggested a four-category taxonomy along which NF tasks can be classified.

They specifically underline the role of active control, awareness of NF and awareness of strategy and implement those factors in what they call: active overt cued tasks, active overt uncued tasks, active covert tasks and passive covert tasks (Munoz-Moldes & Cleeremans, 2020).

Based on this distinction, most published rt-fMRI NF studies would be placed in the first category which is described as including the factors active control, awareness of NF, and awareness of strategy (Munoz-Moldes & Cleeremans, 2020). The second category is similar to the first one with the only difference that cues in the form of cognitive strategies to self-regulate neural activity are absent.

The third category implies tasks that lack the awareness component but still include active control by not informing the participant about the fact that this could be achieved by changes in brain activity. The last category “passive covert tasks” has so far only been used in one study by Ramot et al. (2016). Here, none of the three dimensions is present. Particularly for the last two categories, the role of awareness is important and the authors suggest that for an improvement of the understanding of the learning mechanisms and eventually optimized future tasks additional attention should be paid to this construct. For a proper understanding, it has to be noted that definitions of awareness are diverse and awareness was defined as availability for report and voluntary control of action in their work (Block, 2007; Munoz-Moldes & Cleeremans, 2020).

1.1.3.5 Control groups

In rt-fMRI NF multiple options for control groups are available (Ros et al., 2020; Sorger et al., 2019). Among the most common control groups in rt-fMRI NF are yoke control groups (Weiss, Zamoscik, et al., 2020), control groups without any feedback (Hartwell et al., 2016), active control groups (Weiss, Aslan, et al., 2020), control groups which only differ from the experimental group in the receipt of instructions (Zilverstand et al., 2015) and control groups which receive computer-generated feedback (Mayeli et al., 2020).

A yoke design is characterized by each participant in the yoke control group receiving a signal that was recorded from a participant in the experimental group previously. While Sorger et al. (2019) argue that there is no single best control condition, Fede et

al. (2020) found the application of a yoke control group to be superior. Another option is the inclusion of a control group that does not receive any kind of feedback (Hartwell et al., 2016). An argument for this is that frustration of the participants due to false-feedback is prevented (Hartwell et al., 2016). However, it is questionable whether no feedback at all is less frustrating than this and additionally, whether participants would identify feedback false feedback as false altogether. An alternative is the use of an active control group aiming at control over regional specificity. In this approach the control group receives the feedback signal of a brain region entirely unrelated to the process at hand (Alegria et al., 2017; Gerchen et al., 2018; Weiss, Aslan, et al., 2020). Quite different from the previously mentioned approaches is the display of the same feedback signal in both groups with the only difference that the experimental group is provided with instructions for the regulation based on cognitive strategies, while the control group is told to use their intuition (Zilverstand et al., 2015). It is debatable whether this comparison can provide enough evidence for the actual efficacy of the NF training. Artificial computer-generated feedback is also applicable (Mihara et al., 2012; Sorger et al., 2019). In summary, the amount of consensus on which control group is most suited is limited (Fede et al., 2020).

1.1.3.6 Blinding

As with all psychological studies, the gold standard is the application of a double-blind design (Sitaram et al., 2017; Tursic et al., 2020). However, because of a higher level of complexity in designing and conducting the task, blinding the experimenter is sometimes difficult to implement. Still, our own experience has shown that double-blind procedures can be implemented well via an automated process. Although single-blind studies are not the optimum, the feasibility is increased and they are at present a common procedure in the field of rt-fMRI NF (Ros et al., 2020).

1.1.3.7 Number of sessions

The number of sessions, in particular how much rt-fMRI NF training is needed to achieve learning effects, how far apart these sessions should be in time and whether this also varies dependent on the target are further important aspects to consider (Rance et al., 2018; Strehl, 2014). Regarding the distance in time, it is known from early learning theories that spaced practice is preferred over massed practice (Hull 1943). With regard to rt-fMRI NF this would mean that the intervals between training

sessions should be long enough and that quality is supposed to be preferred over quantity. Though in reality it might be challenging to keep distances in time the same across the whole experiment due to restricted scanner availability. As mentioned previously, the amount of training sessions is associated with symptom improvements (Dudek & Dodell-Feder, 2021). Rance et al. (2018) stated that two sessions is a typical amount of training in clinical rt-fMRI NF studies. According to Auer et al. (2015), multiple sessions are necessary for controlling brain activity. An NF review presented a histogram on the amount of sessions that were conducted over all included NF studies (Fede et al., 2020). They found a median of two NF sessions per study with a general range from one to twelve sessions (Fede et al., 2020). Generally, authors agree that multiple sessions are needed, but that the optimal number is still an open question (Hartwell et al., 2016).

1.1.3.8 Moderators and predictors of NF success

One of the less extensively answered but still very important questions in the NF literature is how NF success can be manipulated. With cumulative knowledge of factors that drive NF success, future rt-fMRI NF studies could take these factors into account which in turn could enhance general NF success. Of course, there are some factors that cannot be addressed actively, but it is nevertheless necessary to shed light on these as they contribute to our understanding of the mechanisms behind rt-fMRI NF.

Interindividual factors such as motivation are important in the context of rt-fMRI NF success as well (Kadosh & Staunton, 2019; Strehl, 2014). Unsurprisingly, enjoyable tasks increase motivation (Strehl, 2014) and therefore a goal in designing the NF study should be to achieve a modest level of attractiveness of the task to increase commitment to it. However, this should be kept within a limit since an abundance of it could distract from the actual task.

In terms of training duration, Dudek and Dodell-Feder (2021) demonstrated that training minutes explained 29.9% of the variance in symptom outcomes as a consequence of rt-fMRI NF. Accordingly, the amount of training is crucial for the success of the method.

An important factor in the ability of self-regulation is the underlying cognitive load. Learning can be decreased by an elevated cognitive load (Weiskopf et al., 2004) so it

should be considered to keep the cognitive load as low as possible when designing the experiment. Further, the feedback modality (visual or auditory) can affect the self-regulation success (Weiskopf et al., 2004).

NF performance can also be driven by external rewards such as money. It was previously demonstrated that providing reward in addition to the NF score itself resulted in more successful performance than just one of the two (Sepulveda et al., 2016). Although motivation and external rewards can boost commitment and could consequently increase NF performance, it should be noted that the study should not be overloaded with such features so that the predefined hypotheses can clearly be tested and are most likely not strongly confounded with other factors.

Importantly, not all participants are equally able to perform rt-fMRI NF training. While some individuals can master these tasks very well, others are more limited in their ability to achieve changes in neural activity by means of rt-fMRI NF (Kadosh & Staunton, 2019; Sitaram et al., 2017; Weiskopf et al., 2004). The variability in the skill to modulate brain activity across individuals is large (Hampson et al., 2012) and there is evidence that NF success can be predicted by structural anatomy (Zhao et al., 2021) and functional activation (Scheinost et al., 2014). It will be an important question for future rt-fMRI NF research to gain further insight into these relationships.

In addition, on a neural level, in a meta-analysis on a number of rt-fMRI NF studies NF success was found to be strongly mediated by anterior insula and basal ganglia activity (Emmert et al., 2016).

While motivation as an NF moderator and training duration can be actively manipulated to a certain degree, the activity of brain areas that accompany NF target activity is not experimentally controllable, but it is worthwhile to take this into account in future NF studies. The same applies to structural and functional predictors of NF performance.

In conclusion, all of the previously described aspects in their sum make a strong difference in the quality of an rt-fMRI NF study and - as a result - can heavily affect the resulting findings. Therefore, it is particularly important to consider the implications of each decision and design aspect when planning a rt-fMRI NF study.

1.1.4 Connectivity NF

The following sections provide an overview on the latest progression in the field of rt-fMRI NF. A progressive alternative to the more common ROI-based NF is functional connectivity (FC)-based NF.

1.1.4.1 Functional connectivity NF

The first years after occurrence of rt-fMRI NF were dedicated to feedback of activation of single regions of interest (ROI) that were supposed to be regulated. Importantly, also in rt-fMRI NF studies targeting specific ROIs evidence for changes in connectivity as a consequence of NF training could be discovered (Paret et al., 2016; Young et al., 2018). With increasing experience in the application of the method the specific conduct of the experiments was refined (Dudek & Dodell-Feder, 2021; Fede et al., 2020; Ros et al., 2020; Sorger et al., 2019) and it was realized that modulation of neural dynamics on a network level could be more effective than rt-fMRI NF training of a single area (Sitaram et al., 2017). From a more holistic perspective, cognitive processes cannot be narrowed down to the activity of an isolated brain region but to coordinated activity amongst brain networks (Koush et al., 2013; Pereira et al., 2019; Scheinost et al., 2020). Therefore, simplistic ROI-based NF approaches might not capture complex connectivity-based processing mechanisms well (Pereira et al., 2019; Ruiz et al., 2014). Furthermore, local signal changes do not translate well to out-of-scanner behavior and the association between two brain regions' activity time courses might be highly effective clinically (Gabrieli et al. 2015, as cited in Scheinost et al., 2020; Watanabe et al., 2017). Based on this, an interest in manipulating brain activity in a broader way based on network theories was developed (Sitaram et al., 2017).

There are different approaches for connectivity-based rt-fMRI NF: Correlation-based NF (FC NF), Dynamic Causal Modeling (DCM), and whole-brain NF are examples for these (Koush et al., 2013; Scheinost et al., 2020; Watanabe et al., 2017).

1.1.4.1.1 Correlation-based FC

The basis of FC-based rt-fMRI NF is more straightforward and the computation of it is faster in comparison to DCM-based approaches (Watanabe et al., 2017). Inspired by Koush et al. (2013), Megumi et al. (2015) conducted one of the first rt-fMRI NF studies using measures of FC as a target. As a basic research question an enhancement in

FC was supposed to be achieved that should be kept for longer time spans. Target connectivity between lateral parietal and primary motor areas was investigated and an up-regulation in target FC of the networks under rest as a consequence of four rt-fMRI NF training sessions was evoked. In particular, this increase in FC lasted for over two months. A high methodological standard such as the duration of the training over four days and the application of a sham control group which was still very rare at that time point was implemented in this study. As a special feature resting-state whole-brain analysis of the degree of connectivity was applied (Hampson et al., 2012; Scheinost et al., 2014), in which for each voxel in the brain a number of voxel this voxel is connected to is extracted. Here, clusters with significant decreases in degree of connectivity were found in the experimental group only. These were located in the left lateral parietal region and posterior cingulate cortex. Taken together, the study provided an optimal first example for the efficacy of rt-fMRI NF training targeting FC.

More evidence for the feasibility of rt-fMRI NF FC was provided by Pereira et al. (2019). In an attempt to modulate motor areas, interhemispheric connectivity of the premotor cortex as a target in rt-fMRI NF FC training with HC was investigated and up-regulation of the network and the ability to retain this over the whole experimental block were found (Pereira et al., 2019).

In another seminal study, Yamashita et al. (2017) tested whether FC of left primary motor cortex and the left lateral parietal cortex could be up-regulated in one group and down-regulated in another. The subsequent changes were then compared between the two groups. Feedback was provided via a disk on the screen and participants were told that disk size was adjustable by performing better at tapping imagery. In fact, the disk size represented FC of left primary motor cortex and left lateral parietal cortex. Performance was rewarded financially in proportion to success. In a mixed-effects model a significant effect for day and interaction between group and day were found, but no effect for group. It is concluded that switch in connectivity across days differed among the groups. In a mixed-effects model split for each group a significant effect of training day in both groups was found and mean FC increased significantly from day one to day four in the increase group and decreased in the decrease group. Accordingly, FC was altered in the aimed direction.

As further evidence for the feasibility of FC NF, Zhao et al. (2019) found in healthy participants with high anxiety that rt-fMRI NF training of the connectivity between

VLPFC and amygdala during threat exposure significantly increased connectivity of target FC and additionally reduced anxiety levels. These findings provide promising evidence that changes in neural activity as a consequence of rt-fMRI NF training also translate to changes on the clinical level. Recently, in individuals with high trait anxiety it was found that an increase in target FC between DLPFC and ACC by means of two sessions of rt-fMRI NF FC could be achieved. Regarding clinical utility, as hypothesized by the authors, these FC changes were correlated with a decrease in anxiety (Morgenroth et al., 2020).

Regarding the application in clinical samples, Sreedharan et al. (2019) examined rt-fMRI NF FC performance by targeting Broca's and Wernicke's area to normalize language areas in patients presenting with expressive aphasia as a consequence of a stroke. Stronger connections in the left hemisphere in direction of normality were found in the experimental group as a result of the training. It is suggested that rt-fMRI NF FC presents with a high utility in language disorders.

The sum of these studies indicates that rt-fMRI NF FC is a promising method to induce changes in large-scale brain networks that could eventually be applied in clinical samples.

1.1.4.1.2 Dynamic causal modelling (DCM) based FC/effective FC

In DCM-based NF, also known as effective connectivity NF, a feedback score is calculated by comparing predefined models of causal brain networks consisting of a few regions and it is reckoned which model is most likely given the acquired data (Koush et al., 2017; Koush et al., 2013). The method dates back to approaches by Friston et al. (2003, 2007) that apply Bayesian model comparison (Koush et al., 2013). It is hypothesis-driven which implies that predefined hypotheses on the neural mechanisms are necessary (Koush et al., 2013). In a proof-of-concept DCM NF experiment Koush et al. (2013) targeted FC between left visual cortex and left parietal cortex and between right visual cortex and right parietal cortex and found that participants were able to regulate the connectivity-based signal. It is suggested that changes in FC could be evoked by shifting visual-spatial attention between the different visual fields (Koush et al., 2013).

A few years later, the same authors built on the previous results and applied the approach to a network involved in emotion regulation including the dorsal mPFC and the amygdala (Koush et al., 2017). It was found that the rt-fMRI NF FC training modulated the dominance of the top-down model in contrast to the bottom-up model. These changes were absent in the control group that received yoke feedback of one of the best performing participants in the experimental group. Additionally, the experimental group presented with a significantly more precipitous learning curve than the control group. In summary, first DCM-based FC approaches deliver strong evidence for its feasibility.

1.1.4.1.3 Whole-brain neurofeedback

A very novel and innovative development in the field of fMRI NF is whole-brain NF (Scheinost et al., 2020). In whole-brain NF connectivity from an already defined predictive network is condensed for construction of a feedback signal (Scheinost et al., 2020). A parcellation for computation of a connectivity matrix and networks (accumulation of edges/elements in the connectivity matrix) for summary of the connectivity (Scheinost et al., 2020) are necessary prerequisites. The edges are combined to a separate number that represents the strength of connectivity in each network (Scheinost et al., 2020). Intermittent in contrast to continuous NF is used (Scheinost et al., 2020). So far, only one study is known that applied this approach successfully.

1.1.4.2 Role of artefacts

Although connectivity NF is highly promising for inducing neural and clinical changes, the method is not mastered as well as ROI based NF due to its novelty and complexity. Measuring FC by means of fMRI presents with many issues such as the confounding influence of head movements that must be disentangled from the data (Geerligs et al., 2017). Heart rate variability was also found to be mistaken with FC, if not corrected for (Heunis et al., 2020; Rangaprakash et al., 2018). Last but not least, respiration is known as another potential source of physiological noise that can be easily mixed up with brain activity (Heunis et al., 2020).

In sum, the risk of including artefacts in the data is higher in FC based NF as opposed to ROI-based approaches and it is even more important to filter out confounding influences (Kasper et al., 2017), specifically in the online training signal.

1.2 Finding a suitable large-scale target network

Beside all other methodological aspects that were depicted previously, the main question an rt-fMRI NF study depends on is which target to regulate. Generally, the available options for qualified target regions are endless and the functional role of the brain area should be considered.

The following sections aim at finding an appropriate large-scale target network with regard to aspects of NF methodology illustrated earlier.

1.2.1 Transition to a transdiagnostic focus

Although the possibilities of NF training in many different mental disorders are highly promising its clinical application is still limited. Among many reasons for this, one main reason is the uncertainty about which approach and target process is most qualified in a given disorder as a huge variety of study designs is available. Secondly, scanning costs are high (Martz et al., 2020) and scanning time can be a scarce resource. Therefore, it is important to utilize the resources available as broadly as possible with a pronounced rationale.

From a neurobiological point of view, it is known that there are brain structures with central functions which are involved in a plurality of mental disorders such as the DLPFC. Additionally, usually it is not one specific area that is related to a specific mental disorder but rather a whole network that is disturbed (Koush et al., 2013; Pereira et al., 2019; Scheinost et al., 2020; Sitaram et al., 2017). Considering these aspects, one of the most practical applications of FC NF training would be transdiagnostic with the target being selected on the basis of networks that are disturbed in many mental disorders rather than based on a disorder-specific view. In general, transdiagnostic approaches focus on the similarities of different mental disorders as opposed to their differences (Dalgleish et al., 2020). If feasible, a transdiagnostic rt-fMRI NF FC approach could have the potential to be applied over a range of disorders and would thereby be more economical to develop and present with a high clinical utility.

1.2.2 Frontostriatal network as a candidate: Anatomy and functional characteristics

There are five anatomically isolated frontostriatal networks that connect parts of the frontal cortex with the striatum, basal ganglia and thalamus and that differ in their functioning (Tekin & Cummings, 2002). Over all circuits, lesions can cause severe impairments such as executive dysfunction, personality changes or apathy (Tekin & Cummings, 2002). The dorsolateral prefrontal circuit as one of the five networks is particularly involved in executive function/cognitive control such as learning new information, planning, shifting behavioral sets, setting-up motor programs, task-appropriate behavior, switching between tasks and memory search strategies (Beste et al., 2012; Ridderinkhof et al., 2004; Tekin & Cummings, 2002). Cognitive control is also required in filtering out irrelevant stimuli from entering consciousness and thereby prevents acquiring an overload of information. The networks are known to act in parallel rather than consecutively (Beste et al., 2012).

1.2.3 Frontostriatal dysfunctioning across mental disorders

A majority of mental disorders is associated with frontostriatal circuit dysfunction (Tekin & Cummings, 2002). For example, MDD is a prevalent heterogeneous psychiatric illness with impairments in functioning and quality of life, depressed mood, anhedonia and altered cognitive function (Schmidt et al., 2011; DSM-5; American Psychiatric Association, 2013). Lesions in the frontal cortex and caudate nucleus were detected in patients with MDD and it was suggested that dysfunctional coordination of limbic-cortical pathways play a key role in the development of it (Tekin & Cummings, 2002). It was previously found that MDD patients present with constricted FC between the ventral striatum and both ventromedial prefrontal cortex and subgenual ACC (Furman et al., 2011) and generally changed FC between striatal and prefrontal regions is often observed in the context of MDD (Mkrtchian et al., 2021).

Obsessive-compulsive disorder (OCD) is characterized by recurrent intrusive and unwanted ideas, thoughts and urges that are called obsessions and compulsions which include repetitive ritualistic cognitive and physical activities (APA, 2013). Evidence promotes the involvement of frontal-subcortical circuit structures in the pathogenesis of OCD (Tekin & Cummings, 2002). An example for this is that increased reward prediction error signals in ACC and putamen have been demonstrated, which points to a hyper-responsive learning network in OCD (Hauser et al., 2017). Reward

prediction errors represent a mismatch of expectations and experiences and are mainly involved in goal-directed behavior (Glimcher, 2011; Hauser et al., 2017; Schultz et al., 1997). Additionally, OCD patients present with reduced right caudate and left orbitofrontal region functional activation (Tekin & Cummings, 2002) and elevated FC along a ventral corticostriatal axis implying the OFC (Harrison et al., 2009).

In terms of substance use and addiction, the thalamus, OFC and limbic parts of the basal ganglia are mentioned frequently (Tekin & Cummings, 2002). In alcohol addiction, increased alcohol dependence severity is correlated to weaker FC between the putamen and prefrontal regions when performing the stop-signal task for response inhibition (Courtney et al., 2013). In cocaine addiction it has been found that reaction times in a stop-signal task were increased in comparison to healthy controls and that the activity of a frontostriatal thalamic network was negatively associated with reactions times in HC but not patients with cocaine use disorder (Wang et al., 2018).

From the domain of neurodevelopmental disorders, a 20-year follow-up study since childhood ADHD diagnosis has shown that patients present with dysfunctions in lateral frontostriatal and superior parietal regions in contrast to healthy controls in sustained attention tasks, emphasizing the perseverance of frontostriatal dysfunctions in adult ADHD (Cubillo et al., 2012).

1.2.3.1 The role of frontostriatal dysfunctioning in Schizophrenia

While frontostriatal networks are relevant in many mental disorders, they have a special role in SCZ. As a background, SCZ is among the most impairing psychiatric disorders (Correll & Schooler, 2020). It puts an immense burden on the affected individual that is noticeable in many aspects of life with difficulties in social interactions or in pursuing a normal profession (Foussias & Remington, 2010) as two examples. The disorder is further characterized by disorganization of thought processes and deficits in planning (APA, 2013). Whereas positive symptoms as part of the disorder can be relatively well treated pharmacologically, negative symptoms e.g. anhedonia and avolition as well as cognitive deficits show a rather low treatment efficacy (Correll & Schooler, 2020).

According to the frontostriatal hypothesis of SCZ, frontostriatal dysfunctioning is a major factor in the development of SCZ (Robbins, 1990). SCZ is described as a

disorder of dysconnectivity (Friston et al., 2016; Friston & Frith, 1995; Pettersson-Yeo et al., 2011) which is manifested by decreased connectivity of frontal with subcortical regions, especially frontostriatal hypoconnectivity (Lin et al., 2018; Shukla et al., 2019; Su et al., 2013). Correlations between reduced ventral striatum – ACC FC and SCZ symptoms (Lin et al., 2018) could be demonstrated. It has also been shown that unmedicated SCZ patients present with decreased activity in the ACC and striatum whereby an increased baseline activity in the striatum was able to predict a better treatment outcome and greater changes in activity were also correlated with better treatment outcomes (Cadena et al., 2018). Shukla et al. (2019) found that FC between striatum and right medial OFC significantly correlated with negative symptom severity. Viher et al. (2019) could show that FC between the left caudate nucleus and bilateral DLPFC and between the left putamen and bilateral supplementary motor area was decreased in SCZ patients. Changed connectivity from DLPFC to caudate nucleus was associated with performance on a fine motor task that was sensitive to psychomotor speed. Based on these studies, the associations of frontostriatal networks and clinical features are particularly strong in SCZ which highlights its suitability for neurobehavioral interventions such as rt-fMRI NF.

In sum, the widespread deviations in frontostriatal networks across a multitude of mental disorders and especially in SCZ underline their potential for transdiagnostic rt-fMRI NF approaches based on networks defined by FC that could be targeted by the method. Furthermore, the encouraging rt-fMRI NF studies in SCZ that were described in chapter 1.1.2.2 suggest that SCZ is a suitable population for the application of rt-fMRI NF FC.

1.3 Aims

This dissertation reviews the developments within the field of real-time fMRI NF and addresses the feasibility of large-scale network FC NF as a specific form of NF. It primarily aims at demonstrating the use of rt-fMRI NF FC to induce changes in FC of a large-scale frontostriatal brain network and point out the methodological challenges that are linked to this. Study one focused on the general feasibility to train the network in a single session of NF training. In study two the methodology was refined based on findings from the previous study and NF training was extended to three training days to demonstrate learning effects in rt-fMRI NF. Both studies have been pre-registered

before the experiments were conducted and the pre-registrations are enclosed in the supplement.

2 STUDY ONE: TESTING THE EFFICACY OF A LARGE-SCALE RTFMRI NF FC APPROACH

2.1 Just a very expensive breathing training? Risk of respiratory artefacts in functional connectivity-based real-time fMRI neurofeedback[1]

2.1.1 Abstract

Real-time functional magnetic resonance imaging neurofeedback (rtfMRI NFB)[2] is a promising method for targeted regulation of pathological brain processes in mental disorders. But most NFB approaches so far have used relatively restricted regional activation as a target, which might not address the complexity of the underlying network changes. Aiming towards advancing novel treatment tools for disorders like schizophrenia, we developed a large-scale network functional connectivity-based rtfMRI NFB approach targeting dorsolateral prefrontal cortex and anterior cingulate cortex connectivity with the striatum.

In a double-blind randomized yoke-controlled single-session feasibility study with N=38 healthy controls, we identified strong associations between our connectivity estimates and physiological parameters reflecting the rate and regularity of breathing. These undesired artefacts are especially detrimental in rtfMRI NFB, where the same data serves as an online feedback signal and offline analysis target.

To evaluate ways to control for the identified respiratory artefacts, we compared model-based physiological nuisance regression and global signal regression (GSR) and found that GSR was the most effective method in our data.

Our results strongly emphasize the need to control for physiological artefacts in connectivity-based rtfMRI NFB approaches and suggest that GSR might be a useful method for online data correction for respiratory artefacts.

[1] Publication: Weiss, F., Zamoscik, V., Schmidt, S. N.L., Halli, P., Kirsch, P., & Gerchen, M. F. (2020). Just a very expensive breathing training? Risk of respiratory artefacts in functional connectivity-based real-time fMRI neurofeedback. *Neuroimage*, *210*, 116580. doi:10.1016/j.neuroimage.2020.116580

[2] The applied abbreviation in this paper deviates from the abbreviation in the second paper and the general thesis, because the other abbreviation is more commonly used.

2.1.2 Introduction

In recent years the development of real-time fMRI neurofeedback (rtfMRI NFB) approaches is transforming fMRI from a knowledge-generating technology into a neurobiological intervention tool for mental disorders (Bagarinao et al., 2006; Kohl et al., 2019; Paret et al., 2019; Weiskopf et al., 2003). In rtfMRI NFB, covert brain processes are displayed in near real-time to make them accessible for targeted regulation by participants in the MRI scanner. However, the plethora of potential confounding noise sources in fMRI, like respiratory artefacts, requires special caution in the development of meaningful rtfMRI NFB approaches.

Lately, a growing number of studies could demonstrate the general ability of rtfMRI NFB to change neural activation patterns related to aberrant brain function in mental disorders (Ramot et al., 2017; Young et al., 2018; Zilverstand et al., 2017). For example, recent rtfMRI NFB studies could show that the NFB procedure induced changes in the activity of targeted brain areas (Karch et al., 2015; Karch et al., 2019; Kirsch et al., 2016). Several rtfMRI NFB studies have been carried out with schizophrenic (SCZ) patients as a target population. It was found that these patients were able to downregulate activity of the superior temporal gyrus (Orlov et al., 2018), upregulate the insula (Ruiz et al., 2013), or control anterior cingulate cortex (ACC) activity (Cordes et al., 2015) during rtfMRI NFB. Additionally, a recent study demonstrated a heightened pairing of default mode network (DMN) and language areas because of rtfMRI NFB (Zweerings et al., 2019).

However, most rtfMRI NFB research has hitherto mainly focused on regional BOLD activation as the target process (Caria et al., 2012; Dyck et al., 2016; Karch et al., 2015; Nicholson et al., 2017; Paret et al., 2016). Although alterations in brain connectivity has been identified as a relevant mechanism in many mental disorders (Braun et al., 2018; Fornito & Bullmore, 2015), much fewer studies have utilized functional connectivity (FC) measures (Megumi et al., 2015; Yamashita et al., 2017; Zhao et al., 2019) and even fewer studies used more complex measures like network-based approaches (Ramot et al., 2017) or effective connectivity (Koush et al., 2017; Koush et al., 2013). Thus, rtfMRI NFB requires further development until fully its potential of translating the results obtained with modern fMRI analysis methods like network

analysis into directed interventions for regulating and normalizing distributed and complex brain processes in mental disorders is achieved.

Despite even higher risks of confounding noise effects accompanying more complex rtfMRI NFB approaches, the development of such methods might be a path worth following to address complex pathological neural phenotypes. These include changes in large scale neural networks in mental disorders such as Major Depressive Disorder (Kaiser et al., 2015), ADHD (Qian et al., 2019), or SCZ. SCZ has been characterized for a long time as a network disorder of brain dysconnectivity (Friston et al., 2016; Friston & Frith, 1995; Pettersson-Yeo et al., 2011), including reduced connectivity of frontal with subcortical regions of which frontostriatal hypoconnectivity is most prominent (Lin et al., 2018; Shukla et al., 2019; Su et al., 2013). Patients with schizophrenia are showing aberrant extra-striatal connectivity during psychosis, for example decreased FC of the putamen with right anterior insula and dorsal prefrontal cortex (Peters et al., 2017). Moreover, there is evidence of a relationship between decreased ventral striatum – ACC FC and SCZ symptoms (Lin et al., 2018). First-episode psychosis patients had lower FC between the putamen and anterior cingulate cortex, and this connectivity was predictive for the further development of negative symptoms and general functioning (Oh et al., 2020). Consequently, disease-related brain networks might be a promising target for connectivity-based rtfMRI NFB in SCZ.

As a first step towards this goal we developed a novel large-scale network connectivity-based rtfMRI NFB approach to target frontostriatal connectivity deficits of the DLPFC and ACC with the striatum in SCZ and applied the method in a preregistered double-blind randomized yoke-controlled single-session pilot study with healthy controls (N=38) which we report here. In this manuscript we were unable to test our preregistered hypotheses, because during analysis we realized the presence of massively confounding physiological, especially respiratory, effects in the data. While we applied online motion parameter regression, spike regression of volumes affected by frame-to-frame movement, and regression of a cerebrospinal fluid (CSF) signal to clean the NFB signal from confounding factors, we did not collect prior measures to address physiological contamination of the feedback signal.

Importantly, the BOLD signal can be influenced by a variety of sources that can be labelled as noise. Examples of noise include structured noise i.e. gross subject movement and physiological sources (e.g. respiration and cardiac features) (Liu, 2016)

as well as random noise (e.g. thermal noise). Unsurprisingly, it is longstanding knowledge that retrospective corrections should be applied to the data to ensure reasonable quality of the findings - a notion pointed out in 1995 (Hu et al., 1995).

Especially rtfMRI NFB methods using FC-based signals as the feedback source face a number of methodological problems which might be more pronounced than in activation-based feedback (Power et al., 2012; Power et al., 2015). In the analysis of FC-fMRI data, particularly physiological artefacts that strongly affect connectivity must be considered (Nikolaou et al., 2016). Physiological features such as heart rate and respiration mostly influence the connectivity of resting state networks (Chu et al., 2018; Nikolaou et al., 2016), underlining the need to control for these factors.

The experimental procedures we based our study on, however, are in line with general procedures in the fMRI field. While motion correction is nowadays applied by default in fMRI analyses, physiological noise correction is still conducted only in a much smaller, although growing, number of studies. This is despite physiological artefacts forming one of the largest proportions of noise in general (Kruger & Glover, 2001). Therefore, in rtfMRI NFB, it is of large interest to subtract as many of these noise sources as possible from the data as the outcome of the whole procedure strongly depends on the validity of the feedback signal. Failure to correct for any of these sources might bias the whole procedure towards training unwanted strategies that are relatively easy to apply for participants, like changing breathing patterns.

Alarmed by our findings of confounding physiological effects, we tried to identify possible ways to control the identified confounding effects in our data and present the results of analyses with two different techniques for physiology correction, namely global signal regression (GSR) (Aguirre et al., 1998; Power et al., 2015) and model-based correction for physiological noise signals with the TAPAS PhysIO Toolbox (Kasper et al., 2017).

2.1.3 Methods

2.1.3.1 Participants

Healthy participants with normal or corrected-to-normal vision, eligible for MRI scanning, and without a history of mental or neurological disorders, prior and current psychiatric diagnoses, pregnancy, or acute intake of any medication except for oral contraceptives were recruited from the student population at Mannheim and Heidelberg. Two participants of the original sample of N=40 had to be excluded from analysis due to technical problems that occurred during scanning and N=38 healthy participants (23 women; age: 23.39 ± 4.24 years; age range: 18-35) were analyzed. Before participation, the experimental procedures were explained and participants provided written informed consent. During the experiment, participants were automatically assigned in a double-blind procedure to one of the two experimental groups: real neurofeedback (real NFB) or yoke neurofeedback (yoke group) through a predefined randomization list. The study was approved by the Ethics Committee of the Medical Faculty Mannheim at the University of Heidelberg, Germany (2018-520N-MA) and all procedures complied with the World Medical Association's Declaration of Helsinki.

2.1.3.2 Preregistration

The study was originally planned for testing the capability of participants to modulate the target network with rtfMRI NFB and was preregistered at the Open Science Foundation (OSF NeCoSchi https://osf.io/d6fre/). The confounding physiological effects described in this paper were not expected a priori and thus not preregistered, thus the reported analyses are exploratory.

2.1.3.3 Data/code availability statement

Raw fMRI data cannot be made publicly available due to protection of sensitive personal data. The summary data the analyses were based on are available at the OSF project site. We further provide the code to estimate the summarizing respiratory parameters at the OSF project site (OSF NeCoSchi https://osf.io/d6fre/).

2.1.3.4 MRI scanning

MRI scanning was conducted at two 3T Siemens Trio TIM Scanners (Siemens Healthineers, Erlangen, Germany) at the Central Institute of Mental Health in Mannheim, Germany. MR images were obtained with a 32-channel head-coil. T1-weighted structural images were acquired with a repetition time of TR = 2.3 s, echo time of TE = 3.03 ms, flip angle = 9°, 192 slices, slice thickness = 1 mm, voxel dimensions = 1 x 1 x 1 mm^3, FOV = 256 x 256 mm and a matrix size = 256 x 256. BOLD signals were measured using an echo planar imaging (EPI) sequence with TR = 1.64s, TE = 30ms, flip angle = 73°. The whole brain was partitioned in 30 axial slices (3 mm of thickness) with a voxel size of 3x3x3 mm^3 and a field of view of 192 mm. All functional runs were acquired with the same EPI sequence.

2.1.3.5 Brain network definition

We focused the rtfMRI NFB approach and our further analyses on a bilateral network comprising the dorsolateral prefrontal cortex (DLPFC), the anterior cingulate cortex (ACC) and the striatum. The ACC was defined based on the Neuromorphometrics Atlas included in SPM12 (Wellcome Department of Cognitive Neurology, London, UK). The DLPFC was extracted from an automatical metaanalysis with Neurosynth (https://neurosynth.org/; (Yarkoni et al., 2011) on the term ‘DLPFC’. In the next step we used the brain parcellations by Schaefer et al. (2018) for the cortical regions and by (Choi et al., 2012) for the striatum and extracted the ROIs that fell into the defined regions, providing 22 DLPFC ROIs, 23 ACC ROIs, and 13 striatum ROIs in both hemispheres which adds up to a total of 58 regions. Both approaches that we used are based on the 7-network parcellation of the cerebral cortex by Yeo et al. (2011). While this network definition procedure is relatively complex, it was chosen because we are aiming toward developing the basis for a flexible network-based neurofeedback approach that enables the estimation of complex graph-theoretical network measures and allows identification of potential sub-networks to further refine target networks.

2.1.3.6 rtfMRI NFB training

MRI scanning was conducted in a single session. Before scanning, participants provided demographic information and answered questionnaires including the German

Version of the Beck Depression Inventory-II (BDI-II) (Beck et al., 1996), the NEO Five-Factor Inventory (NEO-FFI) (Costa & McCrae, 2008), Schizotypal Personality Questionnaire (SPQ) (Raine, 1991) and a sensory inventory (Zamoscik et al., 2017).

The scanning session comprised a 5:21 min T1-weighted anatomical MPRAGE scan, a resting state run, two NFB runs, and a transfer run of 9:29 min each. During the resting state measurement a fixation cross was shown at the center of the screen and participants were instructed to keep their eyes open while not thinking about specific things. A transfer run is often included in NFB experiments to test whether regulation ability transfers to situations without feedback. In our case the transfer run was exactly similar to the resting state run except that participants were instructed to use the strategies learned in the NFB runs to upregulate the target network. During NFB, a thermometer display was presented on the left and right side of a fixation cross and continuously updated every TR. This feedback signal showed averaged dynamic FC of DLPFC/ACC ROIs with striatum ROIs, and participants were instructed to upregulate the feedback signal. Participants in the yoke control group did not receive their own feedback signal, but the saved signal of a participant from the real NFB group. Data of participants from the real group were saved in a first in first out queue, meaning that each recorded signal is used once in the yoke procedure, and a yoke participant always receives the first unused recorded signal. To ensure availability of data for the yoke procedure, the first three participants were assigned to the real feedback group. After completion of the scanning session, participants were interviewed, and group assignment was disclosed. For a graphical representation of the experimental setup and design, please see Figure 1.

2.1.3.7 Online data analysis

All online and offline analyses were conducted in MATLAB (R2017a, MathWorks Inc., Sherborn, MA, USA). Online rtfMRI NFB processing was conducted with in-house software based on SPM12 functions.

During the resting state scan, the acquired anatomical image was segmented and normalized to the SPM 12 TPM template. The inverse projection of the normalization was then applied to map the ROI masks into individual subject space. During rtfMRI NFB scanning each newly acquired volume was directly written to the analysis laptop

and realigned to the first volume of the series. Then, averaged intensity values from all ROIs in this image were extracted and added to the ROI signal time series. At each step a general linear model (GLM) was applied over the whole available data to correct for movement parameters estimated during realignment, a cerebrospinal fluid (CSF) signal, and spikes associated with head movements (framewise displacement (FD) > 0.5mm). Fisher Z-transformed Pearson correlations were then calculated from the last 30 points (i.e. implementing a sliding window size of 30 volumes) of the corrected time courses of all cortical ROIs with all striatal ROIs and averaged to determine the feedback value presented to the participant. To assure availability of sufficient data for the online correction algorithm, the first feedback value was calculated after 37 volumes (60.68 s). The feedback signal was only calculated from windows that included at least 15 volumes not affected by head motions (FD < 0.5mm), otherwise the feedback value would not be estimated and the thermometer display would be kept constant until a window with sufficient data occurs.

2.2.3.8 Offline data analysis

Offline data analysis was conducted with SPM 12 (v7219). The anatomical image was segmented and normalized to SPM 12 TPM template MNI space. The first 15 volumes of the functional data were discarded and the functional images were slice-time corrected, realigned, co-registered to the MPRAGE, normalized, and smoothed with an isotropic Gaussian kernel of 6 mm full width at half maximum.

A first level GLM was set up that included six movement parameters, the CSF signal, and dummy regressors for volumes affected by head motion (framewise displacement > 0.5mm; scan-to-scan global signal change z > 4) and a constant. Runs which exceeded a level of 20% movement-affected volumes were excluded from further analysis (3 runs in 3 subjects). Runs with problems with physiology recording were also excluded from analysis (2 runs in 2 subjects).

To assess possible methods to correct for physiological associations, the analyses were repeated with global gray matter signal regression (GSR), model-based physiological nuisance regression (Physio), as well as both (Physio & GSR) implemented in the first level model. To be more comprehensive we also added a repetition with white matter signal regression (WMR). Then, the time courses from the ROIs used in the online NFB procedure were extracted from the residual images of the

first level analyses and large-scale network connectivity was estimated from averaged Fisher Z-transformed Pearson correlations between the DLPFC/ACC ROIs and the striatal ROIs.

Second level analyses were conducted based on these DLPFC/ACC-striatum large-scale network connectivity values. In all second level analyses, age, gender and scanner were included as covariates. For group comparisons, we used two-sample t-tests implemented in a GLM model, which allowed for the addition of covariates. Associations of connectivity values with physiological parameters were assessed with partial correlations. As our main analyses are aiming at demonstrating confounding physiological effects, we did not correct the reported p-values for multiple comparisons, because multiple comparison correction might potentially hide nuisance effects in this case.

2.1.3.9 Physiological noise correction

Respiration and heart rate were recorded with a pulse finger clip and a respiration belt during MRI at a sampling rate of 50 Hz using built-in equipment (PMU Wireless Physio Control, Siemens Healthineers, Erlangen, Germany). Before estimating physiological parameters, we cut the physiological recordings based on recorded volume triggers so that they were exactly aligned with the analyzed fMRI data. Then we used the TAPAS PhysIO Toolbox (Kasper et al., 2017); http://dx.doi.org/10.1016 /j.jneumeth.2016.10.019) to estimate 20 physiological nuisance regressors, including cardiac, respiratory, cardiac x respiratory interaction (RETROICOR order 1), heart rate variability (HRV; RETROICOR order 3), and respiratory volume per time (RVT; RETROICOR order 4) terms. The derived physiology nuisance regressors were included in the first level GLM of the respective analyses (see above), and an F contrast over all physiology regressors was estimated to test whether physiology correction worked (see Supplementary Figure 1 for six examples.). We also repeated the analyses with a more complex model with 39 regressors that included temporally shifted versions of the respiratory response function before convolution with RVT (shifts from -24 s to 18 s in 6 s steps and additional shifts of -3 s, -1 s, 1 s, and 3 s) and HRV as described by (Bianciardi et al., 2009).

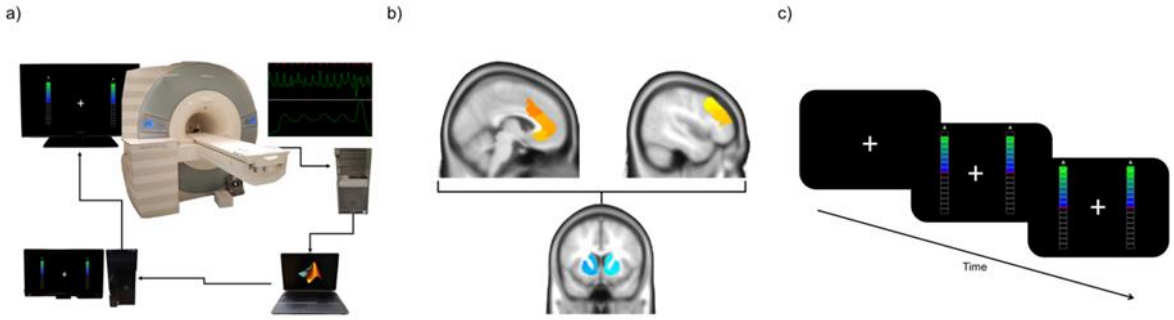


Figure 1. Experimental setup and design. a) rtfMRI Neurofeedback Setup. Acquired images are reconstructed and sent to a laptop running in-house MATLAB scripts to preprocess the images and extract the neurofeedback signal. The feedback is sent to a computer running Presentation software and presented to the participant in the scanner as a thermometer value. During scanning physiology measures (respiration and cardiac) are recorded. b) Target Network. The feedback signal represented the averaged functional connectivity of ROIs in the ACC (upper left) and the DLPFC (upper right) with ROIs in the striatum (below). c) Experimental Design. A fixation cross is shown to the participant in the scanner. After approximately one minute, the feedback signal is displayed as a thermometer value. This value is constantly adapted to represent the latest FC value estimated from the last 30 volumes. During resting state and transfer runs, only the fixation cross is presented.

2.1.3.10 Physiological parameters

While the PhysIO toolbox provides physiology time courses, we also calculated additional summarizing respiratory and cardiac parameters from the time courses that are potentially associated with the BOLD signal (Zamoscik et al., 2018) to test for confounding associations over subjects.

We created a set of respiration parameters: Breath Rate (peaks/breaths per minute), expiratory pause duration, its variance, and expiration-to-inspiration time ratio. For these parameters, expiration was defined as starting at each maximum peak and ending at the lowest local minimum before the next maximum peak, and correspondingly, inspiration was defined as the opposite. With these data we calculated the expiration-to-inspiration time ratio. For detecting expiratory pauses, we calculated the slope of the respiration curve with a sliding window approach (window size of 5 samples) to find clusters of minimum peaks which were then used to determine rough temporal markers for a provisional pause onset. This was recursively extended into both directions based on the slope parameters to determine pause onset and offset. In addition to the expiratory pause duration, we calculated the coefficient of variance (standard deviation divided by the mean) of pause duration (Pause CV; see

Supplementary Figure 5). For a more detailed description of the respiratory parameters please see also (Zamoscik et al., 2018). Additionally, we estimated heart rate (beats per minute) and two heart rate variability parameters, namely the standard deviation of the length of all beat intervals (SDNN [ms]) and the root mean square of successive differences of intervals (RMSSD [ms]).

2.1.4 Results

2.1.4.1 Functional connectivity group comparison

Connectivity estimates of the NFB runs were compared between the experimental groups with a one-sided independent samples t-test in accordance with the preregistered hypotheses. Higher large-scale network connectivity between DLPFC/ACC and striatum was found in the real NFB group in comparison to the yoke group during the first NFB run (NFB1: t(31) = 1.81, p = .040). Comparisons of connectivity estimates of the second NFB and the transfer run did not yield significant effects (NFB2: t(32) = .66, p = .258; transfer: t(30) = .615, p = .277). After including physiological nuisance regressors estimated with the PhysIO toolbox in the first level model, the group comparison of NFB1 remained significant (t(31) = 1.70, p = .049). However, when GSR or GSR and physiology correction combined were applied to the data, the effect in the first NFB run was no longer significant (GSR: t(31) = 1.09, p = .142; GSR and physiology correction: t(31) = .84, p = .205) (see Figure 2 and Supplementary Figure 12).

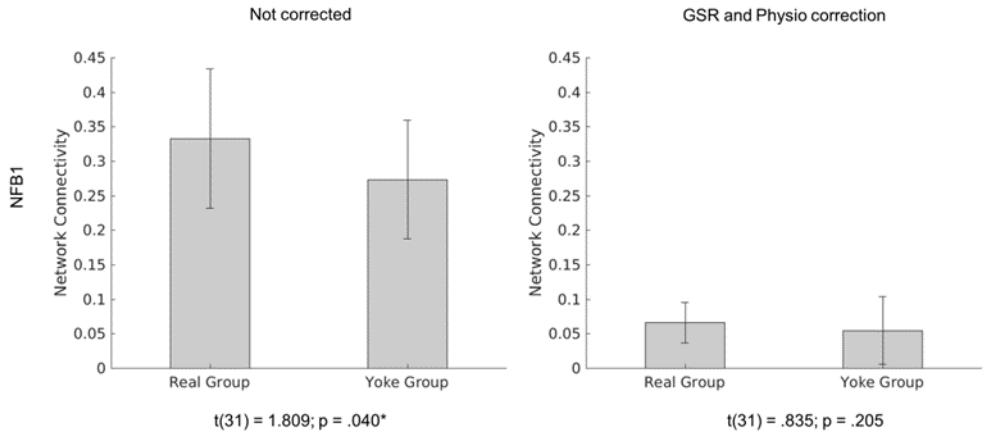


Figure 2. Group effect in the first neurofeedback run (NFB1). a) Group comparison of network connectivity between the real NFB and yoke NFB group uncorrected for physiology. b) Group comparison of network connectivity between the real NFB and yoke NFB group in data with global signal (GSR) and model-based physiology (Physio) correction.

2.1.4.2 Physiological associations

2.1.4.2.1 Correlations between physiology measures and functional connectivity

We then investigated the correlations of physiological measures with the target FC of the respective runs. These analyses were conducted separately for each correction method. Here we report the associations with Breath Rate, Pause CV and RMSSD in the first NFB run. The results for the other parameters and the other runs are similar and are presented in detail in the supplement.

2.1.4.2.2 Data not corrected for physiology

The analyses consistently showed strong significant correlations of respiratory physiological parameters with the target FC during NFB1 (see Figure 3, a and e and Supplementary Figures 4-7 for all runs). Cardiac parameters showed no correlation with our FC estimate in NFB1 (RMSSD: rho = -.026, p = .888) (see Figure 4, a), but a relatively weak association in the resting state data (Supplementary Figure 8-10).

2.1.4.2.3 Physio correction

When physiology correction with nuisance regressors estimated with the PhysIO toolbox was used, only minor changes in the association of physiology and target FC were seen. As shown in Figure 3 (b,f), Breath Rate and Pause CV both exhibited highly significant correlations with the target FC during NFB1 (Breath Rate: rho = -.448, p = .009; Pause CV: rho = .606, p = 1.8659e-04). For further details, see Supplementary Figures 4-7. In line with the uncorrected analyses, the cardiac parameter RMSSD showed almost no association with the respective BOLD signal (rho = -.027, p = .883) as can be seen in Figure 4 (Supplementary Figures 8-10). Using a more complex model with shifted respiratory response functions did only slightly diminish these associations during NFB1 (Breath Rate: rho = -.400, p = .023; Pause CV: rho = .477, p = .006; see Supplementary Figure 14).

2.1.4.2.4 Global signal correction

In comparison to the previous approach, applying GSR in the first level analyses yielded non-significant correlations between target FC and all parameters for respiration (Breath Rate: rho = -.079, p = .662; Pause CV: rho = -.060, p = .74) and

HRV (RMSSD: rho = .047, p = .797; Figure 4 (c, g) during NFB1 (Supplementary Figures 4-10). In an exemplary subject we show a strong association of the global signal with the data (Supplementary Figure 2a). In contrast to GSR, white matter regression (WMR) did not eliminate the associations during NFB1 (Breath Rate: rho = -.27, p = .129; Pause CV: rho = .558, p = 7.4163e-04; see Supplementary Figure 14).

2.1.4.2.5 Global signal and physio correction

The combination of GSR and model-based physiology nuisance regressors likewise resulted in non-significant correlations across all parameters (Breath Rate: rho = -.156, p = .387; Pause CV: rho = .016; p = .93; RMSSD: rho = .007, p = .969), see Figure 3 (d, h) and 4 (d) and Supplementary Figures 4-10. The association of the global signal with the data is much larger than the association of model-based nuisance regressors (Supplementary Figure 2b).

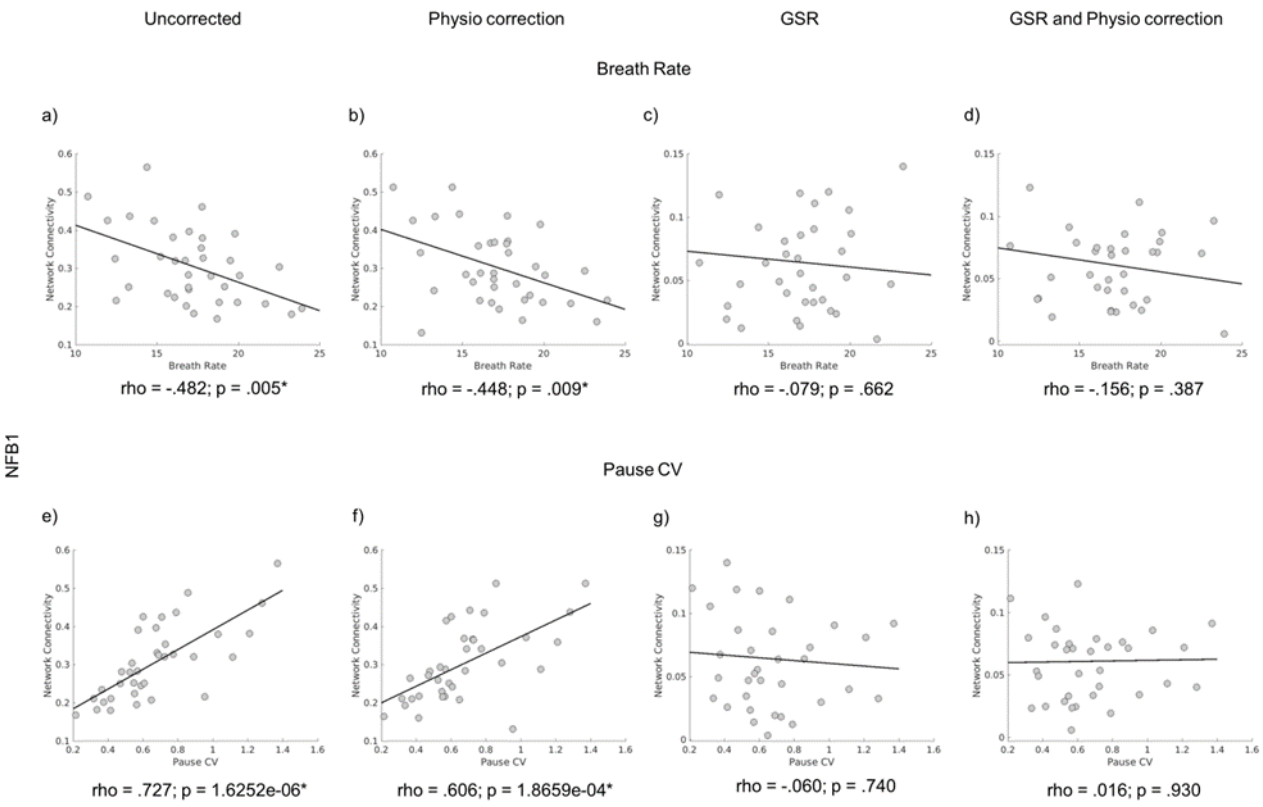


Figure 3. Association of respiratory parameters with network connectivity in the first neurofeedback run (NFB1). Correlations of the respiratory parameters Breath Rate and Pause CV (standard deviation of respiration pause duration divided by its mean), a measure for regularity of breathing, with target network connectivity for the different physiology corrections.

Physio: model-based physiology correction; GSR: global signal regression. Please note that, as expected, GSR overall shifted connectivity estimates towards 0.

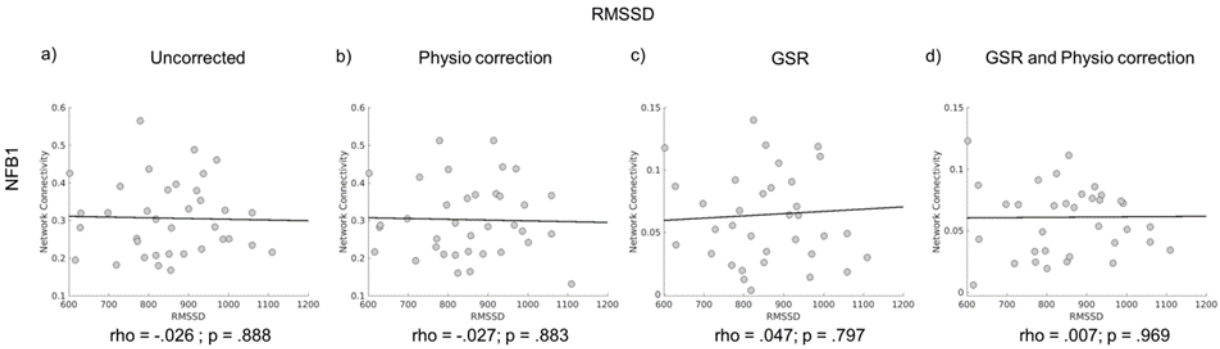


Figure 4. Association of cardiac parameters with network connectivity in NFB1. Correlations of the cardiac parameter RMSSD (root mean square of the successive differences [ms]) with target network connectivity for the different physiology corrections. Physio: model-based physiology correction; GSR: global signal regression. Please note that, as expected, GSR overall shifted connectivity estimates towards 0.

2.1.4.2.6 Change in physiological parameters between runs in the experiment

Because of the associations between respiratory parameters and connectivity estimates we explored whether respiration changed between runs differently in the groups (see also Supplementary Figure 11) to identify training effects on respiratory parameters. We first tested whether the parameters changed between the resting state run and the first NFB run. On a descriptive level, Breath Rate was slightly increased in the real feedback group (mean change: .245) and even more in the yoke feedback group (mean change: 1.22), but the groups did not show significant differences in Breath Rate between rest and NFB1 ($t(31) = -1.007$, $p = .322$). However, the change in Breath Rate was negatively associated with the change in connectivity ($rho = -.533$, $p = .001$). For Pause CV, the change was significantly different between the groups ($t(31) = 2.085$, $p = .045$) with a mean change of .258 in the real group and a mean change of .017 in the yoke group. Furthermore, we found a strong relationship between Pause CV change and change in connectivity ($rho = .757$, $p = 3.3812e-07$).

Between the first NFB run and the second, we did not find significant group differences in the change of these respiratory parameters. On a descriptive level, Breath Rate was reduced in the real feedback group (mean change: -.742) and not in the yoke group

(mean change: .007), but the groups did not significantly differ in Breath Rate change between NFB1 and NFB2 ($t(31) = -1.257$, $p = .218$). However, the change in Breath Rate was negatively associated with the change in connectivity ($rho = -.515$, $p = .003$). For Pause CV, the real group had a mean change of -.064 and the yoke group of .037, and the change was also not significantly different between the groups ($t(31) = -1.092$, $p = .283$). For Pause CV there was no association between parameter change and change in connectivity ($rho = .232$, $p = .201$).

2.1.5 Discussion

We conducted a double-blind randomized yoke-controlled single-session pilot trial originally designed and preregistered to test the feasibility of a newly developed large-scale FC rtfMRI NFB approach targeting DLPFC/ACC-striatum FC. When testing the preregistered hypotheses we found only a weak effect in the first NFB run (NFB1). After applying corrections for physiological artefacts, this effect could not be detected any longer. This allowed us to assess the influence of physiological parameters on our FC estimates in exploratory analyses. In our data we identified worryingly strong relationships between parameters reflecting the rate and regularity of breathing and our target large-scale network FC signature during all runs of the experiment, as well as a putative training effect on the regularity of respiration (Pause CV) from the resting state to the first NFB run.

Of note, our results are based on second-level analyses, where interindividual differences in respiration between participants were strongly associated with differences in large-scale network connectivity. This is relevant in rtfMRI NFB because the comparison of different subjects in different groups is often used as the level of analysis, and differences in respiration or changes in respiration during NFB training might lead to false positive results. We are aware that a single NFB training session may not provide enough data to assess whether our large-scale rtfMRI NFB approach worked, or whether a training effect of physiological parameters occurred. However, we collected a sufficient amount of data to clearly identify the associations of our large-scale network FC measure with respiratory parameters.

With hindsight, these relationships should have been expected. In general, the existence of respiratory artefacts has been well described for fMRI (Caballero-Gaudes & Reynolds, 2017; Chu et al., 2018; Nikolaou et al., 2016; Power et al., 2018; Power et al., 2017). These low-frequency confounds reduce the sensitivity of the signal (Murphy et al., 2013) and are not extracted by standard physiology corrections (Birn et al., 2006) and low pass filtering (Liu et al., 2017).

Physiological artefacts have an especially heavy influence on FC estimates (Nikolaou et al., 2016). Particularly respiration effects are difficult to distinguish from FC due to identical spatial locations and temporal characteristics (Birn et al., 2006). Our target network was more strongly affected by respiration than by cardiac features, suggesting

that in frontostriatal networks cardiac features have little effect on the BOLD response. Here, our results are in line with the finding that cardiac properties generally influence the global BOLD signal only to a small degree (Power et al., 2018). Whereas CSF regions and regions near greater arteries and draining veins are especially vulnerable for noise of cardiac nature (Bhattacharyya & Lowe, 2004; Birn et al., 2006; Caballero-Gaudes & Reynolds, 2017).

Thus, the consequences of respiratory artefacts are concerning and it is important to control for them. This is of even higher importance in the case of rtfMRI NFB, where the fMRI signal or a derived measure is the fed back to be modulated by the participants in addition to being the main analyzed data. If confounding effects in the feedback signal are not corrected, it might be much easier for the participants to manipulate the target signal by changing physiological parameters like breathing patterns instead of regulating brain processes. For example, (Ramot et al., 2017) reported that subjects stated that they have changed their breathing patterns as a behavioral strategy to regulate the neurofeedback signal. Indeed, a common objection against rtfMRI NFB is the danger of confounding noise that might contaminate the feedback signal. Our empirical results suggest that there is indeed a real danger for respiratory artefacts, at least in connectivity-based rtfMRI NFB, which may simply lead to a costly form of breathing training, rather than an effective NFB treatment. To identify ways to control for the unintended associations, we assessed two possible approaches to correct for the identified artefacts in our data, the inclusion of model-based physiological nuisance regressors estimated with the TAPAS PhysIO toolbox, and the simpler but also disputed inclusion of the global gray matter signal in the first level model. While we tested these approaches offline with already acquired data, we chose them because both could be implemented relatively easily in the online rtfMRI NFB procedures.

In general, the PhysIO toolbox and the regression of physiological nuisance parameters seemed to work at least satisfactorily. F-contrasts over all physiological regressors included in the respective first level analyses showed strong relationships with the data, with only very few runs showing weak relationships (see Supplementary Figure 1). However, the model-based physiology corrections resulted in virtually no change of the second level association of connectivity estimates with physiological parameters. This was surprising, as the model-based approach represents the current

state-of-the-art for physiology correction. It nonetheless seems relatively unlikely that this was due to failures in the application of the toolbox or in the toolbox itself. We have carefully double-checked our analyses, and even if physiology correction did not work as perfectly as possible, the large amount of variance removed together with its very small influence on the second level associations makes it unlikely that this would be substantially changed if the removed variance would not be increased much further. Nonetheless, we cannot exclude the possibility that the used physiological recordings are not of optimal quality as the built-in recording devices probably provide suboptimal measurements which could influence the quality of the model-based physiological correction. However, our data quality should be at a comparative standard level and reflect the level available at other sites with similar standard equipment.

In comparison to model-based physiology correction, model-free GSR was capable of attenuating the associations between the same physiology parameters and target measure sufficiently, resulting in non-significant correlations. The association was even further attenuated when a combination of GSR and Physio correction was used. GSR seemed to reduce variance in both groups, as well as the mean difference between groups. After GSR the residual variance in the first NFB run seems to be relatively larger in the yoke group than in the real group (Figure 2), which might be expected. Participants in the yoke group have no control about the brain process they try to regulate. Therefore, they should apply diverse strategies that randomly influence the target process and increase variance.

Several issues might be related to the difference between the model-based and model-free GSR approaches. An apparent difference between GSR and Physio correction is the strength of association with the data, and thus the amount of variance removed by the approaches, which is much larger for GSR (Supplementary Figure 2). This might reflect the ability of the global signal to capture more noise in the data than the model-based physiology nuisance regressors. Another possibility might be that the intraindividual physiology nuisance regressors do not eliminate interindividual differences in the baseline of physiological parameters, which are then taken up by the second level analyses at the group level. Furthermore, it was recently shown that the used measure of respiratory volume over time (RVT) can have inferior correction performance for respiratory artefacts in comparison to alternative approaches like RVT envelope of the waveform or windowed variance in the waveform (Power et al., 2020),

which might provide an explanation for the failure of the physiological model to capture much variance in this study. There is also the option that a real relationship between respiration and the target brain process exists, and that the frontostriatal network is in fact controllable by changing respiration, or that another brain region controls both, the network and respiration.

It should be noted that our results are not in line with the report of Hellrung et al. (2018) who conclude that in activation-based rtfMRI NFB of the amygdala training effects are not mainly driven by physiological artefacts. It might be the case that our large-scale network NFB approach is specifically sensitive for the artefacts.

Taken together, in our analyses GSR was the single most effective method to correct for the undesired physiological associations that we detected in the group data. This fits to prior research emphasizing the ability of GSR to capture not only the activity of voxels located in the major clusters of the brain (Chen et al., 2012), but also reflects noise of different origins i.e. scanner driven-noise, motion, respiratory and heart-rate (Murphy & Fox, 2017).

According to Zarahn et al. (1997) inclusion of the global signal as a covariate decreases the effects of spatially related noise, which again allows for a better detection of effects in fMRI studies (Aguirre et al., 1997). Interestingly, after correlating the global signal with fMRI data, Birn et al. (2006) found that these maps were identical to maps of regions with signs of respiration. These correlations were particularly strong in gray matter regions elucidating the need to take this knowledge into account for fMRI analyses. GSR has also been suggested recently as a method to correct for noise of respiratory nature in modern multi-echo fMRI sequences (Power et al., 2018), although this has been debated (Power et al., 2019; Spreng et al., 2019).

Despite the ability of GSR for noise correction, it remains a controversial technique because it is unclear what the global signal in fact measures. As it is a hard problem to delineate low-frequency artefacts from neural effects, real signal could potentially be removed by the method (Hahamy et al., 2014). Another common remark about GSR is that it might induce spurious negative correlations (Saad et al., 2012), e.g. of task-negative and task-positive networks (Fox et al., 2009). However, similar negative correlations might also appear when only physiological noise correction is applied

(Chang & Glover, 2009) and some negative correlations might reflect real neural signal instead of artefacts (Chai et al., 2012).

However, it has to be noted that GSR may skew results in clinical populations. It was shown that the global signal is changed in SCZ, although it is unclear in which direction the changes occur, since opposing results were reported (Hahamy et al., 2014; Yang et al., 2014). While Caballero-Gaudes & Reynolds (2017) emphasize caution in the application of GSR to task-based activity or FC, it currently seems to be the most effective approach for eliminating global artefacts (Power et al., 2017) and shows superior output than different state-of-the-art correction methods, like e.g. ICA in terms of eliminating motion artefacts (Burgess et al., 2016; Liu et al., 2017).

So far it remains unclear how well GSR would work in rtfMRI NFB, and which correction method is providing the best online signal for learning to modulate brain networks. Thus, further research is needed to replicate the results and better understand the reported effects.

It is also crucial to mention that respiration is indeed truly associated with neural activity (Heck et al., 2016; Herrero et al., 2018; Ito et al., 2014) and the removal of the respiratory signal might thus also result in the loss of actual neural information. The reasons why we consider all physiological associations as artefacts here are that it is not clear how nuisance and real effects could be separated, and that rtfMRI is probably too costly to be the method of choice to measure and feedback signals that are strongly coupled to respiration, which could be assessed in a more direct and cheaper way.

Importantly, it was only possible to identify and examine the relationships between FC and physiology in this study because we recorded physiological parameters during fMRI scanning, otherwise the detected associations might have gone unnoticed. It thus follows as a clear recommendation for fMRI NFB research that physiological signals are recorded during scanning, and that the contamination of the target signal by physiological parameters is assessed and reported. To foster replication and further research on this topic we provide the scripts that we used to estimate physiological parameters as an open source script.

2.1.6 Conclusion

Our results suggest that it might be necessary to account for physiological artefacts in connectivity-based rtfMRI NFB, for example by applying online GSR. Failures in correction of physiological artefacts from online signals might lead to a confounded feedback which undermines the methodology of a study and challenges the validity of the conclusions. Given the massive impact physiological artefacts have on the BOLD signal, caution seems to be needed when interpreting the results of studies that do not use working physiology correction.

Funding

The study was partially financed by grants from the European Union's Horizon 2020 research and innovation programme (grant agreement No 668863 (SyBil-AA)), the Deutsche Forschungsgemeinschaft (Collaborative Research Center TRR 265-C04) and the German Federal Ministry of Education and Research/BMBF (01EA1605) to P.K. The funding sources had no involvement in study design, the collection, analysis and interpretation of data, and in the writing of the report.

Acknowledgements

The authors would like to thank Christian Sojer for recruitment of participants and Ellen Schmucker for MRI scanning.

Declaration of interest

Declarations of interest: none

3 INTERIM DISCUSSION

3.1 Challenges in rt-fMRI NF FC with frontostriatal networks

Study one was designed to implement the technology and test the feasibility of frontostriatal network rt-fMRI NF FC in a healthy population. Unfortunately, feasibility of the method could not be demonstrated sufficiently as strong methodological problems were detected. Specifically, severe respirational artefacts have to be faced when leaving the data uncontrolled. This did not come entirely unexpected, as already a substantial body of general fMRI FC literature is warning about the interpretation of uncorrected FC measures due to the ambiguity of the signal (Heunis et al., 2020; Power et al., 2012).

It was previously reported that the fMRI signal is containing several relevant noise signals (Caballero-Gaudes & Reynolds, 2017). Thermal noise, instrumental drifts, signal changes due to head motion and physiological fluctuations of non-neuronal origin such as cardiac and respiratory noise all affect the BOLD signal, and can be confounding results, for example if they are associated with the tasks (Caballero-Gaudes & Reynolds, 2017; Murphy & Fox, 2017). These non-neural fluctuations add variance in the signals extracted from distinct brain regions and can contaminate measures of FC by enlarging the putative FC if they affect different brain regions similarly (Caballero-Gaudes & Reynolds, 2017). Reducing this noise can be a demanding job, and several methods have been proposed to clean the data from different kinds of noise (Caballero-Gaudes & Reynolds, 2017). The approaches differ widely and e.g. can be focused on denoising motion-related noise and physiological-related noise, attempt to account for multiple kinds of noise sources including motion-related non-neuronal physiological noise such as by means of application of the average white matter (WM) and cerebrospinal fluid (CSF) as nuisance regressors, denoising based on multi-echo fMRI or phase-based denoising methods (Caballero-Gaudes & Reynolds, 2017).

Global signal regression (GSR) in which the average fMRI signal across all voxels in the brain is discarded is proposed to be a promising, but also disputed candidate for data correction (Caballero-Gaudes & Reynolds, 2017). The underlying hypothesis is that processes which are caught globally over the whole brain cannot represent

neuronal activity (Caballero-Gaudes & Reynolds, 2017). Hence, global signals rather equal confounding sources to the BOLD fMRI signal such as instrumental, motion-related and physiological fluctuations, particularly respiratory effects and should be removed from the data (Caballero-Gaudes & Reynolds, 2017; Power et al., 2017). Usage of GSR was first applied in PET studies (Aguirre et al., 1998), before being applied to task-based fMRI and resting state fMRI studies because of its' ability to demonstrate a more persistent and focal pattern of FC (Caballero-Gaudes & Reynolds, 2017).

Although it seems very straightforward to apply this approach, there have been multiple debates on the suitability of the method (Caballero-Gaudes & Reynolds, 2017). A commonly reported concern is that the method removes such a great portion of the signal that the remaining signal is too small to yield effects. Another point of criticism is that the global signal contains neuronal-related BOLD fluctuations as well and by removal of the global signal these would also be removed (Caballero-Gaudes & Reynolds, 2017). A negative bias in the approximated BOLD response is added to the data by application of GSR, meaning that positive BOLD responses are reduced and negative ones are falsely created (Aguirre et al., 1998; Caballero-Gaudes & Reynolds, 2017). The method drives the average correlations in the brain in the direction of zero (Caballero-Gaudes & Reynolds, 2017; Murphy et al., 2009), yet sometimes correlations are changed in an unforeseeable direction (Caballero-Gaudes & Reynolds, 2017; Saad et al., 2012). While authors still did not achieve consensus on whether its' use is recommended or not, there is agreement that its appropriateness and practicality depend on the dataset and research question (Caballero-Gaudes & Reynolds, 2017; Murphy & Fox, 2017). Since the accuracy of the NF signal is of crucial relevance in the conduct of rt-fMRI NF FC and the availability of alternative approaches for online data correction is scarce, GSR is - despite its potential flaws - recommended for use in the context of rt-fMRI NF FC (Weiss, Zamoscik, et al., 2020).

As mentioned earlier, an alternative for denoising is phase-based denoising. While the majority of fMRI studies apply magnitude-only denoising, leading to a redundant loss of efficiency (Calhoun et al., 2012), phase-based denoising is rather seldom (Caballero-Gaudes & Reynolds, 2017). A limitation of this method is that it is very sensitive to large-scale artefacts (bulk motion, respiration, scanner drifts) and more focal confounds (blood flow from large vessels) (Caballero-Gaudes & Reynolds, 2017).

Some independent component analysis approaches are also available that use both (phase- and magnitude-based sources) combined to improve mapping of the BOLD response (Calhoun et al., 2002). Among many solutions for the issues with phase-based denoising, phase variations due to physiological fluctuations can be eliminated by means of RETROICOR (Caballero-Gaudes & Reynolds, 2017) or k-space physiological correction methods like RETROKCOR which are both based on external physiological recordings (Caballero-Gaudes & Reynolds, 2017; Hu et al., 1995). Retrospectively, RETROICOR models cardiac and respiratory fluctuations as quasi-periodic processes by applying a low-order Fourier series with time-varying cardiac and respiratory phases that are matched with the data (Caballero-Gaudes & Reynolds, 2017). These are added to the GLM design matrix as nuisance regressors and are thus corrected (Caballero-Gaudes & Reynolds, 2017; Glover et al., 2000).

Importantly, insights that GSR is relevant in fMRI NF studies are far from novel. One of the first NF studies ever already applied GSR on their data to control for respiratory artefacts (Weiskopf et al., 2003). In 2008, Sitaram et al. (2008) mentioned the importance of correction for respiration and cardiac activity and criticized that this is not implemented online yet. A few years later, Koush et al. (2013) reported in one of their DCM papers that it was checked for differences in heart rate and respiration among the experimental groups. The absence of differences let the authors conclude that control of the feedback signal was not associated to cardio-respiratory artefacts. In their NF study, Yamashita et al. (2017) applied GSR, but did not find an effect of it. This evidence indicates that usage of GSR is - despite its' strong criticism as outlined in the first paper of this dissertation - an eminently necessary method.

The rather recent development of the model-based PhysIO toolbox that also employs RETROICOR (Kasper et al., 2017) emphasized the relevance of correction for artefacts in fMRI studies. Therefore, this method was applied to our data as a first step after discovery of the artefacts in our data set. Though as realized in our study, application of it can result in a false sense of security, because it can be used as an excuse to resign from further examination of whether artefacts were in fact reduced as a result of the use of the toolbox. On the other hand, methodological papers like this draw attention to the importance of correction of the data and are therefore essential.

Clarity on the encountered issues enables correction for those. In this respect, it was learned in the first study that respirational artefacts can be corrected offline by means

of model-based physiology correction (TAPAS) (Kasper et al., 2017) and GSR (Power et al., 2012). While the problems in offline data correction were eventually sufficiently solved, it is still an ongoing task to clean online FC data. Yet, based on the presented literature, GSR could also be a promising candidate in this regard too as it currently is the only available method for direct implementation and further tools for online correction of data are still under development (Misaki & Bodurka, 2021). Hence, study two of this dissertation applied online GSR and a few further adaptions with respect to the first study have been applied.

Aside to the data correction issue, a closer look in the literature revealed that the changes achievable within a single-session are very limited if possible at all (Rance et al., 2018). Therefore, we considered an extension of the NF training from a single session to three training days necessary for the second study. Additionally, it was realized as a result of the study that for demonstrating feasibility of FC NF training, more restricted neural networks should be preferred over complex networks that consist of several regions. Specifically, we restricted the network applied in the first study to the DLPFC and the striatum, thus excluding the ACC. One reason for this is that the ACC might have been a problematic target because of its error-monitoring function (Swick & Turken, 2002) that could drive FC of the frontostriatal network by reacting to failure in the training. Although the ACC has been a valuable target for regulation in previous rt-fMRI NF ROI-based approaches also in the context of SCZ (Cordes et al., 2015), we considered it fundamentally more important to test the controllability of a more simple and straightforward network which does not include abundantly many regions and decided for the DLPFC-striatum network. Notwithstanding, it will also be interesting and important to test an ACC-striatum network in a similar way in future studies.

In summary, the encountered issues with respiratory artefacts in the first paper of this dissertation were not entirely new, but they were advantageous in informing our second rt-fMRI NF FC study. Due to the severity of the detected associations and the enormous consequences for the data if the artefacts remain uncorrected, instead of applying the described approach to an SCZ sample, the second study aimed at solving the issues that appeared in study one again including HC. Based on our findings, in the second study we intended to rigidly control our offline and particularly online data

as they form the training signal by means of model-based physiology correction and GSR.

4 STUDY TWO: THE FEASIBILITY OF A LARGE-SCALE FMRI NETWORK FUNCTIONAL CONNECTIVITY NEUROFEEDBACK WITH IMPROVED METHODOLOGY

4.1 Feasibility of training the dorsolateral prefrontal-striatal network by real-time fMRI neurofeedback[3]

4.1.1 Abstract

Real-time fMRI neurofeedback (rt-fMRI NF) is a promising non-invasive technique that enables volitional control of usually covert brain processes. While most rt-fMRI NF studies so far have demonstrated the ability of the method to evoke changes in brain activity and improve symptoms of mental disorders, a recently evolving field is network-based functional connectivity (FC) rt-fMRI NF. However, FC rt-fMRI NF has methodological challenges such as respirational artefacts that could potentially bias the training if not controlled. In this randomized, double-blind, yoke-controlled, pre-registered FC rt-fMRI NF study with healthy participants (N=40) studied over three training days, we tested the feasibility of an FC rt-fMRI NF approach with online global signal regression (GSR) to control for physiological artefacts for up-regulation of connectivity in the dorsolateral prefrontal-striatal network. While our pre-registered null hypothesis significance tests failed to reach criterion, we estimated the FC training effect at a medium effect size at the end of the third training day after rigorous control of physiological artefacts in the offline data. This hints at the potential of FC rt-fMRI NF for the development of innovative transdiagnostic circuit-specific interventional approaches for mental disorders and the effect should now be confirmed in a well-powered study.

[3] Publication: Weiss, F., Zhang, J., Aslan, A., Kirsch, P., Gerchen, M.F. (2022). Feasibility of training the dorsolateral prefrontal-striatal network by real-time fMRI neurofeedback. *Sci Rep, 12*, 1669. https://doi.org/10.1038/s41598-022-05675-0

4.1.2 Introduction

Functional magnetic resonance imaging (fMRI) is a nowadays almost ubiquitous technique to study the brain and to gain information on alterations in brain functioning in mental disorders. With rising numbers of individuals affected by mental disorders in the recent past (Wu et al., 2021; Czeisler et al., 2020) and a relevant level of non-responders to current treatments, the need not only for basic scientific results but also for novel and innovative therapy approaches based on these insights is high. Orienting towards a more treatment-related usage, fMRI has been progressively used for the application of real-time fMRI neurofeedback (rt-fMRI NF) since its' introduction around the turn of the last century (Cox et al., 1995; deCharms et al., 2005; Weiskopf et al., 2003). In rt-fMRI NF participants are trained to volitionally control a brain process, which is usually not directly accessible, in a predefined direction. The brain process is selected based on its involvement in a mental disorder and evoked changes in activity are expected to be accompanied by an improvement in symptoms (deCharms et al., 2005). rt-fMRI NF specifically profits from the high spatial resolution and whole-brain coverage (Tursic et al., 2020) provided by fMRI. Although the temporal resolution in contrast is relatively low, it still allows for feedback in nearly real-time (Sulzer et al., 2013; Watanabe et al., 2017). As a further advantage of the method, compared to pharmacological treatments, no side effects have been reported (Tsuchiyagaito et al., 2021).

So far, the largest number of rt-fMRI NF studies have been conducted with Region-of-Interest (ROI) based approaches that feedback activity of a single ROI as the training target. Those studies could successfully demonstrate the feasibility of inducing changes in brain functioning (see for example (Kirsch et al., 2016; Zotev et al., 2011)) and in several cases associated improvement in symptoms could be demonstrated (Gerin et al., 2016; Mehler et al., 2018; Scheinost et al., 2013; Young et al., 2017).

Notwithstanding, a transition of the focus to connectivity-based approaches has begun in the past years. First reports suggest that changes in connectivity as a consequence of rtfMRI NF can be achieved (Zhao et al., 2019) and clinical measures might be improved (Bauer et al., 2020). Recently, Morgenroth et al. (2020) targeted DLPFC-ACC connectivity and found increased connectivity in the experimental group which correlated with symptom improvement in high trait anxiety after rt-fMRI NF. Further technological advances are Dynamic Causal Modelling (DCM)-based NF (Koush et al.,

2017; Koush et al., 2013) and whole-brain connectome-based NF (Scheinost et al., 2020).

Particularly for the development of transdiagnostic approaches addressing specific neural circuits that are involved in diverse mental disorders, like frontostriatal networks, connectivity-based rt-fMRI NF shows great promise. Before such approaches should be applied in clinical contexts, it is however necessary to address the methodological problems associated with the technology and gain a better understanding of the effects that can be expected.

In a previous single-session rt-fMRI NF study we found that FC rt-fMRI NF is heavily influenced by physiological artefacts, particularly from respiration (Weiss, Zamoscik, et al., 2020). In subsequent offline analyses of the acquired data we tested whether model-based physiology correction with the TAPAS PhysIO toolbox (Kasper et al., 2017), and global-signal regression (GSR) (Power et al., 2015) can eliminate the influence of these non-neural artefacts, and found that GSR is a promising approach for online physiology correction. GSR also has the further advantage that it can be implemented in a simple and straightforward manner during the estimation of the online feedback signal.

In the present double-blind randomized yoke-controlled pre-registered study we now tested in healthy controls over three training days whether an updated rt-fMRI NF approach with online GSR can be used to train participants to up-regulate FC in a bilateral frontostriatal network comprising the DLPFC and the striatum. With this study we aimed at demonstrating the principle feasibility of this approach, gain insight into the time course of the training effect, obtain basic effect size estimates and by this pave the way for future confirmatory studies and clinical applications of the developed technology to modulate frontostriatal circuitry in the diverse clinical conditions where they are involved.

4.1.3 Methods

4.1.3.1 Participants

40 healthy participants took part in this double-blind randomized yoke-controlled rt-fMRI NF experiment. Participants (24 female, 16 male) were between 19 and 30 years of age (mean: 23.28; SD: 2.39), did not present with any current or prior psychiatric diagnosis, had normal or corrected-to-normal vision, were free of a history of mental and neurological disorders and were not on acute psychopharmacological medication. Female participants were not pregnant. The study was approved by the Ethics Committee of the Medical Faculty Mannheim at the University of Heidelberg, Germany (2018-520N-MA) and complies with the World Medical Association's Declaration of Helsinki.

4.1.3.2 Pre-registration

The study was pre-registered at the Open Science Foundation (OSF NeCoSchi II https://osf.io/znrbk/). Specifically, we tested here the following two pre-registered hypotheses:

1. Averaged correlations between DLPFC and striatum are higher in the real neurofeedback group in comparison to the yoke control group during rt-fMRI neurofeedback sessions (directional).

2. Participants from the real neurofeedback group in contrast to participants from the yoke control group will demonstrate a higher increase of averaged correlations between DLPFC and striatum from the initial resting state period (directional).

4.1.3.3 Data/code availability statement

Due to the protection of sensitive data it is not possible to make the raw fMRI data publicly available. However, synopsis data that built the foundation of the reported analyses are accessible at the OSF project site (OSF NeCoSchi II https://osf.io/znrbk/).

4.1.3.4 MRI Scanning

MRI scanning was administered at a Siemens Biograph Scanner with 3T (Siemens Healthineers, Erlangen, Germany) at the Central Institute of Mental Health in Mannheim, Germany. MR images were acquired with a 32-channel head coil. T1-weighted structural images were obtained with a repetition time (TR) of 2 s, echo time (TE) = 2.58 ms, flip angle = 10°, 192 slices, slice thickness = 0.9 mm, voxel dimensions = 0.4 mm x 0.4 mm x 0.9 mm, FoV = 192 mm. Echo planar imaging (EPI) sequences were acquired with a TR of 1.64 s, TE = 30 ms, flip angle = 73°, 30 slices, slice thickness = 3 mm, voxel dimensions = 3 mm x 3 mm x 3 mm, FoV = 192 mm, GRAPPA factor 2. 343 Volumes were acquired and EPI sequence was the same for all functional runs. Physiological signals were measured with built-in equipment during functional scans.

4.1.3.5 Brain network definition

The rt-fMRI NF approach and all analyses focused on a predefined bilateral network including the dorsolateral prefrontal cortex (DLPFC) and the striatum (see Figure 1a). 22 ROIs in the DLPFC and 12 ROIs in the striatum were extracted from the cortical parcellations by Schaefer et al. (2018) and the striatal parcellation by Choi et al. (2012) which are both based on the 7-network cortex parcellation by Yeo et al. (2011). The DLPFC ROIs were identified based on an automatic metaanalysis with Neurosynth (https://neurosynth.org/; (Yarkoni et al., 2011) with the term “DLPFC”. We selected a broad measure of DLPFC-striatum FC as feedback target instead of targeting specific striatal subnetworks. Cortical regions have widespread projection fields in the striatum which enables them to influence other networks (Shipp, 2017), and the location of maximal connectivity is dynamically changing over the striatum (Gerchen, Weiss, et al., 2021). Thus, with our approach we are aiming at training the general ability of the DLPFC to exert control over striatal processes without focusing on a specific sub-network. However, our approach based on several ROIs within the target regions would allow for the identification of potential sub-networks linked with symptom changes in clinical studies. This could then facilitate refinement of the target networks. It is further important to note that while we used averaged connectivity as a rather simple network measure in this study our approach provides the technological basis

for future NF applications that could take more complex graph-theoretical network measures into account.

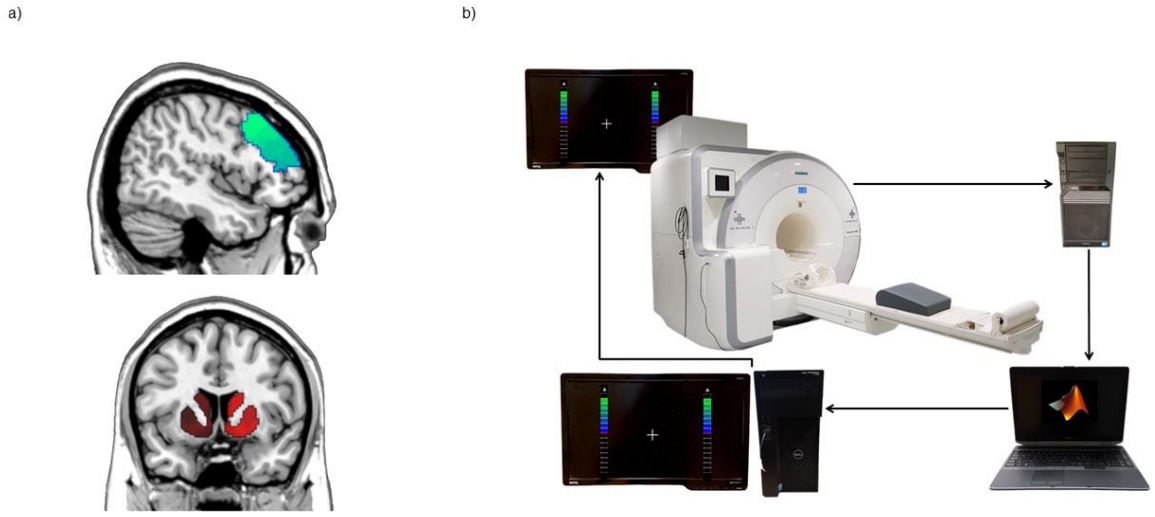


Figure 1. Experimental setup. a) DLPFC-striatum target network. Bilateral ROIs in the DLPFC and the striatum were predefined and projected into the individual anatomy of the participant to extract the online feedback signal during NF training. b) rt-fMRI NF setup. Images are sent to a laptop running in-house MATLAB scripts for pre-processing and extraction of the neurofeedback signal. The feedback signal represents the averaged functional connectivity between the ROIs in the DLPFC and the striatum. The feedback value is forwarded to a computer running Presentation software and is displayed in the scanner as a thermometer value that is continuously updated every TR.

4.1.3.6 rt-fMRI NF training

For a graphic representation of the NF-setup, please refer to Figure 1b. NF training was performed on three separate fMRI scanning days. At the start of the first day, demographic information was collected and questionnaires were answered. The questionnaires included the German Version of the Beck Depression Inventory (BDI-II) (Beck et al., 1996), the Schizotypal Personality Questionnaire (SPQ) (Raine, 1991), the 10 Item Big Five Inventory (BFI-10) (Rammstedt et al., 2012) and a sensory inventory (Zamoscik et al., 2017). The first and the third scanning day comprised a T1-weighted anatomical MPRAGE scan (5 min), an initial resting state scan, two NF runs and a transfer run with each one being of 9:29 min length. On the second training day the third run was also a NF run instead of a transfer run. The initial resting state run was conducted with open eyes while a fixation cross was displayed at the center of the

screen. The transfer run aims at testing for the generalizability of NF learning without a feedback signal and was in essence equal to the resting state run with the difference that participants were instructed to regulate their brain as in the NF run. The NF run included the presentation of a fixation cross which was located in the center of the screen. The fixation cross was surrounded by two thermometer bars that indicated the value to be up-regulated. The value was updated every TR and consisted of the averaged Z-transformation Pearson's correlation of DLPFC and striatal ROIs. Participants were not informed about specific mental strategies but were instructed to try out different strategies and were told to pursue the one they find most successful. Participants were randomly assigned to the real (N=20) or yoke control group (N=20) in a double-blind fashion by an automatic procedure implemented in the MATLAB code with a pre-specified randomization list. Participants in the yoke control group were paired with unique participant from the real group with a first-in-first-out procedure and received in each run the saved feedback signal of this participant from the same run. Data processing and handling were the same for both groups; just the sent feedback signal was automatically replaced in the yoke group. Thus, the staff was unaware of the group identity of the participants. The received feedback signal was consistent between the real and yoke control group in every run. The first three participants were deliberately allocated to the real group to ensure that sufficient recordings for the yoke procedure were available. After each scanning session participants rated their subjective performance, reported used strategies, and indicated which group they thought they belong to. At the end of the last session group allocation was disclosed.

4.1.3.7 Online data analysis

Online and offline analyses were conducted in MATLAB (R2019a, Math Works Inc., Sherborn, MA, USA). In-house software based on SPM functions (Gerchen et al., 2018; Weiss, Zamoscik, et al., 2020) was used for online rt-fMRI NF processing. During the resting state scan the anatomical image was segmented and normalized to the SPM 12 TPM MNI template. ROI masks were then projected into individual subject space. For rt-fMRI processing, during scanning every collected volume was immediately transferred to the laptop where the analysis was run. Each volume was realigned to the first volume of the series and averaged intensity values from all ROIs were extracted and added to the ROI signal time series. A general linear model (GLM) was estimated at every step over all acquired data to correct for movement parameters,

a cerebrospinal fluid (CSF) signal, spikes correlated with head movements (framewise displacement (FD) > 0.5mm) and the global signal. The last 15 volumes were used to calculate Fisher Z-transformed Pearson correlations (i.e. achieving a sliding window size of 15 volumes) between all ROIs in the DLPFC and all ROIs in the striatum. These were then averaged to obtain the online feedback signal. The first feedback value was presented after ~ 1 min (37 volumes = 60.68s). This delay was included to ensure that sufficient data for a stable estimation of the nuisance regression model was available. Only windows including a minimum of 10 volumes that were not influenced by head motions (FD < 0.5 mm) were used for the calculation of the feedback signal. If a window contained insufficient information the feedback value was kept constant.

4.1.3.8 Offline data analysis

SPM 12 (v6906) was used for offline data analysis. The anatomical image was segmented and normalized to SPM 12 TPM MNI space. We removed the first 10 volumes of the functional data. The images were slice-time corrected, realigned to the mean image, and co-registered to the anatomical image. The images were normalized, scaled to a resolution of 2 x 2 x 2 mm and smoothed with an isotropic Gaussian kernel of 6 mm full width at half maximum. In a first level General Linear Model (GLM) six movement parameters, the cerebrospinal fluid (CSF) signal, dummy regressors for volumes affected by head motion identified with the ART toolbox (framewise displacement > 0.5 mm; scan-to-scan global signal change z > 4, physiological nuisance regressors (see next paragraph)) and a constant were included. Runs with >20% movement-affected volumes were excluded from further analyses.

4.1.3.9 Physiological noise correction

A built-in respiration belt and a pulse finger clip (PMU Wireless Physio Control, Siemens Healthineers, Erlangen, Germany) were utilized for recording of respiration and heart rate during MRI scanning with a sampling rate of 400 Hz. To allow for evaluation of physiological parameters, physiological recordings were cut on the basis of recorded volume triggers for precise alignment with the fMRI data. Next, the TAPAS PhysIO Toolbox (Kasper et al., 2017) was applied for estimation of 20 physiological nuisance regressors, including heart rate variability (HRV), respiratory volume per time (RVT) and cardiac x respiratory interaction. Deduced physiology nuisance regressors were implicated in the first level GLM of the analyses.

4.1.3.9 Respiratory parameters

To assure that our results were not confounded by respiratory artifacts (Weiss, Zamoscik, et al., 2020) and demonstrate that GSR and model-based physiology correction worked efficiently in cleaning up the data, we further computed summarizing respiratory parameters from the time courses that are possibly related to the BOLD signal (Weiss, Zamoscik, et al., 2020; Zamoscik et al., 2018). Breath Rate that is defined as peaks/breaths per minute and Pause CV which is the coefficient of variance of respiration pause duration were calculated. For a more detailed description, see Weiss, Zamoscik, et al. (2020) and Zamoscik et al. (2018).

4.1.3.10 Offline connectivity estimates

Offline connectivity was estimated over the same averaged ROI-to-ROI connections as in the online approach, but over the whole available time course. For testing the second hypothesis the connectivity estimates of the initial resting state period of each day were subtracted from the respective estimates of the NF and transfer runs within each participant to normalize the modulation effect with respect to the individual baseline.

4.1.3.11 Second level analyses

Second level analyses were performed based on the DLPFC-striatal large-scale network connectivity values. For each NF and transfer run, connectivity estimates corrected for age and gender as covariates were compared between the two groups with one-sided independent samples t-tests implemented in a GLM model. Hedges'g and its confidence interval were estimated based on the obtained t-values to estimate the effect size per run (Gerchen, Kirsch, et al., 2021). Pearson's correlations were used to assess associations of offline and online connectivity with respiratory parameters.

4.1.4 Results

4.1.4.2 Functional connectivity group comparison

The randomized groups did not differ in terms of age ($t(38) = .1964$, $p = .8454$) and gender, $x^2(1, N = 40) = 1.667$, $p = .197$). Participants were not able to indicate above chance which group they were assigned to (training day1: $x^2(1, N = 39) = .205$, $p = .651$; training day2: $x^2(1, N = 40) = .404$, $p = .525$; training day3: $x^2(1, N = 40) = 1.616$, $p = .204$).

In line with our first pre-registered hypothesis, we investigated whether absolute averaged correlations between DLPFC and striatum were increased in the real group in comparison to the yoke control group. Here, no significant group differences could be found (see Supplementary Table 1). In accordance with our second pre-registered hypothesis we further investigated whether the real group would present with an increased relative FC in the NF runs normalized to individual baseline FC in the initial resting state run of the respective day. On a purely descriptive level, while at the start of the training FC in the target network was similar in the two groups, this began to change on the second training day (Figure 2). From the second NF run on day 2 on, the real group shows higher connectivity than the yoke group and this difference augments until the end of the training. We also did not find any significant group differences with the pre-registered null hypothesis significance tests at the specified criterion of $p<0.05$. In the second NF run on the third scanning day, which was the last of all conducted NF runs, significance testing led to a result of $t(33) = 1.5469$, $p = .0657$ (see Table 1) for an effect with a medium effect size (Hedges' $g = .5206$) that however had an accordingly large 90% confidence interval ranging from a very small to a medium effect (Figure 3).

The effect in the subsequent transfer run was also not significantly different in the two groups, ($t(32) = 1.2275$, $p = .1143$) and had a slightly smaller effect size of Hedges' $g = .4199$.

4.1.4.3 Physiological associations

We further explored the presence of remaining physiological associations after the online and offline processing procedures. In the offline data, over all NF and transfer runs, we identified one run in the real group in which FC was associated with respiration and one run in which the same applied to the yoke group (Pause CV for both) (See Supplementary Table 2). Unfortunately, more physiological associations were present in the online data. Within the real group we found four significant associations (Pause CV: 2, Breath Rate: 2) and the same was true for the yoke group. For further details, please refer to the supplement (Supplementary Table 3). It is important to note here that our effect size estimates are based on the offline data, and that no physiological associations were present in run 2 of day 3 in which we detected moderate evidence against the null (offline association over whole sample: Breath rate: rho = 0.043, p = .803; Pause CV: rho = -.0182, p = .287, see Supplementary Figure 3 for more details).

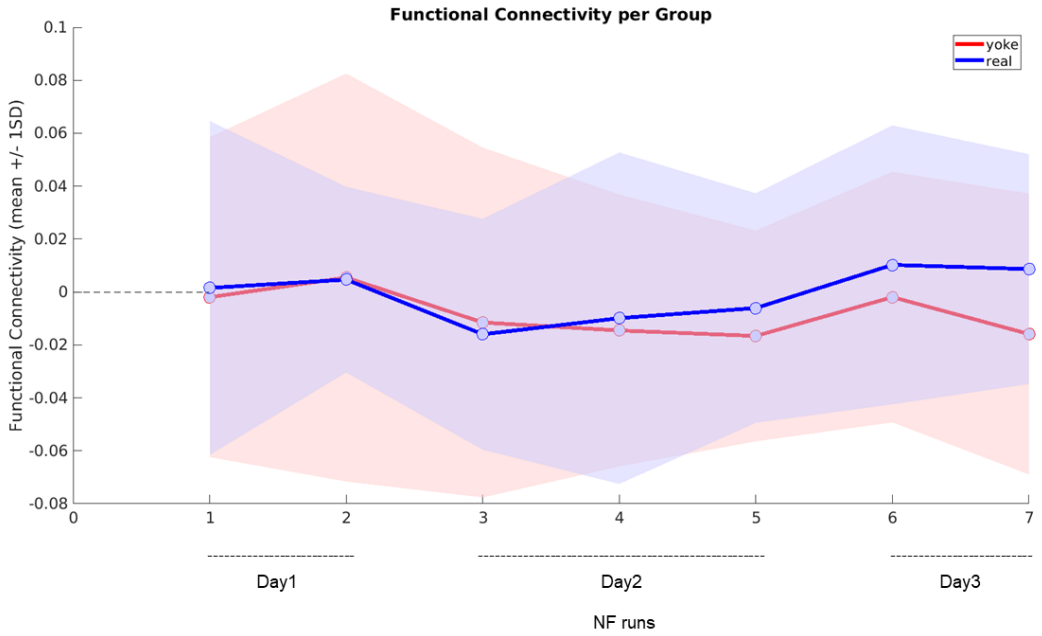


Figure 2. Functional connectivity over runs. Functional connectivity per group during NFB runs normalized by initial resting state FC of the respective day and corrected for age and gender is displayed (blue = real feedback group, red = yoke feedback group). Shaded areas represent ± 1SD from the group mean. 7 NF training runs were conducted over three training days. A moderate group difference was found during NF run 7 (the last NF run of day3).

Runs	**Group Differences Functional Connectivity**
Day1_NF1	t(31) = 0.1685, p = 0.43365
Day1_NF2	t(30) = -0.0370, p = 0.48535
Day1_transfer	t(31) = -0.5548, p = 0.2915
Day2_NF1	t(35) = -0.2495, p = 0.4022
Day2_NF2	t(31) = 0.2456, p = 0.4038
Day2_NF3	t(33) = 0.7689, p = 0.2237
Day3_NF1	t(33) = 0.7515, p = 0.2288
Day3_NF2	t(33) = 1.5469, p = 0.0657
Day3_transfer	t(32) = 1.2275, p = 0.1143

Table 1. Group comparisons of DLPFC-striatum functional connectivity per run corrected for age and gender as covariates. Functional connectivity was normalized by initial resting state activity of the respective day.

4.1.5 Discussion

In this study we tested the feasibility of functional connectivity NF training in a large-scale DLPFC-striatal network and estimated a medium effect size for the difference between the experimental (real) and control (yoke) group at the end of the third training day. Throughout the study we aimed at rigorously controlled experimental procedures including pre-registration, online physiology correction, double-blind randomization, and a yoke control group.

Unfortunately, our pre-registered null hypothesis significance tests failed to reach significance at the specified criterion of $p<0.05$, and thus do not provide clear evidence against the null hypothesis. This is likely due to our study being clearly underpowered. When we planned the study we were not having a good estimate of the expected effect size and thus planned the study with the minimum reasonable sample size with the intention to obtain effect size estimates from the study. It is important to note that it is still rather common for current NF studies to be underpowered (Tursic et al., 2020). While our sample size of N=40 in two groups is small, it is not one of the smallest in the field of rt-fMRI NF (see for example (Canterberry et al., 2013; deCharms et al., 2005; Ruiz et al., 2013)). To address this issue in future research it is important to have realistic estimates of effect sizes in rt-fMRI NF (Thibault et al., 2018). Our medium effect size of Hedges' g = .5206 is similar to the estimation of Dudek & Dodell-Feder (2021) who found medium effect sizes (Hedges' g = .59) for regulation over all included NF studies and Tursic et al. (2020) who showed that an effect size of 0.73 (Cohen's d) can be detected with a power of 95% based on ROI activation and connectivity regulation NF studies.

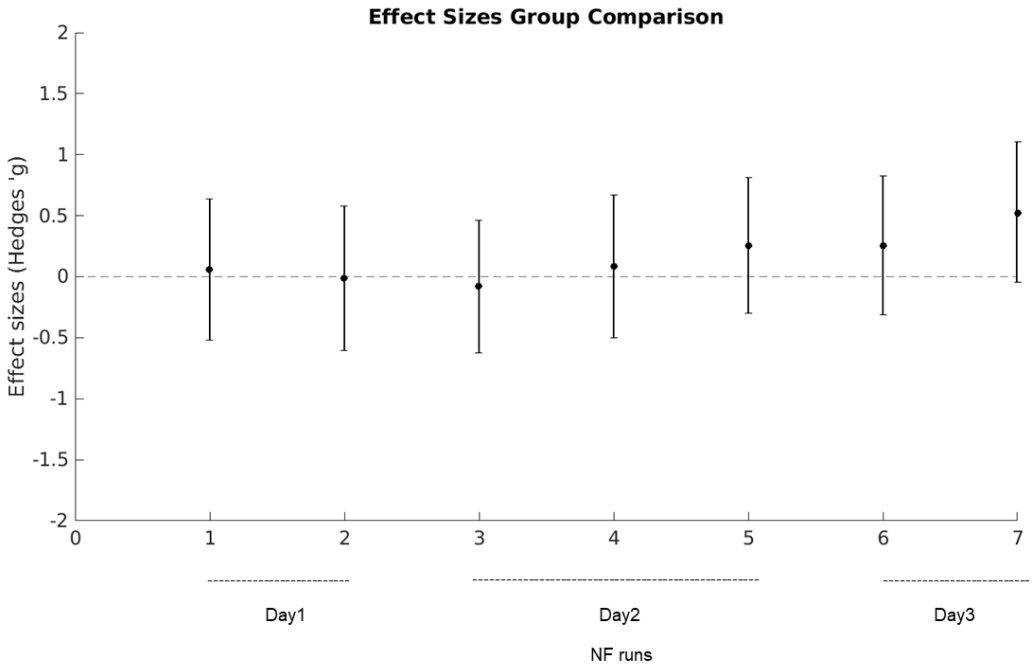


Figure 3. Effect sizes of the group comparison. Effect sizes (Hedges' g with 90% confidence interval) of the group comparison testing for differences in DLPFC-striatum FC between the real feedback group and the yoke control group for each NF run. An effect of g = .5206 was found in NF run 7 at day 3.

The time course of the NF training effect in our data looks promising. Increased FC in the experimental group in contrast to the yoke group is on a purely descriptive level visible from the second NF training run of day 2 and then increases until the end of the third day.

The trend of this time course suggests that the effect might even increase when additional training days would be added. Our findings fit into the framework of a growing number of methodological papers that investigate optimal conditions for rt-fMRI NF (Fede et al., 2020; Heunis et al., 2020; Sorger et al., 2019) and are in line with Morgenroth et al. (2020) and Rance et al. (2018) which demonstrate that NF training effects begin to appear after two training days. Along these lines, in an extensive NF study including 12 training days, Auer et al. (2015) found an increase in training efficiency as far as the 3rd-4th day. Thus, single-session NF training might not be sufficient, although some studies find effects in just one training session (Kohl et al., 2019; MacDuffie et al., 2018; Tsuchiyagaito et al., 2021).

On these grounds, our effect size estimated prepare the ground for an adequately powered confirmatory study to validate the time course and the effect size of the

training effect, which should include at least one further training day. A reason for the non-significant results despite a medium effect size and a decent sample size might be a high signal-to-noise ratio in the data. One potential source of interindividual noise might be the use of a yoke control procedure together with a target process involved in executive control. It could be that perceived group assignment or the dynamics of the feedback signal might have had an influence on the target process in our study, especially in the yoke control group.

Importantly, when comparing NF studies, a differentiation should be made between activation NF and FC NF as the latter is a more complex signal and this probably harder to train.

(Online) control of confounding noise sources is essential for FC rt-fMRI NF (Heunis et al., 2020; Weiss, Zamoscik, et al., 2020). Importantly, unintended sources should optimally be corrected in the data already in the online analyses, because of their potential to bias the training if the feedback value is confounded. In our previous study we identified GSR as a promising approach to clean up the online signal (Weiss, Zamoscik, et al., 2020), and implemented it in the present experiment. GSR is however a controversial method (Murphy et al., 2009; Saad et al., 2012) and has been criticized for being too rigid potentially removing real signal along with noise (Hahamy et al., 2014) and might thus reduce the power of the study, i.e. the probability to uncover real NF effects. On the other side, GSR is one of the most efficient methods for correction of global artefacts and has been recommended for correction of respiratory noise, for example in modern multiband sequences (Power et al., 2018; Power et al., 2017).

Unfortunately, despite the application of online GSR, in some runs associations of the online signal with respiratory measures were still present in the data (see Supplementary Table 3). For additional investigation of physiology correction and a comparison between artefactual physiological artefacts offline and online, we conducted analyses without specific physiology correction (without GSR and Physio) and GSR only (without PhysIO) (see Supplementary Tables 4&5). These findings replicate the findings in the preceding paper (Weiss, Zamoscik, et al., 2020). Most importantly, the offline analyses without GSR and PhysIO correction identified substantial physiological associations in the vast majority of runs (21/28 “corrupt” runs), which is much higher than the number of associations in the online training data. This clearly demonstrates that our online correction machinery with GSR was more effective

in cleaning the data than the usual offline analysis without further correction. GSR thus seems to have a strong incremental value despite its controversial aspects and can be recommended for online use, although it is slightly less effective online. Regarding more insights into the relationship of the different processing strategies, an exemplary network connectivity time course with the different processing strategies is shown (see Supplementary Figure 4). This figure demonstrates the strong influence of GSR on the network connectivity estimate and why it has such a strong influence on the detection of physiological associations. While GSR might be a helpful tool in addressing physiological artefacts, it does not solve the issue completely and further improvements are necessary, for example by online implementation of model-based physiology correction algorithms using simultaneously acquired physiological signals (Misaki & Bodurka, 2021). Such approaches are however technically much more demanding than simple GSR and are not available yet.

It is important to note that in the offline analyses we additionally conducted model-based physiology correction with the TAPAS PhysIO toolbox (Kasper et al., 2017) before we calculated FC estimates. After this additional correction, associations with respiratory measures were widely diminished (Supplementary Table 2), but the reported evidence of a NF training effect was still present. Thus, GSR alone might not already be the optimal method to control online physiological artefacts, but our NF approach seems at least sufficient for generating evidence of FC training effects beyond physiological artefacts.

A further methodological rigor of our experiment is the use of a double-blind yoke-controlled design. A majority of rt-fMRI NF studies is conducted single-blind (see for example (Mehler et al., 2018; Papoutsi et al., 2020; Zilverstand et al., 2017) and double-blind designs are still rare, although they are highly recommended, for example in the CRED-nf protocol for neurofeedback studies (Ros et al., 2020).

While the recommendation for a double-blind procedure is unambiguous, several different control procedures for NF studies are available. It is for example possible to use computer-generated sham feedback in the control group (Mayeli et al., 2020) or employ a within-subjects design in which participants received real feedback in the first session and control feedback in the second session (Bauer et al., 2020). Our selected yoke-control procedure guaranteed that all facets of the conditions apart from control over the ROI signal were matched (Sorger et al., 2019). An important aspect of the

interpretation of results from a yoke-controlled design is whether participants can accurately indicate which group they were assigned to. If participants in the yoke control group are able to correctly identify themselves, this could cause frustration and influence performance. Accordingly, this could artificially inflate group differences. However, as our participants were not able to accurately guess their group identity, this was likely not a problem in our study.

We used a design with continuous NF regulation over whole runs because this corresponded well with the requirements of the FC-based feedback measure. However, the task was demanding and fatigue might have prevented a better performance. It remains open whether a block design with alternating NF and rest blocks would have provided a better training outcome.

We estimated FC based on a sliding window of 15 volumes (24.6s). This is a window size within the normal range in the field of FC NF but can introduce a substantial delay in the feedback signal. We chose this window size because it covers the full waveform of the canonical hemodynamic response function, leaves sufficient data for estimating connectivity even when several volumes are censored, and corresponds to the continuous nature of our training. It can however not be excluded that a shorter window might have provided better learning and future research is needed to empirically address the influence of the window size on NF learning.

Of note, the FC values were normalized by the individual baseline FC in the initial resting state scan of the respective day (hypothesis two). Instead of taking the resting state scan as the baseline it would be another possibility to conduct the transfer run before the NF runs and use this pre-training transfer as baseline. We conclude that it is important and particularly more sensitive to take an individual baseline into account when calculating group differences in FC.

In comparison to the last NF run on training day three, the effect size was smaller in the transfer run without a feedback signal (Supplementary Figures 1 and 2) again suggesting that additional training runs might improve the effects. Nonetheless, our findings might cautiously be interpreted in a way that successful regulation in the transfer task could potentially be achieved after further training.

The patterns in our data, if confirmed, would underline the potential of FC rt-fMRI NF to induce actual changes in FC beyond physiological artefacts and thus provide various

options to develop innovative and transdiagnostic treatment approaches for different mental disorders sharing common neural features like frontostriatal dysconnectivity. However, it is still to demonstrate that changes in FC also lead to changes in behavioral or disorder associated alterations. Given the relatively high scanning costs, rtfMRI NF is however a rather expensive method. On the other hand, rt-fMRI NF is completely non-invasive, is able to address very specific and complex phenotypes in the brain, and has the potential to evoke changes in FC (Long et al., 2021). To avoid unnecessary costs, it will be important to identify predictors of successful NF training (see for example Scheinost et al. (2013); Zhao et al. (2021)) to inform precise individualized treatment approaches.

4.1.6 Conclusion

Our findings extend the hitherto thin but increasing literature on connectivity fMRI NF studies (Megumi et al., 2015; Morgenroth et al., 2020; Tsuchiyagaito et al., 2021; Young et al., 2018) by adding effect size estimates for NF modulations of complex fMRI signals. The moderate effects could only be seen after extensive training and several FC rt-fMRI NF training sessions. Extra caution is needed in controlling the online target signal for artefacts. Overall, our study supports the further exploration of FC-based rt-fMRI NF as a contingently promising method to develop circuit-specific treatment approaches for mental disorders in adequately powered confirmatory studies.

Data availability statement

The data that support the findings of this study are available in the Open Science Foundation (OSF) repository, at: https://osf.io/znrbk/. Supplementary data to this article can be found online at https://doi.org/10.1038/ s41598-022-05675-0.

Acknowledgements

The work was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 402170461 – TRR 265 (C04-project) to P.K. and Falk Kiefer (Heinz et al., 2020). The authors thank Mario Ratzel, Benjamin König and Cagatay Gürsoy for assisting in the fMRI scanning sessions, Sophia Zimmer for recruitment of participants and Elisabeth Horschmann for supporting graphic works.

Author contributions

F.W., P.K., M.F.G. designed the study. M.F.G. developed the rt-fMRI setup. F.W., M.F.G., J.Z. were involved in the data acquisition. F.W., M.F.G., P.K. wrote the manuscript. A.A. revised the manuscript. All authors read and approved the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

5 GENERAL DISCUSSION

The research in this dissertation has shown that rt-fMRI NF FC for large-scale brain networks is feasible and suggests a high clinical utility, also because of its potential to be applied transdiagnostically. However, it was found that in the conduct of rt-fMRI NF FC, several more or less expectable factors with strong impact need to be considered for successful implementation. Rigid data correction for artefacts and sufficient thoughts on duration of the training play a key role here.

To test the feasibility of novel large-scale NF FC approaches that were only recently developed, study one investigated the general feasibility of rt-fMRI NF training of a large-scale frontostriatal network in a single session. The approach presented with an increased degree of complexity and so one of the challenges that appeared was to control for respiratory artefacts sufficiently online and offline and to perform a training of sufficient duration to allow for training effects. Accordingly, study two was informed by the findings of the first study and replicated the approach with online and offline GSR data correction and an extended training duration.

5.1 Summary of study results

5.1.1 Study one: The feasibility of large-scale network regulation by means of rt-fMRI NF FC

In study one the feasibility of single-session rt-fMRI NF training of a large-scale frontostriatal network was assessed. The study was pre-registered, double-blind and randomized with a yoke-controlled design. While the main hypothesis, the demonstration of an NF effect could not be confirmed, it was found that FC measures in NF training are subject to strong respiratory artefacts. Retrospectively, this finding should not have been completely unexpected as in the general fMRI literature several reports warned of the risk of artefacts in FC-based measures (Heunis et al., 2020; Power et al., 2012; Van Dijk et al., 2012). For rt-fMRI NF it was however not very well documented and its consequences especially for the online training were not explicitly discussed. Regarding offline data correction GSR and the TAPAS PhysIO Toolbox qualify for use (Kasper et al., 2017). These can potentially also be applied online.

Application of GSR to the offline data is a rather straightforward and easy-to-implement way of data correction with frequent application in the past (Caballero-Gaudes & Reynolds, 2017; Weiskopf et al., 2003; Yamashita et al., 2017). However, its usage is heavily discussed i.e. because of its potential to induce anticorrelations or because of a putatively decreased ability to demonstrate effects.

In particular, it has to be noted that of course not only the offline data might be confounded by said artefacts, but especially the online data which is even more unintended as they build the basis of the feedback signal and therefore of the whole training. If the training signal is biased by respiration, then respiration in contrast to the FC measures is regulated which is what was found in the first study. This then also implies that offline corrections are worthless in the event of a biased online signal. It was therefore suggested to also apply online data correction in future studies which was likewise recommended by (Misaki & Bodurka, 2021).

Taken together, this study provided a very important insight on the conduct of rt-fMRI NF FC which is the necessity of online and offline data correction, existing problems namely that the NF training can be biased by respiratory artefacts and more importantly a promising solution which is the application of GSR and the PhysIO Toolbox, as the combination of these two was found to eliminate the influence of the artefacts on the training signal.

5.1.2 Study two: The feasibility of large-scale network regulation by means of rt-fMRI NF FC with improved methodology

Study two of this dissertation applied an additionally elaborated method that consulted the findings of the first study. While this study was also conducted randomized, double-blind, yoke-controlled and was pre-registered prior to the start, the data were GSR corrected online as suggested in the first paper. As we did not find a training effect in only one session and because literature suggests multiple sessions (Fede et al., 2020; Rance et al., 2018; Yamashita et al., 2017), the training was extended to three training sessions. Because it is potentially more straightforward to regulate connectivity of two as opposed to three brain areas, the large-scale network was reduced to DLPFC and striatum only.

It was found that with optimized data correction promising learning effects of rt-fMRI NF training could be visualized but that the effects were not strong enough to achieve significance.

Calculation of effect sizes for the different runs demonstrated a medium effect size in the run with the most pronounced group difference. This is in line with reviews by Fede et al. (2020) and Tursic et al. (2020) that demonstrate on average medium effect sizes in rt-fMRI NF studies. Our results further showed that even after online correction for physiology by means of GSR, a few associations with respiration were still present in the online training signal, although the amount was decreased in comparison to after offline correction. Remarkably, despite remaining online artefacts, our training resulted in a medium offline effect which survived maximal physiology correction.

Based on the visualization of the learning curves of the two groups, which in the real group looked like one would expect it from a functioning training, the results also suggest that a further extension of the training days would be strongly advised to deliver significant changes. In a review discussing the number of training sessions it was found that most rt-fMRI NF studies applied two sessions (Fede et al., 2020) which also corresponds with previous recommendations (Auer et al., 2015; Rance et al., 2018). Importantly, discussions regarding the optimal number of NF sessions did not distinguish between ROI and FC approaches, although the latter due to an increased complexity might also require more time.

In sum, study two highlighted the great potential of rt-fMRI NF to induce changes in large-scale frontostriatal networks and suggests further use of the approach as a clinical treatment tool with optimized methodology. This implies an increased amount of training session to allow for training effects and the conduct of power analyses to enable reaching predefined effect sizes.

5.2 Implications of the studies

As mentioned earlier, an important insight from these two studies is that the state of research is still at the start of establishing useful rt-fMRI NF FC. Despite the novelty of the method, we demanded high standards of the quality and rigor of the two studies of this dissertation.

The studies in this dissertation have been pre-registered prior to the start of the recruitment. A sample size of N=40 was applied in both of our studies and the experiments included a state-of-the-art automatic double-blind randomized yoke control procedure. This seems specifically noteworthy as lack of study preregistration to increase transparency in the conduct, small sample sizes, shortage of control conditions, lack of randomized treatment allocation and blinding were listed as serious remarks on the applied best practice standards in rt-fMRI NF (Melnikov, 2021; Trambaiolli et al., 2021). These aspects align with general fMRI study weaknesses (Poldrack, 2012). Regarding sample sizes it was highlighted before that current NF studies are rather small with median sample size of N=20 (Fede et al., 2020).

The necessity of the addition of effect sizes has been mentioned before. Sitaram et al., (2017) claimed that due to the still initial development of ROI-based NF training, there are many aspects which are unclear yet and fundamentally criticized the absence of expectations regarding effect sizes which was also noted for fMRI studies in general (Ioannidis, 2005; Poldrack, 2012). In this regard, measures of effect sizes for each run in our rt-fMRI NF FC training were provided in the second study. As outlined earlier, it was shown that medium effect sizes can be detected in rt-fMRI NF studies which aligns with previous research (Fede et al., 2020; Tursic et al., 2020).

Summing up, in the studies we aimed at addressing important methodological topics such as open science practices and controlling for a multitude of biases such as experimenter bias to assure interpretability of results (Poldrack, 2012; Sitaram et al., 2017).

Rt-fMRI NF has been applied in a multitude of different study populations, clinical and healthy participants. Both of the studies in this dissertation studied healthy controls. The field is overrepresented with pilot trials since 64% of all current rt-fMRI NF studies include healthy volunteers only (Fede et al., 2020; Melnikov, 2021). However, an argument for the usage of healthy controls as a target group is that a NF approach should be tested in terms of feasibility in a healthy population first, before it is applied in more sensitive clinical populations that also present with higher recruitment costs.

Although the two studies in this dissertation were conducted in healthy participants, our specific method was developed to be further used in clinical populations. Therefore, it is relevant to consider whether efficacy is distinct for different mental disorders

(Sitaram et al., 2017). NF findings on the same psychiatric disorder investigated in different labs diverge strongly (Melnikov, 2021). In MDD for example, a meta-analysis by Trambaiolli et al. (2021) demonstrated good face validity with significant clinical changes and cognitive and neural improvements as a consequence of rt-fMRI NF training, while Melnikov (2021) found only some efficacy. Further research is needed to investigate the clinical efficacy in different subcategories such as SCZ, since evidence for the efficacy of rt-fMRI NF in individual mental disorders is thin.

However, nowadays an increasing amount of research work is done in the direction of optimizing comparability of various NF studies and methodological rigor is increasingly addressed which was previously mentioned as a general weakness of fMRI studies (Poldrack, 2012). As an example, statistical power of fMRI studies is elevated. Excellent reviews (Thibault et al., 2018), meta analyses summarizing the majority of NF studies (Fede et al., 2020; Tursic et al., 2020) and methodological papers (Heunis et al., 2020; Sorger et al., 2019) are increasingly available and a union of NF researchers published the “Consensus on the reporting and experimental design of clinical and cognitive –behavioral neurofeedback studies (CRED-NF)” list that draws attention to important points to address in NF research and suggests how methodology in future NF studies could be improved (Ros et al., 2020). While the points listed there such as preregistration of experimental protocols have been increasingly addressed in general fMRI studies over the last years, increasing awareness for these aspects within the field of NF will support addressing the issues sufficiently also in this domain. Due to the stated increasing methodological focus within the field of rt-fMRI NF (see for example (Heunis et al., 2020; Ros et al., 2020)), one can confidently say that the field is heading in a good direction to eventually realize the clinical utility of rt-fMRI NF studies. Nevertheless, future rt-fMRI NF studies should increase methodological rigor (Heunis et al., 2020).

Pharmaceutical treatments possess a probability of evoking side effects, but for non-invasive approaches such as rt-fMRI NF this is not as strongly expected. However, generally interventions that have an effect must also present with some sort of side effects. Hence, if a NF approach is in fact working, side effects should also be possible e.g. by overtraining a network. Therefore, it is recommended to investigate this. In this regard, no side effects were reported in the two studies of this dissertation. This is in agreement with another rt-fMRI NF study (Tsuchiyagaito et al., 2021) and yields

information on the previously proposed question whether side effects as a consequence of rt-fMRI NF are likely (Sitaram et al., 2017). Nevertheless, the possibility exists that previous NF studies did not examine this with sufficient rigor.

5.3 A framework for well-controlled rtfMRI NF

It was mentioned previously that rt-fMRI NF is still under development especially with regard to optimal practices. Based on the crucial relevance of methodological aspects of study results, an optimization of the methodology in the form of an agenda including aspects which improve the quality of the work is required. The agenda should support in the process of designing the rt-fMRI NF study, help in the conduct but also facilitate the steps after the data collection. Notably, this is not a complete list of points of consideration with regard to rt-fMRI NF study designs but a selection of important aspects based on the findings of this dissertation which should serve as an add-on to other quality checks (Ros et al., 2020; Sorger et al., 2019). To distinguish it from the CRED-NF checklist (Ros et al., 2020), the focus of the proposed procedure is on optimal conduct of rt-fMRI NF studies but on a more fine-scale level emphasizing details we found to be important rather than an overview on all imaginable aspects.

Generally, the issues that should be carefully addressed in future FC NF studies as documented in the agenda can be broken down into two domains: Artefact correction of data and power.

Regarding the first topic, carefully for respiration corrected offline data eventually allow for better interpretability of the findings and rule out the remaining risk of artefact-based effects (Caballero-Gaudes & Reynolds, 2017). Therefore, rt-fMRI NF studies should carefully correct the online and offline data for global artefacts and artefacts due to respiration and cardiac activity. It was found that GSR and the TAPAS PhysIO Toolbox are suitable candidates for the implementation of this offline. Aside GSR, further approaches for application online are in development (Misaki & Bodurka, 2021). Of note, the associations could only be explored due to the conducted physiology recordings, so future rt-fMRI NF FC studies should clearly consider recording peripheral physiological signals.

In terms of power, the acquired knowledge about necessary effect sizes can be used to design adequately powered future rt-fMRI NF studies that account for those effect sizes. So far, this was relinquished for the larger part in previous rt-fMRI NF studies,

but it should be applied in future studies to improve scientific quality in the field and increase sample sizes (Heunis et al., 2020; Sitaram et al., 2017). Of note, for an intervention to be effective it should present with a rather strong effect, otherwise it is not worth it. However, effects can also manifest strongly intraindividually, although not being significant on the group level and it needs to be emphasized that these two should not be mixed up.

With regard to study design, clear hypotheses on how the networks are associated with behavioral functioning should be available prior to the start of the experiment. Aspects like this should be recorded in writing in the format of a pre-registration which is published e.g. on an appropriate portal and which is adhered to in the conduct of the study and later analysis of data. Adding to the network aspect, it is further advantageous if there is an understanding about accompanying brain processes during execution of NF.

Apart from our main findings, a few other points of improvement in the conduct of rt-fMRI NF FC were salient. Many studies reported that participants exist who generally lack the ability of voluntarily controlling their brains (Fede et al., 2020; Hampson et al., 2012; Sitaram et al., 2017). It was also found that successful NF performance is predictable on the basis of resting state activity measures (Scheinost et al., 2014) and anatomical landmarks i.e. putamen volume (Zhao et al., 2021). In the case of successful replication of these studies, these two aspects could be rationally combined by carefully (pre-)testing participants for their ability regarding NF training using predictors such as the latter mentioned ones. By that, future NF studies could reconsider excluding individuals in which no NF effect is expected. Of note, this could prevent unfounded usage of monetary resources such as scanning costs and thereby maximize efficacy and economical use. A further approach aiming at the latter goal would be to investigate why rt-fMRI NF is not working for some participants and how the amount of responders could be increased.

In summary, all of the previously mentioned aspects (i.e. pre-registration, estimation of effect sizes, careful power analyses, a clear target network, physiology recordings, rigid data correction online and offline and usage of predictors) can build the content of an agenda to optimize rt-fMRI NF. For a visualization of the agenda, please refer to Figure 1.

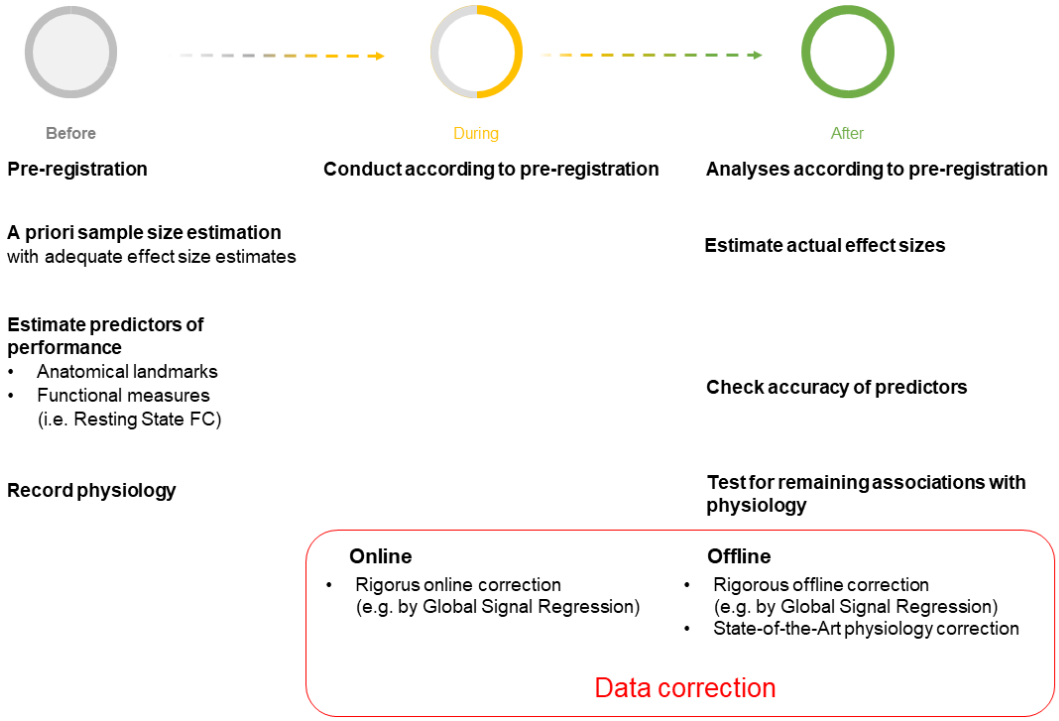


Figure 1. Points of consideration during the course of a real-time fMRI neurofeedback experiment. The left column displays important aspects when designing the experiment. The middle column shows aspects of relevance during the experiment and the right column outlines important steps after the experiment. Note: FC indicates functional connectivity.

A few further ideas of improvement are proposed that go beyond the scope of that proposal.

As one of the weaknesses in rt-fMRI NF lack of replication has been described (Melnikov, 2021) which is expressed in a huge variance of methodological implementations (e.g. different kinds of feedback were used, control groups differed etc.) and decreased comparability. Along these lines, outcomes can either be behaviorally-based or based on changes in brain functioning. These are two very different levels. Aiming at increasing clarity about available results in a faster way and to prevent misunderstandings, future rt-fMRI NF studies should clearly state in the abstract and if possible, also in the title of the research paper that results include “neural changes”, “behavioral changes” or “neural and behavioral changes”.

Effects in the experimental group are strongly desired in experimental research, but sometimes undesirable effects occur in the control group as well. It is suggested that placebo effects in a sham control group are a highly relevant point of criticism in rt-fMRI NF. Some participants are very likely to try all available resources in terms of

brain activity to match the content displayed. A study reported neural changes in the desired direction in the control group although participants were unaware of the actual feedback (Paret et al., 2014). A possible interpretation could be that NF might not be required for regulation of this particular brain area (amygdala). Another study also reported an improvement on the behavioral level in the control group which could be driven by placebo effects such as by the additional attention participants receive as a result of participation (deCharms et al., 2005). While an improvement in symptoms will be generally appreciated by researchers, this is less so in the context of control groups. However, to eliminate the issue of placebo effects due to additional attention or an increased effort and to disentangle them from pure NF effects remains a challenge, but drawing attention to this topic and mentioning it in protocols where it is relevant is a necessary first step which should be undertaken.

To summarize, there are many options to possibly improve rt-fMRI NF conduct e.g. by optimizing data correction methods for artefacts and increasing sample sizes and one should not hesitate to address these.

5.4 Limitations and implications for future studies

Rt-fMRI NF presents with many strengths such as a high spatial resolution and allows for unprecedented options e.g. the visualization of brain activity applying this high spatial resolution, notably, it also has some drawbacks. As mentioned previously, a main limitation of the two studies in this dissertation is that no power analyses have been conducted beforehand to guarantee sufficient sample sizes. This was due to the lack of realistic effect size estimates that could have been used for the power analyses. To provide a realistic chance to find effects, power analyses are clearly needed. Future studies should thus conduct power analyses in advance and by that sample sizes are likely to increase. We have also found that online GSR is still not optimal in clearing the data from respiratory artefacts, but no better tool is currently available for correction of the data. Until no (better) techniques for online correction are available, it is recommended to apply online GSR in future rt-fMRI NF FC studies.

Although we already extended the training from a single training session to three training sessions in the second study, we did not consider that FC NF training is a very specific and probably very complex process which could even require an extended

training in contrast to simple ROI-based approaches. Therefore, future rt-fMRI NF FC studies should consider a minimum of three training sessions.

One characteristic of proof-of-concept/feasibility studies is to include healthy controls in the sample to save more rarely available resources such as diagnosed patients for more elaborate studies (Fede et al., 2020; Yoo et al., 2012). Therefore, the number of studies including healthy participants is enhanced (Fede et al., 2020). However, it is worth questioning whether healthy controls provide a suitable model for those (actual) patients. It is possible that healthy controls do not react as strongly (or not at all) to e.g. symptom provoking pictures that are sometimes used for the conduct of the rt-fMRI NF FC experiment. Another option is that due to a ceiling effect participants cannot be differentiated from each other or that the pathological process is completely absent in HC. On the other hand, the potential use of maladaptive strategies would be less harmful for HC than diagnosed patients. Furthermore, the more complex a network or target process is, the harder it might be to regulate it. For this reason, future studies are advised, as one of the first steps in rt-fMRI NF FC research to define precise target processes.

5.5 Implications for transdiagnostic approaches

As outlined in chapter 1.2.1, transdiagnostic approaches focus on the similarities of different mental disorders rather than their differences (Dalgleish et al., 2020). In a transdiagnostic approach typically many disorders within a specific problem area (e.g. eating disorders) are targeted or multiple problem areas (Martin et al., 2018). One important rationale for this is that mental health symptoms can be best described on a continuum rather than in an all-or-none way (Dalgleish et al., 2020). Therefore, application of rt-fMRI NF transdiagnostically would not only present with high clinical utility, but would also be logical from an economical perspective, because a higher number of participants could potentially be reached.

Precise target processes must not necessarily be associated with a single mental disorder, but accurately measurable transdiagnostic target processes are also an option. The two studies described in this dissertation applied frontostriatal networks as a target for regulation. Frontostriatal networks are considered a reasonable transdiagnostic approach for rt-fMRI NF FC and it was described previously that they play a key role in a majority of mental disorders. The association of these networks

with clinical functioning was e.g. outlined in a study with high-trait anxiety in which an increase in target FC between DLPFC and ACC was achieved after two NF sessions and these changes were correlated with a decrease in anxiety (Morgenroth et al., 2020). This underlines the strong potential of such an approach also in clinical contexts aiming at additionally inducing changes on the behavioral level. To summarize, application of frontostriatal networks as target opens up many new and unique options of use in individual psychiatric disorders, but also beyond many of them combined – in a transdiagnostic approach.

To sum up, this dissertation has shown that regulation of a large-scale frontostriatal network by means of rt-fMRI NF FC is feasible and presents with strong potential with regard to potential applications. However, for optimal conduct of rt-fMRI NF FC data correction for artefacts as important parameter that might be considered nontrivial at first must be addressed. With regard to general methodological aspects, a rigid and well-controlled design is preferred.

6 SUMMARY

Real-time functional magnetic resonance imaging neurofeedback (rt-fMRI NF) is a non-invasive approach to voluntarily control brain activity. Whereas the first decades since its development have been focussed on Region-of-Interest (ROI)-based rt-fMRI NF that feeds back activity of a single brain region as a target brain process, new approaches are constantly developed. One of these promising new approaches is functional connectivity (FC)-based rt-fMRI NF which is characterized by its focus on network-based brain connectivity. This approach has the advantage to be applicable in a plurality of mental disorders, since brain structures with central functions are known to be disturbed in a variety of mental disorders. Secondly, most mental disorders are characterized by distortions in whole brain networks as opposed to specific areas. Therefore, a strong advantage of the approach would be its application in a transdiagnostic way. Due to their involvement in a variety of mental disorders i.e. schizophrenia, frontostriatal networks are a suitable target network for the application of such approaches. The present dissertation attempted to investigate the feasibility of rt-fMRI NF FC targeting large-scale frontostriatal networks.

In a first pre-registered, randomized, double-blind and yoke-controlled proof-of-concept study conducted in a single training session, it was found that application of a large-scale frontostriatal target network including dorsolateral prefrontal cortex (DLPFC), anterior cingulate cortex (ACC) and striatum can be strongly biased by respiratory artefacts. Following up on this, a way to remove the influence of these respiratory artefacts in the offline analyses was found. Global Signal Regression (GSR) and a model-based physiological nuisance regression approach were able to eliminate these detrimental associations and clean the data sufficiently. It was concluded that GSR might be adequate for online data correction for respiratory artefacts.

In the subsequent second study, the methodology was elaborated on the basis of the findings from the first study. This included an extension of the training duration from one to three training sessions for demonstration of learning effects and, most importantly, application of GSR to the offline and specifically also the online data. Furthermore, the network was reduced to DLPFC and striatum only to increase adjustability. While the pre-registered significant effect of the NF training could not be

proven, we found evidence for a medium effect size in the last NF run on the third training day.

In summary, the present dissertation aimed at enhancing the understanding of rt-fMRI NF FC and address its feasibility. It is concluded that extensive NF training is a promising method for generating circuit-specific treatment approaches for mental disorders that could translate to symptom improvements. However, careful correction of data and preceding power analyses to allow for reasonable sample sizes are required in future rt-fMRI NF FC studies.

7 LITERATURVERZEICHNIS

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1997). Empirical analyses of BOLD fMRI statistics. II. Spatially smoothed data collected under null-hypothesis and experimental conditions. *Neuroimage*, *5*(3), 199-212. https://doi.org/10.1006/nimg.1997.0264

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). The inferential impact of global signal covariates in functional neuroimaging analyses. *Neuroimage*, *8*(3), 302-306. https://doi.org/10.1006/nimg.1998.0367

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Alegria, A. A., Wulff, M., Brinson, H., Barker, G. J., Norman, L. J., Brandeis, D., Stahl, D., David, A. S., Taylor, E., Giampietro, V., & Rubia, K. (2017). Real-time fMRI neurofeedback in adolescents with attention deficit hyperactivity disorder. *Hum Brain Mapp*, *38*(6), 3190-3209. https://doi.org/10.1002/hbm.23584

Auer, T., Schweizer, R., & Frahm, J. (2015). Training Efficiency and Transfer Success in an Extended Real-Time Functional MRI Neurofeedback Training of the Somatomotor Cortex of Healthy Subjects. *Front Hum Neurosci*, *9*, 547. https://doi.org/10.3389/fnhum.2015.00547

Bagarinao, E., Nakai, T., & Tanaka, Y. (2006). Real-time functional MRI: development and emerging applications. *Magn Reson Med Sci*, *5*(3), 157-165. https://doi.org/10.2463/mrms.5.157

Bandettini, P. A., Wong, E. C., Hinks, R. S., Tikofsky, R. S., & Hyde, J. S. (1992). Time course EPI of human brain function during task activation. *Magn Reson Med*, *25*(2), 390-397. https://doi.org/10.1002/mrm.1910250220

Bauer, C. C. C., Okano, K., Ghosh, S. S., Lee, Y. J., Melero, H., Angeles, C. L., Nestor, P. G., Del Re, E. C., Northoff, G., Niznikiewicz, M. A., & Whitfield-Gabrieli, S. (2020). Real-time fMRI neurofeedback reduces auditory hallucinations and modulates resting state connectivity of involved brain regions: Part 2: Default mode network -preliminary evidence. *Psychiatry Res*, *284*, 112770. https://doi.org/10.1016/j.psychres.2020.112770

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck depression inventory-II*. Psychological Corporation.

Bennett, C. M., & Miller, M. B. (2013). fMRI reliability: influences of task and experimental design. *Cogn Affect Behav Neurosci*, *13*(4), 690-702. https://doi.org/10.3758/s13415-013-0195-1

Beste, C., Ness, V., Lukas, C., Hoffmann, R., Stuwe, S., Falkenstein, M., & Saft, C. (2012). Mechanisms mediating parallel action monitoring in fronto-striatal circuits. *Neuroimage*, *62*(1), 137-146. https://doi.org/10.1016/j.neuroimage.2012.05.019

Bhattacharyya, P. K., & Lowe, M. J. (2004). Cardiac-induced physiologic noise in tissue is a direct observation of cardiac-induced fluctuations. *Magn Reson Imaging*, *22*(1), 9-13. https://doi.org/10.1016/j.mri.2003.08.003

Bianciardi, M., Fukunaga, M., van Gelderen, P., Horovitz, S. G., de Zwart, J. A., Shmueli, K., & Duyn, J. H. (2009). Sources of functional magnetic resonance imaging signal fluctuations in the human brain at rest: a 7 T study. *Magn Reson Imaging*, *27*(8), 1019-1029. https://doi.org/10.1016/j.mri.2009.02.004

Birn, R. M., Diamond, J. B., Smith, M. A., & Bandettini, P. A. (2006). Separating respiratory-variation-related fluctuations from neuronal-activity-related

fluctuations in fMRI. *Neuroimage*, *31*(4), 1536-1548. https://doi.org/10.1016/j.neuroimage.2006.02.048

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behav Brain Sci*, *30*(5-6), 481-499; discussion 499-548. https://doi.org/10.1017/S0140525X07002786

Bones, O., Hopkins, K., Krishnan, A., & Plack, C. J. (2014). Phase locked neural activity in the human brainstem predicts preference for musical consonance. *Neuropsychologia*, *58*, 23-32. https://doi.org/10.1016/j.neuropsychologia.2014.03.011

Booth, J. R., Macwhinney, B., Thulborn, K. R., Sacco, K., Voyvodic, J., & Feldman, H. M. (1999). Functional organization of activation patterns in children: whole brain fMRI imaging during three different cognitive tasks. *Prog Neuropsychopharmacol Biol Psychiatry*, *23*(4), 669-682. https://doi.org/10.1016/s0278-5846(99)00025-1

Braun, U., Schaefer, A., Betzel, R. F., Tost, H., Meyer-Lindenberg, A., & Bassett, D. S. (2018). From Maps to Multi-dimensional Network Mechanisms of Mental Disorders. *Neuron*, *97*(1), 14-31. https://doi.org/10.1016/j.neuron.2017.11.007

Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*, *10*(3), 186-198. https://doi.org/10.1038/nrn2575

Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nat Rev Neurosci*, *13*(5), 336-349. https://doi.org/10.1038/nrn3214

Burgess, G. C., Kandala, S., Nolan, D., Laumann, T. O., Power, J. D., Adeyemo, B., Harms, M. P., Petersen, S. E., & Barch, D. M. (2016). Evaluation of Denoising Strategies to Address Motion-Correlated Artifacts in Resting-State Functional Magnetic Resonance Imaging Data from the Human Connectome Project. *Brain Connect*, *6*(9), 669-680. https://doi.org/10.1089/brain.2016.0435

Caballero-Gaudes, C., & Reynolds, R. C. (2017). Methods for cleaning the BOLD fMRI signal. *Neuroimage*, *154*, 128-149. https://doi.org/10.1016/j.neuroimage.2016.12.018

Cadena, E. J., White, D. M., Kraguljac, N. V., Reid, M. A., & Lahti, A. C. (2018). Evaluation of fronto-striatal networks during cognitive control in unmedicated patients with schizophrenia and the effect of antipsychotic medication. *NPJ Schizophr*, *4*(1), 8. https://doi.org/10.1038/s41537-018-0051-y

Calhoun, V. D., Adali, T., Pearlson, G. D., van Zijl, P. C., & Pekar, J. J. (2002). Independent component analysis of fMRI data in the complex domain. *Magn Reson Med*, *48*(1), 180-192. https://doi.org/10.1002/mrm.10202

Canterberry, M., Hanlon, C. A., Hartwell, K. J., Li, X., Owens, M., LeMatty, T., Prisciandaro, J. J., Borckardt, J., Saladin, M. E., Brady, K. T., & George, M. S. (2013). Sustained reduction of nicotine craving with real-time neurofeedback: exploring the role of severity of dependence. *Nicotine Tob Res*, *15*(12), 2120-2124. https://doi.org/10.1093/ntr/ntt122

Caria, A., Sitaram, R., & Birbaumer, N. (2012). Real-time fMRI: a tool for local brain regulation. *Neuroscientist*, *18*(5), 487-501. https://doi.org/10.1177/1073858411407205

Chai, X. J., Castanon, A. N., Ongur, D., & Whitfield-Gabrieli, S. (2012). Anticorrelations in resting state networks without global signal regression. *Neuroimage*, *59*(2), 1420-1428. https://doi.org/10.1016/j.neuroimage.2011.08.048

Chang, C., & Glover, G. H. (2009). Effects of model-based physiological noise correction on default mode network anti-correlations and correlations.

Neuroimage, *47*(4), 1448-1459. https://doi.org/10.1016/j.neuroimage.2009.05.012

Chen, G., Chen, G., Xie, C., Ward, B. D., Li, W., Antuono, P., & Li, S. J. (2012). A method to determine the necessity for global signal regression in resting-state fMRI studies. *Magn Reson Med*, *68*(6), 1828-1835. https://doi.org/10.1002/mrm.24201

Choi, E. Y., Yeo, B. T., & Buckner, R. L. (2012). The organization of the human striatum estimated by intrinsic functional connectivity. *J Neurophysiol*, *108*(8), 2242-2263. https://doi.org/10.1152/jn.00270.2012

Chu, P. P. W., Golestani, A. M., Kwinta, J. B., Khatamian, Y. B., & Chen, J. J. (2018). Characterizing the modulation of resting-state fMRI metrics by baseline physiology. *Neuroimage*, *173*, 72-87. https://doi.org/10.1016/j.neuroimage.2018.02.004

Cordes, J. S., Mathiak, K. A., Dyck, M., Alawi, E. M., Gaber, T. J., Zepf, F. D., Klasen, M., Zvyagintsev, M., Gur, R. C., & Mathiak, K. (2015). Cognitive and neural strategies during control of the anterior cingulate cortex by fMRI neurofeedback in patients with schizophrenia. *Front Behav Neurosci*, *9*, 169. https://doi.org/10.3389/fnbeh.2015.00169

Correll, C. U., & Schooler, N. R. (2020). Negative Symptoms in Schizophrenia: A Review and Clinical Guide for Recognition, Assessment, and Treatment. *Neuropsychiatr Dis Treat*, *16*, 519-534. https://doi.org/10.2147/NDT.S225643

Costa, P. T., Jr., & McCrae, R. R. (2008). The Revised NEO Personality Inventory (NEO-PI-R). In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment, Vol. 2. Personality measurement and testing* (pp. 179–198). Sage Publications, Inc. https://doi.org/10.4135/9781849200479.n9

Courtney, K. E., Ghahremani, D. G., & Ray, L. A. (2013). Fronto-striatal functional connectivity during response inhibition in alcohol dependence. *Addict Biol*, *18*(3), 593-604. https://doi.org/10.1111/adb.12013

Cox, R. W., Jesmanowicz, A., & Hyde, J. S. (1995). Real-time functional magnetic resonance imaging. *Magn Reson Med*, *33*(2), 230-236. https://doi.org/10.1002/mrm.1910330213

Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron*, *29*(2), 537-545. https://doi.org/10.1016/s0896-6273(01)00225-2

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*(2), 194-215. https://doi.org/10.1016/j.cortex.2011.04.007

Czeisler, M. É., Lane, R. I., Petrosky, E., Wiley, J. F., Christensen, A., Njai, R., Weaver, M. D., Robbins, R., Facer-Childs, E. R., Barger, L. K., Czeisler, C. A., Howard, M. E., & Rajaratnam, S. (2020). Mental Health, Substance Use, and Suicidal Ideation During the COVID-19 Pandemic - United States, June 24-30, 2020. *MMWR. Morbidity and mortality weekly report*, *69*(32), 1049–1057. https://doi.org/10.15585/mmwr.mm6932a1

Dalgleish, T., Black, M., Johnston, D., & Bevan, A. (2020). Transdiagnostic approaches to mental health problems: Current status and future directions. *J Consult Clin Psychol*, *88*(3), 179-195. https://doi.org/10.1037/ccp0000482

Damaraju, E., Allen, E. A., Belger, A., Ford, J. M., McEwen, S., Mathalon, D. H., Mueller, B. A., Pearlson, G. D., Potkin, S. G., Preda, A., Turner, J. A., Vaidya,

J. G., van Erp, T. G., & Calhoun, V. D. (2014). Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *Neuroimage Clin*, *5*, 298-308. https://doi.org/10.1016/j.nicl.2014.07.003

Dattani, S., Ritchie, H., & Roser, M. (2021). *Mental Health*. OurWorldInData. Retrieved February 27, 2022, from: https://ourworldindata.org/mental-health [Online Resource]

deCharms, R. C., Maeda, F., Glover, G. H., Ludlow, D., Pauly, J. M., Soneji, D., Gabrieli, J. D., & Mackey, S. C. (2005). Control over brain activation and pain learned by using real-time functional MRI. *Proc Natl Acad Sci U S A*, *102*(51), 18626-18631. https://doi.org/10.1073/pnas.0505210102

Dudek, E., & Dodell-Feder, D. (2021). The efficacy of real-time functional magnetic resonance imaging neurofeedback for psychiatric illness: A meta-analysis of brain and behavioral outcomes. *Neurosci Biobehav Rev*, *121*, 291-306. https://doi.org/10.1016/j.neubiorev.2020.12.020

Dyck, M. S., Mathiak, K. A., Bergert, S., Sarkheil, P., Koush, Y., Alawi, E. M., Zvyagintsev, M., Gaebler, A. J., Shergill, S. S., & Mathiak, K. (2016). Targeting Treatment-Resistant Auditory Verbal Hallucinations in Schizophrenia with fMRI-Based Neurofeedback - Exploring Different Cases of Schizophrenia. *Front Psychiatry*, *7*, 37. https://doi.org/10.3389/fpsyt.2016.00037

Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An FMRI study of social exclusion. *Science*, *302*(5643), 290-292. https://doi.org/10.1126/science.1089134

Emmert, K., Breimhorst, M., Bauermann, T., Birklein, F., Rebhorn, C., Van De Ville, D., & Haller, S. (2017). Active pain coping is associated with the response in real-time fMRI neurofeedback during pain. *Brain Imaging Behav*, *11*(3), 712-721. https://doi.org/10.1007/s11682-016-9547-0

Emmert, K., Kopel, R., Koush, Y., Maire, R., Senn, P., Van De Ville, D., & Haller, S. (2017). Continuous vs. intermittent neurofeedback to regulate auditory cortex activity of tinnitus patients using real-time fMRI - A pilot study. *Neuroimage Clin*, *14*, 97-104. https://doi.org/10.1016/j.nicl.2016.12.023

Emmert, K., Kopel, R., Sulzer, J., Bruhl, A. B., Berman, B. D., Linden, D. E. J., Horovitz, S. G., Breimhorst, M., Caria, A., Frank, S., Johnston, S., Long, Z., Paret, C., Robineau, F., Veit, R., Bartsch, A., Beckmann, C. F., Van De Ville, D., & Haller, S. (2016). Meta-analysis of real-time fMRI neurofeedback studies using individual participant data: How is brain regulation mediated? *Neuroimage*, *124*(Pt A), 806-812. https://doi.org/10.1016/j.neuroimage.2015.09.042

Fede, S. J., Dean, S. F., Manuweera, T., & Momenan, R. (2020). A Guide to Literature Informed Decisions in the Design of Real Time fMRI Neurofeedback Studies: A Systematic Review. *Front Hum Neurosci*, *14*, 60. https://doi.org/10.3389/fnhum.2020.00060

Fornito, A., & Bullmore, E. T. (2015). Reconciling abnormalities of brain network structure and function in schizophrenia. *Curr Opin Neurobiol*, *30*, 44-50. https://doi.org/10.1016/j.conb.2014.08.006

Foussias, G., & Remington, G. (2010). Negative symptoms in schizophrenia: avolition and Occam's razor. *Schizophr Bull*, *36*(2), 359-369. https://doi.org/10.1093/schbul/sbn094

Fox, M. D., Zhang, D., Snyder, A. Z., & Raichle, M. E. (2009). The global signal and observed anticorrelated resting state brain networks. *J Neurophysiol*, *101*(6), 3270-3283. https://doi.org/10.1152/jn.90777.2008

Friston, K., Brown, H. R., Siemerkus, J., & Stephan, K. E. (2016). The dysconnection hypothesis (2016). *Schizophr Res*, *176*(2-3), 83-94. https://doi.org/10.1016/j.schres.2016.07.014

Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., & Penny, W. (2007). Variational free energy and the Laplace approximation. *Neuroimage*, *34*(1), 220-234. https://doi.org/10.1016/j.neuroimage.2006.08.035

Friston, K. J., & Frith, C. D. (1995). Schizophrenia: a disconnection syndrome? *Clin Neurosci*, *3*(2), 89-97. https://www.ncbi.nlm.nih.gov/pubmed/7583624

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *Neuroimage*, *19*(4), 1273-1302. https://doi.org/10.1016/s1053-8119(03)00202-7

Furman, D. J., Hamilton, J. P., & Gotlib, I. H. (2011). Frontostriatal functional connectivity in major depressive disorder. *Biol Mood Anxiety Disord*, *1*(1), 11. https://doi.org/10.1186/2045-5380-1-11

Gandara, V., Pineda, J. A., Shu, I. W., & Singh, F. (2020). A Systematic Review of the Potential Use of Neurofeedback in Patients With Schizophrenia. *Schizophr Bull Open*, *1*(1), sgaa005. https://doi.org/10.1093/schizbullopen/sgaa005

Gonzalez, C., Dana, J., Koshino, H., Just, M. (2005). The framing effect and risky decisions: Examining cognitive functions with fMRI. *Journal of Economic Psychology, 26.* https://doi.org/10.1016/j.joep.2004.08.004

Geerligs, L., Tsvetanov, K. A., Cam, C., & Henson, R. N. (2017). Challenges in measuring individual differences in functional connectivity using fMRI: The case of healthy aging. *Hum Brain Mapp*, *38*(8), 4125-4156. https://doi.org/10.1002/hbm.23653

Gerchen, M. F., Kirsch, M., Bahs, N., Halli, P., Gerhardt, S., Schafer, A., Sommer, W. H., Kiefer, F., & Kirsch, P. (2018). The SyBil-AA real-time fMRI neurofeedback study: protocol of a single-blind randomized controlled trial in alcohol use disorder. *BMC Psychiatry*, *18*(1), 12. https://doi.org/10.1186/s12888-018-1604-3

Gerchen, M. F., Kirsch, P., & Feld, G. B. (2021). Brain-wide inferiority and equivalence tests in fMRI group analyses: Selected applications. *Hum Brain Mapp*, *42*(18), 5803-5813. https://doi.org/10.1002/hbm.25664

Gerchen, M. F., Weiss, F., Kirsch, M., Rentsch, A., Halli, P., Kiefer, F., & Kirsch, P. (2021). Dynamic frontostriatal functional peak connectivity (in alcohol use disorder). *Hum Brain Mapp*, *42*(1), 36-46. https://doi.org/10.1002/hbm.25201

Gerin, M. I., Fichtenholtz, H., Roy, A., Walsh, C. J., Krystal, J. H., Southwick, S., & Hampson, M. (2016). Real-Time fMRI Neurofeedback with War Veterans with Chronic PTSD: A Feasibility Study. *Front Psychiatry*, *7*, 111. https://doi.org/10.3389/fpsyt.2016.00111

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc Natl Acad Sci U S A*, *108 Suppl 3*, 15647-15654. https://doi.org/10.1073/pnas.1014269108

Glover, G. H., Li, T. Q., & Ress, D. (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn Reson Med*, *44*(1), 162-167. https://doi.org/10.1002/1522-2594(200007)44:1<162::aid-mrm23>3.0.co;2-e

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105-2108. https://doi.org/10.1126/science.1062872

Haase, L., Cerf-Ducastel, B., & Murphy, C. (2009). Cortical activation in response to pure taste stimuli during the physiological states of hunger and satiety. *Neuroimage*, *44*(3), 1008-1021. https://doi.org/10.1016/j.neuroimage.2008.09.044

Hahamy, A., Calhoun, V., Pearlson, G., Harel, M., Stern, N., Attar, F., Malach, R., & Salomon, R. (2014). Save the global: global signal connectivity as a tool for studying clinical populations with functional magnetic resonance imaging. *Brain Connect*, *4*(6), 395-403. https://doi.org/10.1089/brain.2014.0244

Hampson, M., Stoica, T., Saksa, J., Scheinost, D., Qiu, M., Bhawnani, J., Pittenger, C., Papademetris, X., & Constable, T. (2012). Real-time fMRI biofeedback targeting the orbitofrontal cortex for contamination anxiety. *J Vis Exp*(59). https://doi.org/10.3791/3535

Harrison, B. J., Soriano-Mas, C., Pujol, J., Ortiz, H., Lopez-Sola, M., Hernandez-Ribas, R., Deus, J., Alonso, P., Yucel, M., Pantelis, C., Menchon, J. M., & Cardoner, N. (2009). Altered corticostriatal functional connectivity in obsessive-compulsive disorder. *Arch Gen Psychiatry*, *66*(11), 1189-1200. https://doi.org/10.1001/archgenpsychiatry.2009.152

Hartwell, K. J., Hanlon, C. A., Li, X., Borckardt, J. J., Canterberry, M., Prisciandaro, J. J., Moran-Santa Maria, M. M., LeMatty, T., George, M. S., & Brady, K. T. (2016). Individualized real-time fMRI neurofeedback to attenuate craving in nicotine-dependent smokers. *J Psychiatry Neurosci*, *41*(1), 48-55. https://doi.org/10.1503/jpn.140200

Hauser, T. U., Iannaccone, R., Dolan, R. J., Ball, J., Hattenschwiler, J., Drechsler, R., Rufer, M., Brandeis, D., Walitza, S., & Brem, S. (2017). Increased fronto-striatal reward prediction errors moderate decision making in obsessive-compulsive disorder. *Psychol Med*, *47*(7), 1246-1258. https://doi.org/10.1017/S0033291716003305

Heck, D. H., McAfee, S. S., Liu, Y., Babajani-Feremi, A., Rezaie, R., Freeman, W. J., Wheless, J. W., Papanicolaou, A. C., Ruszinko, M., Sokolov, Y., & Kozma, R. (2016). Breathing as a Fundamental Rhythm of Brain Function. *Front Neural Circuits*, *10*, 115. https://doi.org/10.3389/fncir.2016.00115

Heeger, D. J., & Ress, D. (2002). What does fMRI tell us about neuronal activity? *Nat Rev Neurosci*, *3*(2), 142-151. https://doi.org/10.1038/nrn730

Heinz, A., Kiefer, F., Smolka, M. N., Endrass, T., Beste, C., Beck, A., Liu, S., Genauck, A., Romund, L., Banaschewski, T., Bermpohl, F., Deserno, L., Dolan, R. J., Durstewitz, D., Ebner-Priemer, U., Flor, H., Hansson, A. C., Heim, C., Hermann, D., Kiebel, S., Kirsch, P., Kirschbaum, C., Koppe, G., Marxen, M., Meyer-Lindenberg, A., Nagel, W. E., Noori, H. R., Pilhatsch, M., Priller, J., Rietschel, M., Romanczuk-Seiferth, N., Schlagenhauf, F., Sommer, W. H., Stallkamp, J., Strohle, A., Stock, A. K., Winterer, G., Winter, C., Walter, H., Witt, S., Vollstadt-Klein, S., Rapp, M. A., Tost, H., & Spanagel, R. (2020). Addiction Research Consortium: Losing and regaining control over drug intake (ReCoDe)-From trajectories to mechanisms and interventions. *Addict Biol*, *25*(2), e12866. https://doi.org/10.1111/adb.12866

Hellrung, L., Dietrich, A., Hollmann, M., Pleger, B., Kalberlah, C., Roggenhofer, E., Villringer, A., & Horstmann, A. (2018). Intermittent compared to continuous real-time fMRI neurofeedback boosts control over amygdala activation. *Neuroimage*, *166*, 198-208. https://doi.org/10.1016/j.neuroimage.2017.10.031

Hellrung, L., Borchardt, V., Götting, F.N, Stadler, J., Tempelmann, C., Tobler, P.N., Walter, M., van der Meer J.N. (2018). *Motion and physiological noise*

effects on amygdala real-time fMRI neurofeedback learning. BioRxiv. https://doi.org/10.1101/366138

Herrero, J. L., Khuvis, S., Yeagle, E., Cerf, M., & Mehta, A. D. (2018). Breathing above the brain stem: volitional control and attentional modulation in humans. *J Neurophysiol*, *119*(1), 145-159. https://doi.org/10.1152/jn.00551.2017

Heunis, S., Lamerichs, R., Zinger, S., Caballero-Gaudes, C., Jansen, J. F. A., Aldenkamp, B., & Breeuwer, M. (2020). Quality and denoising in real-time functional magnetic resonance imaging neurofeedback: A methods review. *Hum Brain Mapp*, *41*(12), 3439-3467. https://doi.org/10.1002/hbm.25010

Hu, X., Le, T. H., Parrish, T., & Erhard, P. (1995). Retrospective estimation and correction of physiological fluctuation in functional MRI. *Magn Reson Med*, *34*(2), 201-212. https://doi.org/10.1002/mrm.1910340211

Hull, C. L. (1943). Principles of Behavior. New York: Appleton-Century-Crofts.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med*, *2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Ito, J., Roy, S., Liu, Y., Cao, Y., Fletcher, M., Lu, L., Boughter, J. D., Grun, S., & Heck, D. H. (2014). Whisker barrel cortex delta oscillations and gamma power in the awake mouse are linked to respiration. *Nat Commun*, *5*, 3572. https://doi.org/10.1038/ncomms4572

Johnson, K. A., Hartwell, K., LeMatty, T., Borckardt, J., Morgan, P. S., Govindarajan, K., Brady, K., & George, M. S. (2012). Intermittent "real-time" fMRI feedback is superior to continuous presentation for a motor imagery task: a pilot study. *J Neuroimaging*, *22*(1), 58-66. https://doi.org/10.1111/j.1552-6569.2010.00529.x

Kadosh, K. C., & Staunton, G. (2019). A systematic review of the psychological factors that influence neurofeedback learning outcomes. *Neuroimage*, *185*, 545-555. https://doi.org/10.1016/j.neuroimage.2018.10.021

Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D., & Pizzagalli, D. A. (2015). Large-Scale Network Dysfunction in Major Depressive Disorder: A Meta-analysis of Resting-State Functional Connectivity. *JAMA Psychiatry*, *72*(6), 603-611. https://doi.org/10.1001/jamapsychiatry.2015.0071

Karch, S., Keeser, D., Hummer, S., Paolini, M., Kirsch, V., Karali, T., Kupka, M., Rauchmann, B. S., Chrobok, A., Blautzik, J., Koller, G., Ertl-Wagner, B., & Pogarell, O. (2015). Modulation of Craving Related Brain Responses Using Real-Time fMRI in Patients with Alcohol Use Disorder. *PLoS One*, *10*(7), e0133034. https://doi.org/10.1371/journal.pone.0133034

Karch, S., Paolini, M., Gschwendtner, S., Jeanty, H., Reckenfelderbaumer, A., Yaseen, O., Maywald, M., Fuchs, C., Rauchmann, B. S., Chrobok, A., Rabenstein, A., Ertl-Wagner, B., Pogarell, O., Keeser, D., & Ruther, T. (2019). Real-Time fMRI Neurofeedback in Patients With Tobacco Use Disorder During Smoking Cessation: Functional Differences and Implications of the First Training Session in Regard to Future Abstinence or Relapse. *Front Hum Neurosci*, *13*, 65. https://doi.org/10.3389/fnhum.2019.00065

Kasper, L., Bollmann, S., Diaconescu, A. O., Hutton, C., Heinzle, J., Iglesias, S., Hauser, T. U., Sebold, M., Manjaly, Z. M., Pruessmann, K. P., & Stephan, K. E. (2017). The PhysIO Toolbox for Modeling Physiological Noise in fMRI Data. *J Neurosci Methods*, *276*, 56-72. https://doi.org/10.1016/j.jneumeth.2016.10.019

Kirsch, M., Gruber, I., Ruf, M., Kiefer, F., & Kirsch, P. (2016). Real-time functional magnetic resonance imaging neurofeedback can reduce striatal cue-reactivity to alcohol stimuli. *Addict Biol*, *21*(4), 982-992. https://doi.org/10.1111/adb.12278

Kober, S. E., Witte, M., Ninaus, M., Neuper, C., & Wood, G. (2013). Learning to modulate one's own brain activity: the effect of spontaneous mental strategies. *Front Hum Neurosci*, *7*, 695. https://doi.org/10.3389/fnhum.2013.00695

Kohl, S. H., Veit, R., Spetter, M. S., Gunther, A., Rina, A., Luhrs, M., Birbaumer, N., Preissl, H., & Hallschmid, M. (2019). Real-time fMRI neurofeedback training to improve eating behavior by self-regulation of the dorsolateral prefrontal cortex: A randomized controlled trial in overweight and obese subjects. *Neuroimage*, *191*, 596-609. https://doi.org/10.1016/j.neuroimage.2019.02.033

Koizumi, A., Amano, K., Cortese, A., Shibata, K., Yoshida, W., Seymour, B., Kawato, M., & Lau, H. (2016). Fear reduction without fear through reinforcement of neural activity that bypasses conscious exposure. *Nat Hum Behav*, *1*. https://doi.org/10.1038/s41562-016-0006

Koush, Y., Meskaldji, D. E., Pichon, S., Rey, G., Rieger, S. W., Linden, D. E., Van De Ville, D., Vuilleumier, P., & Scharnowski, F. (2017). Learning Control Over Emotion Networks Through Connectivity-Based Neurofeedback. *Cereb Cortex*, *27*(2), 1193-1202. https://doi.org/10.1093/cercor/bhv311

Koush, Y., Rosa, M. J., Robineau, F., Heinen, K., S, W. R., Weiskopf, N., Vuilleumier, P., Van De Ville, D., & Scharnowski, F. (2013). Connectivity-based neurofeedback: dynamic causal modeling for real-time fMRI. *Neuroimage*, *81*, 422-430. https://doi.org/10.1016/j.neuroimage.2013.05.010

Kruger, G., & Glover, G. H. (2001). Physiological noise in oxygenation-sensitive magnetic resonance imaging. *Magn Reson Med*, *46*(4), 631-637. https://doi.org/10.1002/mrm.1240

Lacroix, J. M. (1981). The acquisition of autonomic control through biofeedback: the case against an afferent process and a two-process alternative. *Psychophysiology*, *18*(5), 573-587. https://doi.org/10.1111/j.1469-8986.1981.tb01828.x

Lin, P., Wang, X., Zhang, B., Kirkpatrick, B., Ongur, D., Levitt, J. J., Jovicich, J., Yao, S., & Wang, X. (2018). Functional dysconnectivity of the limbic loop of frontostriatal circuits in first-episode, treatment-naive schizophrenia. *Hum Brain Mapp*, *39*(2), 747-757. https://doi.org/10.1002/hbm.23879

Lindquist, M. A., Xu, Y., Nebel, M. B., & Caffo, B. S. (2014). Evaluating dynamic bivariate correlations in resting-state fMRI: a comparison study and a new approach. *Neuroimage*, *101*, 531-546. https://doi.org/10.1016/j.neuroimage.2014.06.052

Liu, T. T. (2016). Noise contributions to the fMRI signal: An overview. *Neuroimage*, *143*, 141-151. https://doi.org/10.1016/j.neuroimage.2016.09.008

Liu, T. T., Nalci, A., & Falahpour, M. (2017). The global signal in fMRI: Nuisance or Information? *Neuroimage*, *150*, 213-229. https://doi.org/10.1016/j.neuroimage.2017.02.036

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869-878. https://doi.org/10.1038/nature06976

Long, Z., Guo, Z., Guo, Z., Zhang, H., & Yao, L. (2021). Dynamic Functional Network Connectivity Changes Associated with fMRI Neurofeedback of Right Premotor Cortex. *Brain Sci*, *11*(5). https://doi.org/10.3390/brainsci11050582

Lorenzetti, V., Melo, B., Basilio, R., Suo, C., Yucel, M., Tierra-Criollo, C. J., & Moll, J. (2018). Emotion Regulation Using Virtual Environments and Real-Time fMRI Neurofeedback. *Front Neurol*, *9*, 390. https://doi.org/10.3389/fneur.2018.00390

MacDuffie, K. E., MacInnes, J., Dickerson, K. C., Eddington, K. M., Strauman, T. J., & Adcock, R. A. (2018). Single session real-time fMRI neurofeedback has a

lasting impact on cognitive behavioral therapy strategies. *Neuroimage Clin*, *19*, 868-875. https://doi.org/10.1016/j.nicl.2018.06.009

Mainero, C., Caramia, F., Pozzilli, C., Pisani, A., Pestalozza, I., Borriello, G., Bozzao, L., & Pantano, P. (2004). fMRI evidence of brain reorganization during attention and memory tasks in multiple sclerosis. *Neuroimage*, *21*(3), 858-867. https://doi.org/10.1016/j.neuroimage.2003.10.004

Martin, P., Murray, L.L., Darnell, D., Dorsey, S. (2018). Transdiagnostic treatment approaches for greater public health impact: Implementing principles of evidence-based mental health interventions. *Clin Psychol Sci Pract., 25.* https://doi.org/10.1111/cpsp.12270

Martz, M. E., Hart, T., Heitzeg, M. M., & Peltier, S. J. (2020). Neuromodulation of brain activation associated with addiction: A review of real-time fMRI neurofeedback studies. *Neuroimage Clin*, *27*, 102350. https://doi.org/10.1016/j.nicl.2020.102350

Mayeli, A., Misaki, M., Zotev, V., Tsuchiyagaito, A., Al Zoubi, O., Phillips, R., Smith, J., Stewart, J. L., Refai, H., Paulus, M. P., & Bodurka, J. (2020). Self-regulation of ventromedial prefrontal cortex activation using real-time fMRI neurofeedback-Influence of default mode network. *Hum Brain Mapp*, *41*(2), 342-352. https://doi.org/10.1002/hbm.24805

Megumi, F., Yamashita, A., Kawato, M., & Imamizu, H. (2015). Functional MRI neurofeedback training on connectivity between two regions induces long-lasting changes in intrinsic functional network. *Front Hum Neurosci*, *9*, 160. https://doi.org/10.3389/fnhum.2015.00160

Mehler, D. M. A., Sokunbi, M. O., Habes, I., Barawi, K., Subramanian, L., Range, M., Evans, J., Hood, K., Luhrs, M., Keedwell, P., Goebel, R., & Linden, D. E. J. (2018). Targeting the affective brain-a randomized controlled trial of real-time fMRI neurofeedback in patients with depression. *Neuropsychopharmacology*, *43*(13), 2578-2585. https://doi.org/10.1038/s41386-018-0126-5

Melnikov, M. Y. (2021). The Current Evidence Levels for Biofeedback and Neurofeedback Interventions in Treating Depression: A Narrative Review. *Neural Plast*, *2021*, 8878857. https://doi.org/10.1155/2021/8878857

Mier, W., & Mier, D. (2015). Advantages in functional imaging of the brain. *Front Hum Neurosci*, *9*, 249. https://doi.org/10.3389/fnhum.2015.00249

Mihara, M., Miyai, I., Hattori, N., Hatakenaka, M., Yagura, H., Kawano, T., Okibayashi, M., Danjo, N., Ishikawa, A., Inoue, Y., & Kubota, K. (2012). Neurofeedback using real-time near-infrared spectroscopy enhances motor imagery related cortical activation. *PLoS One*, *7*(3), e32234. https://doi.org/10.1371/journal.pone.0032234

Misaki, M., & Bodurka, J. (2021). The impact of real-time fMRI denoising on online evaluation of brain activity and functional connectivity. *J Neural Eng*, *18*(4). https://doi.org/10.1088/1741-2552/ac0b33

Mkrtchian, A., Evans, J. W., Kraus, C., Yuan, P., Kadriu, B., Nugent, A. C., Roiser, J. P., & Zarate, C. A., Jr. (2021). Ketamine modulates fronto-striatal circuitry in depressed and healthy individuals. *Mol Psychiatry*, *26*(7), 3292-3301. https://doi.org/10.1038/s41380-020-00878-1

Morgenroth, E., Saviola, F., Gilleen, J., Allen, B., Luhrs, M., M, W. E., & Allen, P. (2020). Using connectivity-based real-time fMRI neurofeedback to modulate attentional and resting state networks in people with high trait anxiety. *Neuroimage Clin*, *25*, 102191. https://doi.org/10.1016/j.nicl.2020.102191

Munoz-Moldes, S., & Cleeremans, A. (2020). Delineating implicit and explicit processes in neurofeedback learning. *Neurosci Biobehav Rev*, *118*, 681-688. https://doi.org/10.1016/j.neubiorev.2020.09.003

Murphy, K., Birn, R. M., & Bandettini, P. A. (2013). Resting-state fMRI confounds and cleanup. *Neuroimage*, *80*, 349-359. https://doi.org/10.1016/j.neuroimage.2013.04.001

Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., & Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage*, *44*(3), 893-905. https://doi.org/10.1016/j.neuroimage.2008.09.036

Murphy, K., & Fox, M. D. (2017). Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage*, *154*, 169-173. https://doi.org/10.1016/j.neuroimage.2016.11.052

Nicholson, A. A., Rabellino, D., Densmore, M., Frewen, P. A., Paret, C., Kluetsch, R., Schmahl, C., Theberge, J., Neufeld, R. W., McKinnon, M. C., Reiss, J., Jetly, R., & Lanius, R. A. (2017). The neurobiology of emotion regulation in posttraumatic stress disorder: Amygdala downregulation via real-time fMRI neurofeedback. *Hum Brain Mapp*, *38*(1), 541-560. https://doi.org/10.1002/hbm.23402

Nikolaou, F., Orphanidou, C., Papakyriakou, P., Murphy, K., Wise, R. G., & Mitsis, G. D. (2016). Spontaneous physiological variability modulates dynamic functional connectivity in resting-state functional magnetic resonance imaging. *Philos Trans A Math Phys Eng Sci*, *374*(2067). https://doi.org/10.1098/rsta.2015.0183

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D., & Gross, J. J. (2004). For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. *Neuroimage*, *23*(2), 483-499. https://doi.org/10.1016/j.neuroimage.2004.06.030

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci U S A*, *87*(24), 9868-9872. https://doi.org/10.1073/pnas.87.24.9868

Oh, S., Kim, M., Kim, T., Lee, T. Y., & Kwon, J. S. (2020). Resting-state functional connectivity of the striatum predicts improvement in negative symptoms and general functioning in patients with first-episode psychosis: A 1-year naturalistic follow-up study. *Aust N Z J Psychiatry*, *54*(5), 509-518. https://doi.org/10.1177/0004867419885452

Orlov, N. D., Giampietro, V., O'Daly, O., Lam, S. L., Barker, G. J., Rubia, K., McGuire, P., Shergill, S. S., & Allen, P. (2018). Real-time fMRI neurofeedback to down-regulate superior temporal gyrus activity in patients with schizophrenia and auditory hallucinations: a proof-of-concept study. *Transl Psychiatry*, *8*(1), 46. https://doi.org/10.1038/s41398-017-0067-5

Papoutsi, M., Magerkurth, J., Josephs, O., Pepes, S. E., Ibitoye, T., Reilmann, R., Hunt, N., Payne, E., Weiskopf, N., Langbehn, D., Rees, G., & Tabrizi, S. J. (2020). Activity or connectivity? A randomized controlled feasibility study evaluating neurofeedback training in Huntington's disease. *Brain Commun*, *2*(1), fcaa049. https://doi.org/10.1093/braincomms/fcaa049

Paret, C., Goldway, N., Zich, C., Keynan, J. N., Hendler, T., Linden, D., & Cohen Kadosh, K. (2019). Current progress in real-time functional magnetic resonance-based neurofeedback: Methodological challenges and achievements. *Neuroimage*, *202*, 116107. https://doi.org/10.1016/j.neuroimage.2019.116107

Paret, C., Kluetsch, R., Ruf, M., Demirakca, T., Hoesterey, S., Ende, G., & Schmahl, C. (2014). Down-regulation of amygdala activation with real-time fMRI neurofeedback in a healthy female sample. *Front Behav Neurosci*, *8*, 299. https://doi.org/10.3389/fnbeh.2014.00299

Paret, C., Kluetsch, R., Zaehringer, J., Ruf, M., Demirakca, T., Bohus, M., Ende, G., & Schmahl, C. (2016). Alterations of amygdala-prefrontal connectivity with real-time fMRI neurofeedback in BPD patients. *Soc Cogn Affect Neurosci*, *11*(6), 952-960. https://doi.org/10.1093/scan/nsw016

Paret, C., Zahringer, J., Ruf, M., Gerchen, M. F., Mall, S., Hendler, T., Schmahl, C., & Ende, G. (2018). Monitoring and control of amygdala neurofeedback involves distributed information processing in the human brain. *Hum Brain Mapp*, *39*(7), 3018-3031. https://doi.org/10.1002/hbm.24057

Patel, K., Sutherland, H., Henshaw, J., Taylor, J. R., Brown, C. A., Casson, A. J., Trujillo-Barreton, N. J., Jones, A. K. P., & Sivan, M. (2020). Effects of neurofeedback in the management of chronic pain: A systematic review and meta-analysis of clinical trials. *Eur J Pain*, *24*(8), 1440-1457. https://doi.org/10.1002/ejp.1612

Pereira, J., Direito, B., Sayal, A., Ferreira, C., & Castelo-Branco, M. (2019). Self-Modulation of Premotor Cortex Interhemispheric Connectivity in a Real-Time Functional Magnetic Resonance Imaging Neurofeedback Study Using an Adaptive Approach. *Brain Connect*, *9*(9), 662-672. https://doi.org/10.1089/brain.2019.0697

Peters, H., Riedl, V., Manoliu, A., Scherr, M., Schwerthoffer, D., Zimmer, C., Forstl, H., Bauml, J., Sorg, C., & Koch, K. (2017). Changes in extra-striatal functional connectivity in patients with schizophrenia in a psychotic episode. *Br J Psychiatry*, *210*(1), 75-82. https://doi.org/10.1192/bjp.bp.114.151928

Pettersson-Yeo, W., Allen, P., Benetti, S., McGuire, P., & Mechelli, A. (2011). Dysconnectivity in schizophrenia: where are we now? *Neurosci Biobehav Rev*, *35*(5), 1110-1124. https://doi.org/10.1016/j.neubiorev.2010.11.004

Poldrack, R. A. (2012). The future of fMRI in cognitive neuroscience. *Neuroimage*, *62*(2), 1216-1220. https://doi.org/10.1016/j.neuroimage.2011.08.007

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*, *59*(3), 2142-2154. https://doi.org/10.1016/j.neuroimage.2011.10.018

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., Schlaggar, B. L., & Petersen, S. E. (2011). Functional network organization of the human brain. *Neuron*, *72*(4), 665-678. https://doi.org/10.1016/j.neuron.2011.09.006

Power, J. D., Lynch, C. J., Dubin, M. J., Silver, B. M., Martin, A., & Jones, R. M. (2020). Characteristics of respiratory measures in young adults scanned at rest, including systematic changes and "missed" deep breaths. *Neuroimage*, *204*, 116234. https://doi.org/10.1016/j.neuroimage.2019.116234

Power, J. D., Lynch, C. J., Gilmore, A. W., Gotts, S. J., & Martin, A. (2019). Reply to Spreng et al.: Multiecho fMRI denoising does not remove global motion-associated respiratory signals. *Proc Natl Acad Sci U S A*, *116*(39), 19243-19244. https://doi.org/10.1073/pnas.1909852116

Power, J. D., Plitt, M., Gotts, S. J., Kundu, P., Voon, V., Bandettini, P. A., & Martin, A. (2018). Ridding fMRI data of motion-related influences: Removal of signals with distinct spatial and physical bases in multiecho data. *Proc Natl Acad Sci U S A*, *115*(9), E2105-E2114. https://doi.org/10.1073/pnas.1720985115

Power, J. D., Plitt, M., Laumann, T. O., & Martin, A. (2017). Sources and implications of whole-brain fMRI signals in humans. *Neuroimage*, *146*, 609-625. https://doi.org/10.1016/j.neuroimage.2016.09.038

Power, J. D., Schlaggar, B. L., & Petersen, S. E. (2015). Recent progress and outstanding issues in motion correction in resting state fMRI. *Neuroimage*, *105*, 536-551. https://doi.org/10.1016/j.neuroimage.2014.10.044

Qian, X., Castellanos, F. X., Uddin, L. Q., Loo, B. R. Y., Liu, S., Koh, H. L., Poh, X. W. W., Fung, D., Guan, C., Lee, T. S., Lim, C. G., & Zhou, J. (2019). Large-scale brain functional network topology disruptions underlie symptom heterogeneity in children with attention-deficit/hyperactivity disorder. *Neuroimage Clin*, *21*, 101600. https://doi.org/10.1016/j.nicl.2018.11.010

Raine, A. (1991). The SPQ: a scale for the assessment of schizotypal personality based on DSM-III-R criteria. *Schizophr Bull*, *17*(4), 555-564. https://doi.org/10.1093/schbul/17.4.555

Rammstedt, B., Kemper, C. J., Klein, M. C., Beierlein, C. & Kovaleva, A. (2012). Eine kurze Skala zur Messung der fünf Dimensionen der Persönlichkeit: Big-Five-Inventory-10 (BFI-10). *Gesis, Working Papers*.

Ramot, M., Grossman, S., Friedman, D., & Malach, R. (2016). Covert neurofeedback without awareness shapes cortical network spontaneous connectivity. *Proc Natl Acad Sci U S A*, *113*(17), E2413-2420. https://doi.org/10.1073/pnas.1516857113

Ramot, M., Kimmich, S., Gonzalez-Castillo, J., Roopchansingh, V., Popal, H., White, E., Gotts, S. J., & Martin, A. (2017). Direct modulation of aberrant brain network connectivity through real-time NeuroFeedback. *Elife*, *6*. https://doi.org/10.7554/eLife.28974

Rance, M., Walsh, C., Sukhodolsky, D. G., Pittman, B., Qiu, M., Kichuk, S. A., Wasylink, S., Koller, W. N., Bloch, M., Gruner, P., Scheinost, D., Pittenger, C., & Hampson, M. (2018). Time course of clinical change following neurofeedback. *Neuroimage*, *181*, 807-813. https://doi.org/10.1016/j.neuroimage.2018.05.001

Rangaprakash, D., Wu, G. R., Marinazzo, D., Hu, X., & Deshpande, G. (2018). Hemodynamic response function (HRF) variability confounds resting-state fMRI functional connectivity. *Magn Reson Med*, *80*(4), 1697-1713. https://doi.org/10.1002/mrm.27146

Retter, T. L., Jiang, F., Webster, M. A., Michel, C., Schiltz, C., & Rossion, B. (2021). Varying Stimulus Duration Reveals Consistent Neural Activity and Behavior for Human Face Individuation. *Neuroscience*, *472*, 138-156. https://doi.org/10.1016/j.neuroscience.2021.07.025

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*(5695), 443-447. https://doi.org/10.1126/science.1100301

Robbins, T. W. (1990). The case of frontostriatal dysfunction in schizophrenia. *Schizophr Bull*, *16*(3), 391-402. https://doi.org/10.1093/schbul/16.3.391

Rockstroh, B., Elbert, T., Birbaumer, N., & Lutzenberger, W. (1990). Biofeedback-produced hemispheric asymmetry of slow cortical potentials and its behavioural effects. *Int J Psychophysiol*, *9*(2), 151-165. https://doi.org/10.1016/0167-8760(90)90069-p

Ros, T., Enriquez-Geppert, S., Zotev, V., Young, K. D., Wood, G., Whitfield-Gabrieli, S., Wan, F., Vuilleumier, P., Vialatte, F., Van De Ville, D., Todder, D., Surmeli, T., Sulzer, J. S., Strehl, U., Sterman, M. B., Steiner, N. J., Sorger, B., Soekadar, S. R., Sitaram, R., Sherlin, L. H., Schonenberg, M., Scharnowski, F., Schabus, M., Rubia, K., Rosa, A., Reiner, M., Pineda, J. A., Paret, C., Ossadtchi, A.,

Nicholson, A. A., Nan, W., Minguez, J., Micoulaud-Franchi, J. A., Mehler, D. M. A., Luhrs, M., Lubar, J., Lotte, F., Linden, D. E. J., Lewis-Peacock, J. A., Lebedev, M. A., Lanius, R. A., Kubler, A., Kranczioch, C., Koush, Y., Konicar, L., Kohl, S. H., Kober, S. E., Klados, M. A., Jeunet, C., Janssen, T. W. P., Huster, R. J., Hoedlmoser, K., Hirshberg, L. M., Heunis, S., Hendler, T., Hampson, M., Guggisberg, A. G., Guggenberger, R., Gruzelier, J. H., Gobel, R. W., Gninenko, N., Gharabaghi, A., Frewen, P., Fovet, T., Fernandez, T., Escolano, C., Ehlis, A. C., Drechsler, R., Christopher deCharms, R., Debener, S., De Ridder, D., Davelaar, E. J., Congedo, M., Cavazza, M., Breteler, M. H. M., Brandeis, D., Bodurka, J., Birbaumer, N., Bazanova, O. M., Barth, B., Bamidis, P. D., Auer, T., Arns, M., & Thibault, R. T. (2020). Consensus on the reporting and experimental design of clinical and cognitive-behavioural neurofeedback studies (CRED-nf checklist). *Brain*, *143*(6), 1674-1685. https://doi.org/10.1093/brain/awaa009

Ruiz, S., Buyukturkoglu, K., Rana, M., Birbaumer, N., & Sitaram, R. (2014). Real-time fMRI brain computer interfaces: self-regulation of single brain regions to networks. *Biol Psychol*, *95*, 4-20. https://doi.org/10.1016/j.biopsycho.2013.04.010

Ruiz, S., Lee, S., Soekadar, S. R., Caria, A., Veit, R., Kircher, T., Birbaumer, N., & Sitaram, R. (2013). Acquired self-control of insula cortex modulates emotion recognition and brain network connectivity in schizophrenia. *Hum Brain Mapp*, *34*(1), 200-212. https://doi.org/10.1002/hbm.21427

Saad, Z. S., Gotts, S. J., Murphy, K., Chen, G., Jo, H. J., Martin, A., & Cox, R. W. (2012). Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. *Brain Connect*, *2*(1), 25-32. https://doi.org/10.1089/brain.2012.0080

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X. N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex*, *28*(9), 3095-3114. https://doi.org/10.1093/cercor/bhx179

Schafer, R. J., & Moore, T. (2011). Selective attention from voluntary control of neurons in prefrontal cortex. *Science*, *332*(6037), 1568-1571. https://doi.org/10.1126/science.1199892

Scharnowski, F., Veit, R., Zopf, R., Studer, P., Bock, S., Diedrichsen, J., Goebel, R., Mathiak, K., Birbaumer, N., & Weiskopf, N. (2015). Manipulating motor performance and memory through real-time fMRI neurofeedback. *Biol Psychol*, *108*, 85-97. https://doi.org/10.1016/j.biopsycho.2015.03.009

Scheinost, D., Hsu, T. W., Avery, E. W., Hampson, M., Constable, R. T., Chun, M. M., & Rosenberg, M. D. (2020). Connectome-based neurofeedback: A pilot study to improve sustained attention. *Neuroimage*, *212*, 116684. https://doi.org/10.1016/j.neuroimage.2020.116684

Scheinost, D., Stoica, T., Saksa, J., Papademetris, X., Constable, R. T., Pittenger, C., & Hampson, M. (2013). Orbitofrontal cortex neurofeedback produces lasting changes in contamination anxiety and resting-state connectivity. *Transl Psychiatry*, *3*, e250. https://doi.org/10.1038/tp.2013.24

Scheinost, D., Stoica, T., Wasylink, S., Gruner, P., Saksa, J., Pittenger, C., & Hampson, M. (2014). Resting state functional connectivity predicts neurofeedback response. *Front Behav Neurosci*, *8*, 338. https://doi.org/10.3389/fnbeh.2014.00338

Schmidt, H. D., Shelton, R. C., & Duman, R. S. (2011). Functional biomarkers of depression: diagnosis, treatment, and pathophysiology.

Neuropsychopharmacology, *36*(12), 2375-2394. https://doi.org/10.1038/npp.2011.151

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593-1599. https://doi.org/10.1126/science.275.5306.1593

Seeley, W. W., Crawford, R. K., Zhou, J., Miller, B. L., & Greicius, M. D. (2009). Neurodegenerative diseases target large-scale human brain networks. *Neuron*, *62*(1), 42-52. https://doi.org/10.1016/j.neuron.2009.03.024

Sepulveda, P., Sitaram, R., Rana, M., Montalba, C., Tejos, C., & Ruiz, S. (2016). How feedback, motor imagery, and reward influence brain self-regulation using real-time fMRI. *Hum Brain Mapp*, *37*(9), 3153-3171. https://doi.org/10.1002/hbm.23228

Sheline, Y. I., Barch, D. M., Donnelly, J. M., Ollinger, J. M., Snyder, A. Z., & Mintun, M. A. (2001). Increased amygdala response to masked emotional faces in depressed subjects resolves with antidepressant treatment: an fMRI study. *Biol Psychiatry*, *50*(9), 651-658. https://doi.org/10.1016/s0006-3223(01)01263-x

Shipp, S. (2017). The functional logic of corticostriatal connections. *Brain Struct Funct*, *222*(2), 669-706. https://doi.org/10.1007/s00429-016-1250-9

Shukla, D. K., Chiappelli, J. J., Sampath, H., Kochunov, P., Hare, S. M., Wisner, K., Rowland, L. M., & Hong, L. E. (2019). Aberrant Frontostriatal Connectivity in Negative Symptoms of Schizophrenia. *Schizophr Bull*, *45*(5), 1051-1059. https://doi.org/10.1093/schbul/sby165

Sitaram, R., Weiskopf, N., Caria, A., Veit, R., Erb, M. & Birbaumer, N. (2008). fMRI Brain-Computer Interfaces. A tutorial on methods and applications. *IEEE SIGNAL PROCESSING MAGAZINE, 25(1),* 95-106. https://doi.org/10.1109/MSP.2008.4408446

Sitaram, R., Ros, T., Stoeckel, L., Haller, S., Scharnowski, F., Lewis-Peacock, J., Weiskopf, N., Blefari, M. L., Rana, M., Oblak, E., Birbaumer, N., & Sulzer, J. (2017). Closed-loop brain training: the science of neurofeedback. *Nat Rev Neurosci*, *18*(2), 86-100. https://doi.org/10.1038/nrn.2016.164

Skinner, B.F. (1938). *The behavior of organisms: An experimental analysis*. Appleton-Century.

Sorger, B., Scharnowski, F., Linden, D. E. J., Hampson, M., & Young, K. D. (2019). Control freaks: Towards optimal selection of control conditions for fMRI neurofeedback studies. *Neuroimage*, *186*, 256-265. https://doi.org/10.1016/j.neuroimage.2018.11.004

Spreng, R. N., Fernandez-Cabello, S., Turner, G. R., & Stevens, W. D. (2019). Take a deep breath: Multiecho fMRI denoising effectively removes head motion artifacts, obviating the need for global signal regression. *Proc Natl Acad Sci U S A*, *116*(39), 19241-19242. https://doi.org/10.1073/pnas.1909848116

Sreedharan, S., Arun, K. M., Sylaja, P. N., Kesavadas, C., & Sitaram, R. (2019). Functional Connectivity of Language Regions of Stroke Patients with Expressive Aphasia During Real-Time Functional Magnetic Resonance Imaging Based Neurofeedback. *Brain Connect*, *9*(8), 613-626. https://doi.org/10.1089/brain.2019.0674

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

Su, T. W., Lan, T. H., Hsu, T. W., Biswal, B. B., Tsai, P. J., Lin, W. C., & Lin, C. P. (2013). Reduced neuro-integration from the dorsolateral prefrontal cortex to the whole brain and executive dysfunction in schizophrenia patients and their

relatives. *Schizophr Res*, *148*(1-3), 50-58. https://doi.org/10.1016/j.schres.2013.05.005

Sugiura, M., Kawashima, R., Nakagawa, M., Okada, K., Sato, T., Goto, R., Sato, K., Ono, S., Schormann, T., Zilles, K., & Fukuda, H. (2000). Correlation between human personality and neural activity in cerebral cortex. *Neuroimage*, *11*(5 Pt 1), 541-546. https://doi.org/10.1006/nimg.2000.0564

Sukhodolsky, D. G., Walsh, C., Koller, W. N., Eilbott, J., Rance, M., Fulbright, R. K., Zhao, Z., Bloch, M. H., King, R., Leckman, J. F., Scheinost, D., Pittman, B., & Hampson, M. (2020). Randomized, Sham-Controlled Trial of Real-Time Functional Magnetic Resonance Imaging Neurofeedback for Tics in Adolescents With Tourette Syndrome. *Biol Psychiatry*, *87*(12), 1063-1070. https://doi.org/10.1016/j.biopsych.2019.07.035

Sulzer, J., Haller, S., Scharnowski, F., Weiskopf, N., Birbaumer, N., Blefari, M. L., Bruehl, A. B., Cohen, L. G., DeCharms, R. C., Gassert, R., Goebel, R., Herwig, U., LaConte, S., Linden, D., Luft, A., Seifritz, E., & Sitaram, R. (2013). Real-time fMRI neurofeedback: progress and challenges. *Neuroimage*, *76*, 386-399. https://doi.org/10.1016/j.neuroimage.2013.03.033

Swick, D., & Turken, A. U. (2002). Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proc Natl Acad Sci U S A*, *99*(25), 16354-16359. https://doi.org/10.1073/pnas.252521499

Sylvester, C. Y., Wager, T. D., Lacey, S. C., Hernandez, L., Nichols, T. E., Smith, E. E., & Jonides, J. (2003). Switching attention and resolving interference: fMRI measures of executive functions. *Neuropsychologia*, *41*(3), 357-370. https://doi.org/10.1016/s0028-3932(02)00167-7

Taschereau-Dumouchel, V., Cortese, A., Chiba, T., Knotts, J. D., Kawato, M., & Lau, H. (2018). Towards an unconscious neural reinforcement intervention for common fears. *Proc Natl Acad Sci U S A*, *115*(13), 3470-3475. https://doi.org/10.1073/pnas.1721572115

Tekin, S., & Cummings, J. L. (2002). Frontal-subcortical neuronal circuits and clinical neuropsychiatry: an update. *J Psychosom Res*, *53*(2), 647-654. https://doi.org/10.1016/s0022-3999(02)00428-2

Thibault, R. T., MacPherson, A., Lifshitz, M., Roth, R. R., & Raz, A. (2018). Neurofeedback with fMRI: A critical systematic review. *Neuroimage*, *172*, 786-807. https://doi.org/10.1016/j.neuroimage.2017.12.071

Tian, L., Jiang, T., Liang, M., Zang, Y., He, Y., Sui, M., & Wang, Y. (2008). Enhanced resting-state brain activities in ADHD patients: a fMRI study. *Brain Dev*, *30*(5), 342-348. https://doi.org/10.1016/j.braindev.2007.10.005

Trambaiolli, L. R., Kohl, S. H., Linden, D. E. J., & Mehler, D. M. A. (2021). Neurofeedback training in major depressive disorder: A systematic review of clinical efficacy, study quality and reporting practices. *Neurosci Biobehav Rev*, *125*, 33-56. https://doi.org/10.1016/j.neubiorev.2021.02.015

Tsuchiyagaito, A., Misaki, M., Zoubi, O. A., Tulsa, I., Paulus, M., & Bodurka, J. (2021). Prevent breaking bad: A proof of concept study of rebalancing the brain's rumination circuit with real-time fMRI functional connectivity neurofeedback. *Hum Brain Mapp*, *42*(4), 922-940. https://doi.org/10.1002/hbm.25268

Tursic, A., Eck, J., Luhrs, M., Linden, D. E. J., & Goebel, R. (2020). A systematic review of fMRI neurofeedback reporting and effects in clinical populations. *Neuroimage Clin*, *28*, 102496. https://doi.org/10.1016/j.nicl.2020.102496

Van Dijk, K. R., Sabuncu, M. R., & Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage*, *59*(1), 431-438. https://doi.org/10.1016/j.neuroimage.2011.07.044

Viher, P. V., Docx, L., Van Hecke, W., Parizel, P. M., Sabbe, B., Federspiel, A., Walther, S., & Morrens, M. (2019). Aberrant fronto-striatal connectivity and fine motor function in schizophrenia. *Psychiatry Res Neuroimaging*, *288*, 44-50. https://doi.org/10.1016/j.pscychresns.2019.04.010

Wang, W., Worhunsky, P. D., Zhang, S., Le, T. M., Potenza, M. N., & Li, C. R. (2018). Response inhibition and fronto-striatal-thalamic circuit dysfunction in cocaine addiction. *Drug Alcohol Depend*, *192*, 137-145. https://doi.org/10.1016/j.drugalcdep.2018.07.037

Watanabe, T., Sasaki, Y., Shibata, K., & Kawato, M. (2017). Advances in fMRI Real-Time Neurofeedback. *Trends Cogn Sci*, *21*(12), 997-1010. https://doi.org/10.1016/j.tics.2017.09.010

Weiskopf, N. (2012). Real-time fMRI and its application to neurofeedback. *Neuroimage*, *62*(2), 682-692. https://doi.org/10.1016/j.neuroimage.2011.10.009

Weiskopf, N., Scharnowski, F., Veit, R., Goebel, R., Birbaumer, N., & Mathiak, K. (2004). Self-regulation of local brain activity using real-time functional magnetic resonance imaging (fMRI). *J Physiol Paris*, *98*(4-6), 357-373. https://doi.org/10.1016/j.jphysparis.2005.09.019

Weiskopf, N., Veit, R., Erb, M., Mathiak, K., Grodd, W., Goebel, R., & Birbaumer, N. (2003). Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fMRI): methodology and exemplary data. *Neuroimage*, *19*(3), 577-586. https://doi.org/10.1016/s1053-8119(03)00145-9

Weiss, F., Aslan, A., Zhang, J., Gerchen, M. F., Kiefer, F., & Kirsch, P. (2020). Using mind control to modify cue-reactivity in AUD: the impact of mindfulness-based relapse prevention on real-time fMRI neurofeedback to modify cue-reactivity in alcohol use disorder: a randomized controlled trial. *BMC Psychiatry*, *20*(1), 309. https://doi.org/10.1186/s12888-020-02717-7

Weiss, F., Zamoscik, V., Schmidt, S. N. L., Halli, P., Kirsch, P., & Gerchen, M. F. (2020). Just a very expensive breathing training? Risk of respiratory artefacts in functional connectivity-based real-time fMRI neurofeedback. *Neuroimage*, *210*, 116580. https://doi.org/10.1016/j.neuroimage.2020.116580

Wolfenden, L., Ezzati, M., Larijani, B., & Dietz, W. (2019). The challenge for global health systems in preventing and managing obesity. *Obes Rev*, *20 Suppl 2*, 185-193. https://doi.org/10.1111/obr.12872

Wu, T., Jia, X., Shi, H., Niu, J., Yin, X., Xie, J., & Wang, X. (2021). Prevalence of mental health problems during the COVID-19 pandemic: A systematic review and meta-analysis. *J Affect Disord*, *281*, 91-98. https://doi.org/10.1016/j.jad.2020.11.117

Yamashita, A., Hayasaka, S., Kawato, M., & Imamizu, H. (2017). Connectivity Neurofeedback Training Can Differentially Change Functional Connectivity and Cognitive Performance. *Cereb Cortex*, *27*(10), 4960-4970. https://doi.org/10.1093/cercor/bhx177

Yang, G. J., Murray, J. D., Repovs, G., Cole, M. W., Savic, A., Glasser, M. F., Pittenger, C., Krystal, J. H., Wang, X. J., Pearlson, G. D., Glahn, D. C., & Anticevic, A. (2014). Altered global brain signal in schizophrenia. *Proc Natl Acad Sci U S A*, *111*(20), 7438-7443. https://doi.org/10.1073/pnas.1405289111

Yang, Z., Gu, S., Honnorat, N., Linn, K. A., Shinohara, R. T., Aselcioglu, I., Bruce, S., Oathes, D. J., Davatzikos, C., Satterthwaite, T. D., Bassett, D. S., & Sheline, Y. I. (2018). Network changes associated with transdiagnostic depressive symptom improvement following cognitive behavioral therapy in MDD and

PTSD. *Mol Psychiatry*, *23*(12), 2314-2323. https://doi.org/10.1038/s41380-018-0201-7

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods*, *8*(8), 665-670. https://doi.org/10.1038/nmeth.1635

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zollei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol*, *106*(3), 1125-1165. https://doi.org/10.1152/jn.00338.2011

Yoo, J. J., Hinds, O., Ofen, N., Thompson, T. W., Whitfield-Gabrieli, S., Triantafyllou, C., & Gabrieli, J. D. (2012). When the brain is prepared to learn: enhancing human learning using real-time fMRI. *Neuroimage*, *59*(1), 846-852. https://doi.org/10.1016/j.neuroimage.2011.07.063

Yoo, S. S., & Jolesz, F. A. (2002). Functional MRI for neurofeedback: feasibility study on a hand motor task. *Neuroreport*, *13*(11), 1377-1381. https://doi.org/10.1097/00001756-200208070-00005

Young, K. D., Prause, N., Lazzaro, S., & Siegle, G. J. (2020). Low Cost MR Compatible Haptic Stimulation with Application to fMRI Neurofeedback. *Brain Sci*, *10*(11). https://doi.org/10.3390/brainsci10110790

Young, K. D., Siegle, G. J., Misaki, M., Zotev, V., Phillips, R., Drevets, W. C., & Bodurka, J. (2018). Altered task-based and resting-state amygdala functional connectivity following real-time fMRI amygdala neurofeedback training in major depressive disorder. *Neuroimage Clin*, *17*, 691-703. https://doi.org/10.1016/j.nicl.2017.12.004

Young, K. D., Siegle, G. J., Zotev, V., Phillips, R., Misaki, M., Yuan, H., Drevets, W. C., & Bodurka, J. (2017). Randomized Clinical Trial of Real-Time fMRI Amygdala Neurofeedback for Major Depressive Disorder: Effects on Symptoms and Autobiographical Memory Recall. *Am J Psychiatry*, *174*(8), 748-755. https://doi.org/10.1176/appi.ajp.2017.16060637

Young, K. D., Zotev, V., Phillips, R., Misaki, M., Yuan, H., Drevets, W. C., & Bodurka, J. (2014). Real-time FMRI neurofeedback training of amygdala activity in patients with major depressive disorder. *PLoS One*, *9*(2), e88785. https://doi.org/10.1371/journal.pone.0088785

Zamoscik, V., Niemeyer, C., Gerchen, M. F., Fenske, S. C., Witthoft, M., & Kirsch, P. (2017). [Sensory Inventory (SI): self-assessment of sensory sensitivity for adults and adolescents]. *Fortschr Neurol Psychiatr*, *85*(9), 541-551. https://doi.org/10.1055/s-0043-117885 (Sensorik Inventar (SI) - Selbstbeurteilung der sensorischen Sensitivitat fur Erwachsene und Jugendliche.)

Zamoscik, V. E., Schmidt, S. N. L., Gerchen, M. F., Samsouris, C., Timm, C., Kuehner, C., & Kirsch, P. (2018). Respiration pattern variability and related default mode network connectivity are altered in remitted depression. *Psychol Med*, *48*(14), 2364-2374. https://doi.org/10.1017/S0033291717003890

Zarahn, E., Aguirre, G. K., & D'Esposito, M. (1997). Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. *Neuroimage*, *5*(3), 179-197. https://doi.org/10.1006/nimg.1997.0263

Zhang, J., Cheng, W., Liu, Z., Zhang, K., Lei, X., Yao, Y., Becker, B., Liu, Y., Kendrick, K. M., Lu, G., & Feng, J. (2016). Neural, electrophysiological and anatomical basis of brain-network variability and its characteristic changes in mental disorders. *Brain*, *139*(Pt 8), 2307-2321. https://doi.org/10.1093/brain/aww143

Zhao, Y., Li, J., Liu, X., Song, Y., Wang, R., Yang, Z., & Liu, J. (2016). Altered spontaneous neural activity in the occipital face area reflects behavioral deficits in developmental prosopagnosia. *Neuropsychologia*, *89*, 344-355. https://doi.org/10.1016/j.neuropsychologia.2016.05.027

Zhao, Z., Yao, S., Li, K., Sindermann, C., Zhou, F., Zhao, W., Li, J., Luhrs, M., Goebel, R., Kendrick, K. M., & Becker, B. (2019). Real-Time Functional Connectivity-Informed Neurofeedback of Amygdala-Frontal Pathways Reduces Anxiety. *Psychother Psychosom*, *88*(1), 5-15. https://doi.org/10.1159/000496057

Zhao, Z., Yao, S., Zweerings, J., Zhou, X., Zhou, F., Kendrick, K. M., Chen, H., Mathiak, K., & Becker, B. (2021). Putamen volume predicts real-time fMRI neurofeedback learning success across paradigms and neurofeedback target regions. *Hum Brain Mapp*, *42*(6), 1879-1887. https://doi.org/10.1002/hbm.25336

Zilverstand, A., Sorger, B., Sarkheil, P., & Goebel, R. (2015). fMRI neurofeedback facilitates anxiety regulation in females with spider phobia. *Front Behav Neurosci*, *9*, 148. https://doi.org/10.3389/fnbeh.2015.00148

Zilverstand, A., Sorger, B., Slaats-Willemse, D., Kan, C. C., Goebel, R., & 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. https://doi.org/10.1371/journal.pone.0170795

Zotev, V., Krueger, F., Phillips, R., Alvarez, R. P., Simmons, W. K., Bellgowan, P., Drevets, W. C., & Bodurka, J. (2011). Self-regulation of amygdala activation using real-time FMRI neurofeedback. *PLoS One*, *6*(9), e24522. https://doi.org/10.1371/journal.pone.0024522

Zotev, V., Phillips, R., Misaki, M., Wong, C. K., Wurfel, B. E., Krueger, F., Feldner, M., & Bodurka, J. (2018). Real-time fMRI neurofeedback training of the amygdala activity with simultaneous EEG in veterans with combat-related PTSD. *Neuroimage Clin*, *19*, 106-121. https://doi.org/10.1016/j.nicl.2018.04.010

Zweerings, J., Hummel, B., Keller, M., Zvyagintsev, M., Schneider, F., Klasen, M., & Mathiak, K. (2019). Neurofeedback of core language network nodes modulates connectivity with the default-mode network: A double-blind fMRI neurofeedback study on auditory verbal hallucinations. *Neuroimage*, *189*, 533-542. https://doi.org/10.1016/j.neuroimage.2019.01

8 PUBLIKATIONEN

Weiss, F., Zamoscik, V., Schmidt, S. N. L., Halli, P., Kirsch, P., & Gerchen, M. F. (2020). Just a very expensive breathing training? Risk of respiratory artefacts in functional connectivity-based real-time fMRI neurofeedback. *Neuroimage, 210,* 116580. doi:10.1016/j.neuroimage.2020.116580

Weiss, F., Aslan, A., Zhang, J., Gerchen, M. F., Kiefer, F., & Kirsch, P. (2020). Using mind control to modify cue-reactivity in AUD: the impact of mindfulness-based relapse prevention on real-time fMRI neurofeedback to modify cue-reactivity in alcohol use disorder: a randomized controlled trial. *BMC Psychiatry, 20(1),* 309. https://doi.org/10.1186/s12888-020-02717-7

Gerchen, M. F., Weiss, F., Kirsch, M., Rentsch, A., Halli, P., Kiefer, F., & Kirsch, P. (2021). Dynamic frontostriatal functional peak connectivity (in alcohol use disorder). *Hum Brain Mapp, 42(1)*, 36-46. https://doi.org/10.1002/hbm.25201

Weiss, F., Zhang, J., Aslan, A., Kirsch, P., Gerchen, M.F. (2022). Feasibility of training the dorsolateral prefrontal-striatal network by real-time fMRI neurofeedback. *Sci Rep, 12,* 1669. https://doi.org/10.1038/s41598-022-05675-0

Doñamayor, N., Ebrahimi, C., Arndt, V. A., Weiss, F., Schlagenhauf, F., Endrass, T. (in revision). Goal-directed and habitual control in substance use: State of the art and future directions.

9 LEBENSLAUF

PERSONALIEN

Name und Vorname:	Franziska Weiß
Geburtsdatum:	25.01.1995
Geburtsort:	Moers

SCHULISCHER WERDEGANG

08/2005 – 06/2013	Lise-Meitner-Gymnasium, Geldern Abitur

UNIVERSITÄRER WERDEGANG

WS 2013/2014	Beginn des Bachelor-Studiums (Psychologie) An der Universität Maastricht (NL) Bachelorarbeit: *Amygdala structure and function between obsessive-compulsive disorder patients and healthy controls*
01.07.2016	Bachelor of Science
WS 2016/2017	Beginn des Master-Studiums (Psychologie) An der Universität Maastricht (NL) Masterarbeit: *The relationship between executive functioning and verbal and visual memory in Amyotrophic Lateral Sclerosis*
01.09.2017	Master of Science

BERUFLICHER WERDEGANG

2017 – 2021	Wissenschaftliche Mitarbeiterin (Doktorandin) an der Abteilung Klinische Psychologie, Zentralinstitut für Seelische Gesundheit, Mannheim
01/2021 – 03/2021	Forschungsaufenthalt in der Abteilung für Biomedizinische Bildgebung an der Yale School of Medicine der Yale University

10 DANKSAGUNG

Zuallererst möchte ich mich bei meinem Doktorvater Prof. Dr. Peter Kirsch für das Vertrauen bedanken, meine Doktorarbeit in seiner Abteilung durchführen zu dürfen. Ich verdanke ihm vier spannende Jahre, in denen ich viel von ihm lernen konnte und die Vielschichtigkeit internationaler psychologischer Forschung erfahren durfte. Ein besonderer Dank gilt ihm für die jederzeit mögliche Hilfe, die immer offenstehende Tür und seinen Rückhalt, auf den man sich bei ihm immer verlassen konnte. Ich hatte stets das Gefühl, dass er sich Zeit für mich nimmt und mein Anliegen wichtig war. Seine außerordentliche Unterstützung ermöglichten mir außerdem unvergessliche Forschungserfahrungen in Form von Konferenzen, Begutachtungen und nicht zuletzt einem Forschungsaufenthalt in Yale. Auch in persönlicher Hinsicht war die Zeit eine bereichernde Erfahrung, welche meine wissenschaftlichen Ambitionen stärkte.

Ein großer Dank gilt ebenfalls meinem engen Betreuer Dr. Martin Fungisai Gerchen, der mich mit größtem Engagement betreute und ohne dessen unermüdlichen Einsatz ein Abschluss der Promotion in dieser Form kaum möglich gewesen wäre. Dabei war es nicht nur seine sehr ausgeprägte fachliche Kompetenz, die mich dabei nicht „schwimmen ließ“, sondern insbesondere die Priorität, die er den Dingen einräumte und seine Art, mich immer wieder zu motivieren und die gemeinsamen Ziele zu verdeutlichen. Er war zu jeder Zeit ansprechbar und aufgrund seiner klaren Lösungsorientiertheit und seines Optimismus konnten mögliche Probleme besonders schnell aus der Welt geschafft werden. Außerdem schaffte er eine vertrauensvolle Atmosphäre und hatte immer einen passenden Rat oder ein offenes Ohr. Die fruchtbare Zusammenarbeit mit ihm zeigte mir, was exzellente Strukturierung und tolle Teamarbeit leisten können.

Bei meinen ehemaligen KollegInnen aus der Abteilung für Klinische Psychologie bedanke ich mich ebenfalls. Insbesondere möchte ich dabei Dr. Vera Zamoscik und Ellen Schmucker hervorheben, die mich besonders unterstützten, sowie Jingying Zhang mit der die Arbeit immer Spaß machte. Wir waren zwar nicht allzu lange zusammen am Institut, aber dennoch war mir Vanessa Renner weit über diese gemeinsame Zeit hinaus eine große Unterstützung. Ebenso danke ich Zhimin Yan für seine hilfreichen inhaltlichen Anmerkungen. Auch Acelya Aslan aus der

Suchtforschung gebührt ein großer Dank für ihre emotionale Unterstützung. Das gemeinsame Paper schreiben bleibt mir in schöner Erinnerung.

Ich danke außerdem Frau Prof. Dr. Hampson dafür, dass ich sie in Yale besuchen durfte und für die große Mühe, die sie im Zuge dessen in meine Betreuung steckte. Ich blicke auf eine unvergessliche Erfahrung zurück.

Ausdrücklich bedanken möchte ich mich auch bei allen ProbandInnen für ihre Teilnahme an meinen Studien, was einen wichtigen Beitrag zur Entstehung dieser Arbeit darstellte.

Obwohl fachlich sehr fremd, sind weite Teile dieser Arbeit Maximilian Schipke zu verdanken. Seine bedingungslose Unterstützung hat mich darin bestärkt, auf diesem Weg nicht aufzugeben.

Mein tiefer Dank geht auch an meine Familie für den Rückhalt während dieser Zeit.

Ausdrücklich bedanken möchte ich mich außerdem bei meinen Freunden, insbesondere Laura Dienst und Elisabeth Horschmann für ihre beständige Unterstützung und Motivation.

Michaela Krieger war meine Bezugsperson während des Studiums, hat mich schon während meiner Masterarbeit in Dublin besonders moralisch unterstützt und auch während dieses Lebensabschnittes konnte ich mich immer auf sie verlassen und ich verdanke ihr das ein oder andere Korrekturlesen.

Zu guter Letzt möchte ich mich bei der mittlerweile verstorbenen Oma Mia dafür bedanken, dass sie immer an mich geglaubt hat. Nachdem sie jahrelang begeistert von meinen Bachelor-, und Masterverleihungen erzählte, wäre es mir eine Freude gewesen, sie auch an meiner Promotionsfeier teilhaben zu lassen. Ihr soll diese Arbeit gewidmet sein.

11 SUPPLEMENT

11.1 Pre-registrations

11.1.1 NeCoSchi OSF (https://osf.io/bejcq)

Study Information

Title

Regulation of Large-Scale Frontostriatal Brain Network Connectivity by Real-Time fMRI Neurofeedback

Research Questions

1. Are healthy controls able to upregulate brain network functional connectivity between dorsolateral prefrontal cortex (DLPFC) and anterior cingulate cortex (ACC) and the striatum during real-time fMRI (rtfMRI) neurofeedback?

2. Are healthy controls able to learn upregulation of network functional connectivity between DLPFC and ACC and the striatum by real-time fMRI (rtfMRI) neurofeedback?

3. Is the ability to upregulate network functional connectivity between DLPFC and ACC and the striatum associated with schizotypal personality traits in healthy controls?

Hypotheses

1. Averaged correlations between DLPFC and ACC and the striatum are higher in the real neurofeedback group in comparison to the yoke control group during rtfMRI neurofeedback sessions. (directional)

2. The increase of averaged correlations between DLPFC and ACC and the striatum from the initial resting state period to the transfer run is higher in the real neurofeedback group in comparison to the yoke control group. (directional)

3.1 The subscales of the SPQ questionnaire are correlated with DLPFC/ACC-Striatum network functional connectivity during neurofeedback sessions.

3.2 The subscales of the SPQ questionnaire are correlated with changes in DLPFC/ACC-Striatum network functional connectivity between the initial resting state scan and the transfer run.

Sampling Plan

Existing Data

Registration prior to creation of data

Explanation of existing data

No data existed prior to registration.

Data collection procedures

We will recruit healthy human participants from the population of Mannheim, Germany. Healthy participants will be between 18 and 65 years, eligible for MRI scanning, and free of neurological or psychiatric disorders. The study will be explained in detail before participation and participants will provide written informed consent. Scanning will be conducted in a single MRI scanning session of 1.5 hours and the participants will receive 20€ compensation for participation.

Sample size

The study will comprise N=40 healthy human subjects who will be randomly assigned to an experimental neurofeedback (n=20) or a yoke control group (n=20).

Sample size rationale

The study is planned as a pilot study and uses a sample size of n=20 per group to ensure a minimal group size suitable for fMRI data analysis. It will itself serve as a basis for power calculations for follow-up studies.

Stopping rule

Not applicable.

Variables

Manipulated variables

The sample will be randomly split into a real real-time fMRI (rtfRMI) neurofeedback group and a yoke control group who will receive feedback of a participant from the other group. The feedback signal will show averaged functional connectivity between bilateral dorsolateral prefrontal cortex (DLPFC) and anterior cingulate cortex (ACC) and the striatum in a sliding window of the last 30 scans (49.2 s). Subjects in both groups will be instructed to upregulate the presented signal.

Measured variables

First, we will conduct a resting state fMRI scan with eyes open and a fixation cross presented on the screen. Then, two rtfMRI neurofeedback training runs with feedback about frontostriatal connectivity that should be upregulated will be conducted. Last, a transfer run with only a fixation cross and the instruction to upregulate the brain process as in the two prior neurofeedback runs but without a feedback signal will be measured.

Covariates for analyses will be age and gender. The German version of the Schizotypal Personality Questionnaire (SPQ) will be collected prior to MRI scanning.

Indices

The main measure will be averaged Fisher Z-transformed Pearson correlations between fMRI time series of bilateral DLPFC and ACC and the striatum.

SPQ subscales will be calculated as described in Klein et al. (1997).

Klein, C., Andresen, B., & Jahn, T. (1997). Psychometric assessment of the schizotypal personality according to DSM-III-R criteria: Psychometric properties of an authorized German translation of Raine's "Schizotypal Personality Questionnaire" (SPQ)]. Diagnostica, 43(4), 347-369.

Design Plan

Study type

Experiment - A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized controlled trials.

Blinding

For studies that involve human subjects, they will not know the treatment group to which they have been assigned.

Research personnel who interact directly with the study subjects (either human or non-human subjects) will not be aware of the assigned treatments.

Study design

The sample will be randomly split into a real real-time fMRI (rtfRMI) neurofeedback group and a yoke control group.

Randomization

Randomization will be conducted based on the order of inclusion into the study with a predefined randomized sequence with the constraint that the first 3 participants will be assigned to the real neurofeedback group to assure sufficient data for the yoke control procedure. Yoke participants will receive the neurofeedback signal from the oldest saved real neurofeedback run which has not been presented as yoke data. If a participant is assigned to the yoke control group without unused real data being available, it will be switched to the real neurofeedback group.

Analysis Plan

Statistical models

Group differences in the averaged Fisher Z-transformed Person correlations of the DLPFC/ACC-striatum network in the neurofeedback runs will be tested with and ANCOVA/t-tests with age and gender as covariates.

Group differences in the difference of averaged Fisher Z-transformed Person correlations of the DLPFC/ACC-striatum network between the resting state and the transfer runs will be tested with ANCOVA/t-tests with age and gender as covariates.

Association of SPQ subscales with DLPFC/ACC-striatum network functional connectivity will be tested with partial correlations corrected for age and gender.

Transformations

fMRI data will be preprocessed with SPM12 (or a newer version). Preprocessing will include segmentation of the anatomical image and normalization to MNI space. Functional images will be slice-time corrected, realigned to the mean image, co-registered to the anatomical image, normalized to MNI space, resampled, and smoothed. Functional data will be corrected for nuisance signals from white matter and cerebrospinal fluid and movement artifacts. Volumes effected by small movements (greater than 0.5mm) will be identified and scrubbed.

Follow-up analyses

We will test whether DLPFC/ACC-striatum functional network connectivity is associated with the BOLD time course of other voxels in the brain with a voxelwise SPM analysis using the dynamic time course of DLPFC/ACC-striatum connectivity as a predictor.

We will assess whether subnetworks that are consistently upregulated within the DLPFC/ACC-striatum network can be identified by testing for group differences between the neurofeedback and the yoke control group in each connection from the DLPFC and the ACC with the striatum.

Inference criteria

We will use a statistical threshold of p=0.05. For hypothesis 1 & 2 we will use one tailed tests. For hypothesis 3 we will use two tailed tests Bonferroni correct for multiple comparisons.

In the follow-up analyses we will use a corrected cluster-level inference threshold of p=0.05 with a cluster defining threshold of p=0.001 in imaging analyses and a threshold of p=0.05 Bonferroni corrected for the number of tests in subnetwork analyses. Imaging analyses will be conducted with one tailed tests.

Data exclusion

Participants who have incidental findings in their anatomical images that prevent data analysis will be excluded. Furthermore, fMRI runs with more than 3mm movement, more than 3° rotation at any scan, or more than 33% of volumes affected by small movements (greater than 0.5mm) will be excluded from analysis.

Missing data

Missing data will not be interpolated and not be included into the analysis.

Exploratory analysis

- Association of SPQ subscales with potentially identified subnetworks

- Group differences in dynamic and spectral characteristics of DLPFC/ACC-striatum functional network connectivity

- Association of DLPFC/ACC-striatum functional network connectivity with frontostriatal functional connectivity topography

- Prediction of the ability to regulate DLPFC/ACC-striatum functional network

11.1.2 NeCoSchi II OSF (https://osf.io/vdn3j)

Study Information

Hypotheses

1. Averaged correlations between DLPFC and striatum are higher in the real neurofeedback group in comparison to the yoke control group during rtfMRI neurofeedback sessions (directional).

2. Participants from the real neurofeedback group in contrast to participants from the yoke control group will demonstrate a higher increase of averaged correlations between DLPFC and striatum from the initial resting state period (directional).

3.1 The subscales of the Schizotypal Personality questionnaire (SPQ) (Klein et al., 1997) are correlated with DLPFC-striatum network functional connectivity during neurofeedback sessions.

3.2 The subscales of the SPQ are correlated with changes in DLPFC-striatum network functional connectivity between the initial resting state scan and the transfer run.

Design Plan

Study type

Experiment - A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized controlled trials.

Blinding

For studies that involve human subjects, they will not know the treatment group to which they have been assigned.

Personnel who interact directly with the study subjects (either human or non-human subjects) will not be aware of the assigned treatments. (Commonly known as “double blind”)

Is there any additional blinding in this study?

No response

Study design

This is a double-blind randomized yoke controlled study. Participants in this sample will be randomly allocated to either real-time fMRI (rtfMRI) neurofeedback group or a yoke control group.

Randomization

Randomization will be conducted based on the order of inclusion into the study with a predefined randomized sequence with the constraint that the first 3 participants will be assigned to the real neurofeedback group to assure sufficient data for the yoke control procedure. Yoke participants will receive the neurofeedback signal from the oldest saved real neurofeedback run which has not been presented as yoke data. If a participant is assigned to the yoke control group without unused real data being available, it will be switched to the real neurofeedback group.

Registration prior to creation of data

Data collection procedures

Healthy human participants from the area of Mannheim, Germany, aged between 18 and 65 years will be recruited. Participants will be eligible for MRI scanning, and free of neurological or psychiatric disorders. The study will be explained in detail before participation and participants will provide written informed consent. Scanning will be conducted in three MRI scanning sessions of 1.5 hours. The first and the third MRI scanning session will consist of a T1-weighted anatomical MPRAGE scan, a resting state run, two neurofeedback runs and a transfer run, whereas the second MRI scanning session includes three neurofeedback sessions and no transfer run. The participants will receive 50€ compensation for participation.

Sample size

The study will include N=40 healthy human subjects who will be randomly assigned to an experimental neurofeedback (n=20) or a yoke control group (n=20).

Sample size rationale

The study is planned as a pilot study and uses a sample size of n=20 per group to ensure a minimal group size suitable for fMRI data analysis. It will itself serve as a basis for power calculations for follow-up studies.

Stopping rule

Not applicable

Variables

Manipulated variables

The sample will be randomly separated into a real real-time fMRI (rtfMRI) neurofeedback group and a yoke control group who will receive feedback of a participant from the other group. The feedback signal will show averaged functional connectivity between bilateral dorsolateral prefrontal cortex (DLPFC) and the striatum in a sliding window of the last 15 scans. Subjects will be instructed to upregulate the presented signal.

Measured variables

On each training day, first, we will conduct a resting state fMRI scan with eyes open and a fixation cross presented on the screen. Then, two rtfMRI neurofeedback training runs with feedback about frontostriatal connectivity that should be upregulated will be conducted. Last, a transfer run with only a fixation cross and the instruction to upregulate the brain process as in the two prior neurofeedback runs but without a feedback signal will be measured. The second scanning session includes three neurofeedback runs instead of the transfer run. Covariates for analyses will be age and gender. Demographics, the German version of the Schizotypal Personality Questionnaire (SPQ), the Beck Depression Inventory (BDI II), the Sensory Inventory and the 10 Item Big Five Inventory (BFI-10) will be collected prior to MRI scanning.

Indices

The main measure will be averaged Fisher Z-transformed Pearson correlations between fMRI time series of bilateral DLPFC and the striatum.

SPQ subscales will be calculated as described in Klein et al. (1997).

Klein, C., Andresen, B., & Jahn, T. (1997). Psychometric assessment of the schizotypal personality according to DSM-III-R criteria: Psychometric properties of an authorized German translation of Raine's "Schizotypal Personality Questionnaire" (SPQ)]. Diagnostica, 43(4), 347-369.

Analysis Plan

Statistical models

Group differences in the averaged Fisher Z-transformed Person correlations of the DLPFC-striatum network in the neurofeedback runs will be tested with an ANCOVA/t-tests including age and gender as covariates.
Group differences in the difference of averaged Fisher Z-transformed Pearson correlations of the DLPFC-striatum network between the resting state and the transfer runs will be tested with ANCOVA/t-tests with age and gender as covariates.

Association of SPQ subscales with DLPFC/striatum network functional connectivity will be tested with partial correlations corrected for age and gender.

Transformations

fMRI data will be preprocessed with SPM12. Preprocessing will include segmentation of the anatomical image and normalization to MNI space. Functional images will be slice-time corrected, realigned to the mean image, co-registered to the anatomical image, normalized to MNI space, resampled, and smoothed. Functional data will be corrected for nuisance signals from white matter and cerebrospinal fluid and movement artifacts. Volumes effected by small movements (greater than 0.5mm) will be identified and scrubbed.

Inference criteria

We will use a statistical threshold of p=0.05. For hypothesis 1 & 2 we will use one tailed tests. For hypothesis 3 we will use two tailed tests.
In the follow-up analyses we will use a corrected cluster-level inference threshold of p=0.05 with a cluster defining threshold of p=0.001 in imaging analyses and a threshold

of p=0.05 Bonferroni corrected for the number of tests in subnetwork analyses. Imaging analyses will be conducted with one tailed tests.

Data exclusion

Participants who have incidental findings in their anatomical images that prevent data analysis will be excluded. Furthermore, fMRI runs with more than 3mm movement, more than 3° rotation at any scan, or more than 25% of volumes affected by small movements (greater than 0.5mm) will be excluded from analysis.

Missing data

Missing data will not be interpolated and not be included into the analysis.

Exploratory analysis

- Association of SPQ subscales with potentially identified subnetworks

- Group differences in dynamic and spectral characteristics of DLPFC-striatum functional network connectivity

- Association of DLPFC-striatum functional network connectivity with frontostriatal functional connectivity topography

- Prediction of the ability to regulate DLPFC-striatum functional network connectivity by resting state markers

Further exploratory hypotheses:

- The subscales of the sensory inventory (Zamoscik et al., 2017) are correlated with DLPFC-striatum network functional connectivity during neurofeedback sessions

- The Big-Five-Inventory-10 (Rammstedt et al., 2007) is correlated with DLPFC-striatum network functional connectivity furing neurofeedback sessions.

- The BDI II (Hautzinger et al., 2006) is correlated with DLPFC -striatum network functional connectivity during neurofeedback sessions.

11.2 Supplements

11.2.1 Supplementary figures NeCoSchi

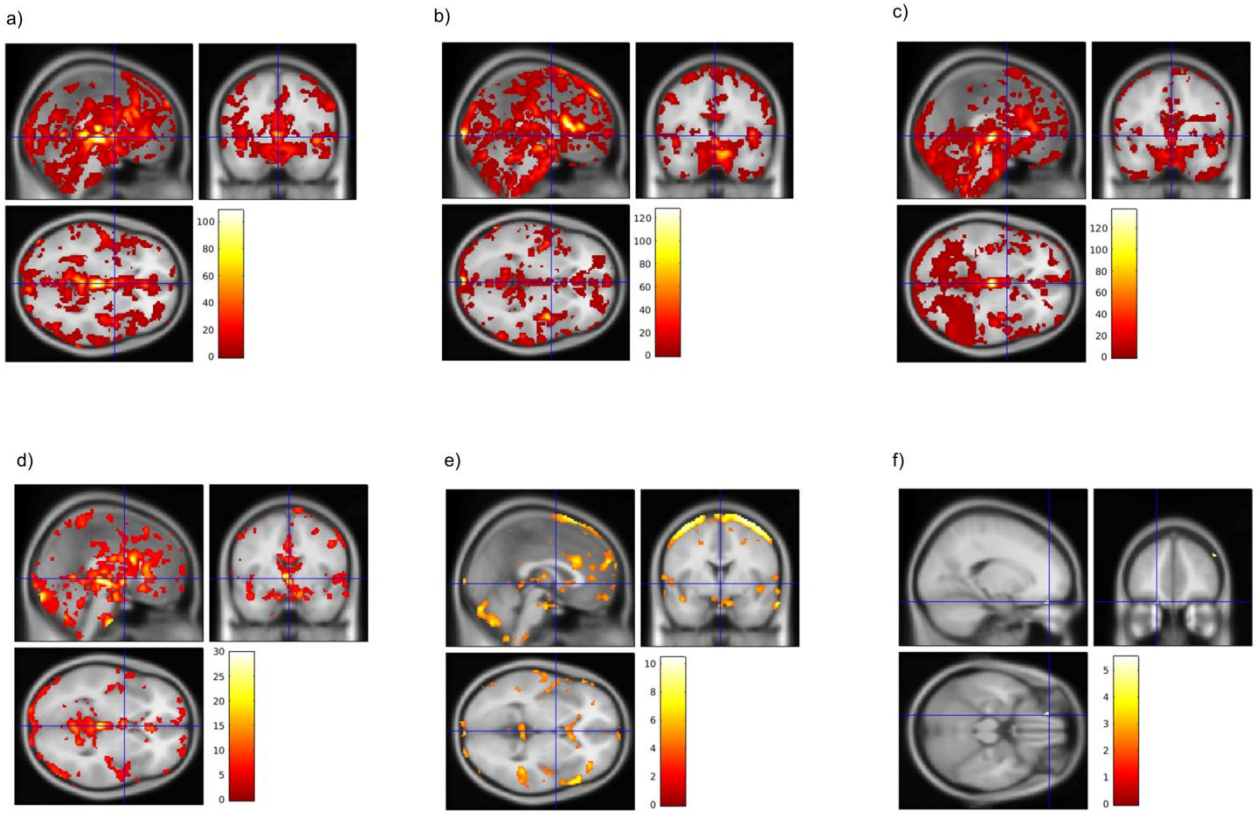


Figure 1. Quality check of model-based physiology regressors obtained with the PhysIO toolbox. Six examples for model-based physiology correction effects during NFB1. a) to c) show the subjects with highest F-values for the F contrast over all physiological nuisance regressors in increased order from left to right. d) to f) represent the subjects with lowest F-values in decreased order from left to right. All other participants lie between the effects depicted in d) and a).

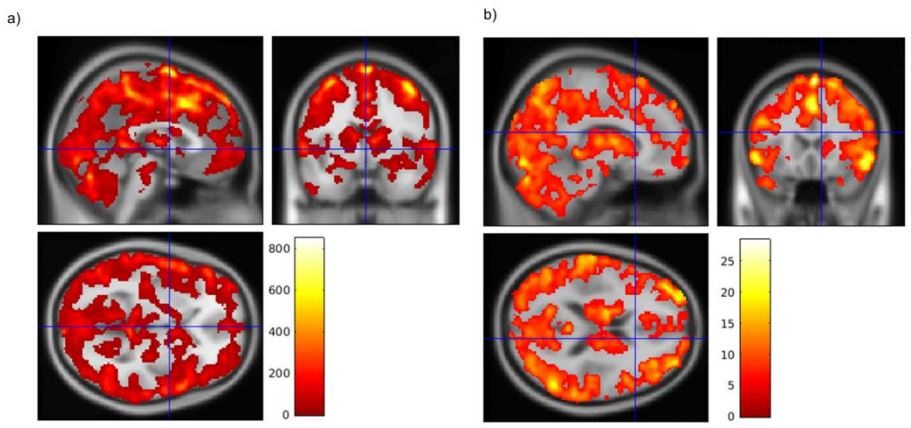


Figure 2. Global signal effect and comparison to model-based physiology effect. Association of the global signal with fMRI data in an example participant not specifically selected (first participant of the data set). a) shows the effect of the global signal, b) shows the results of the t-contrast global signal regressor > model-based nuisance regressors, demonstrating a substantially larger effect of global signal regression.

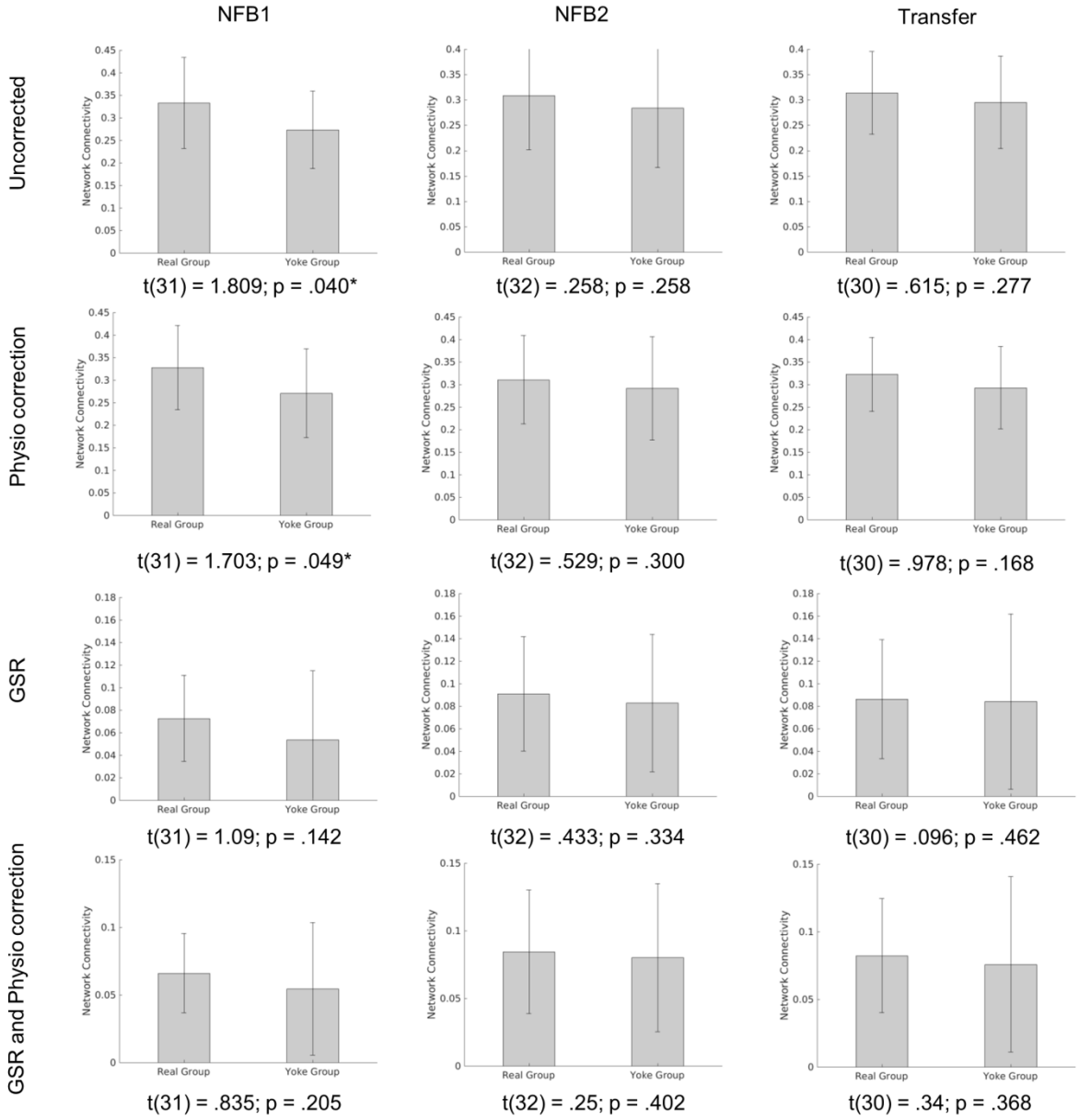


Figure 3. Neurofeedback effects across all runs. Group comparison of network connectivity between the real NFB and the yoke NFB group, uncorrected for physiology, with model-based physiology correction (Physio), with global signal regression (GSR), and with the combination of Physio correction and GSR.

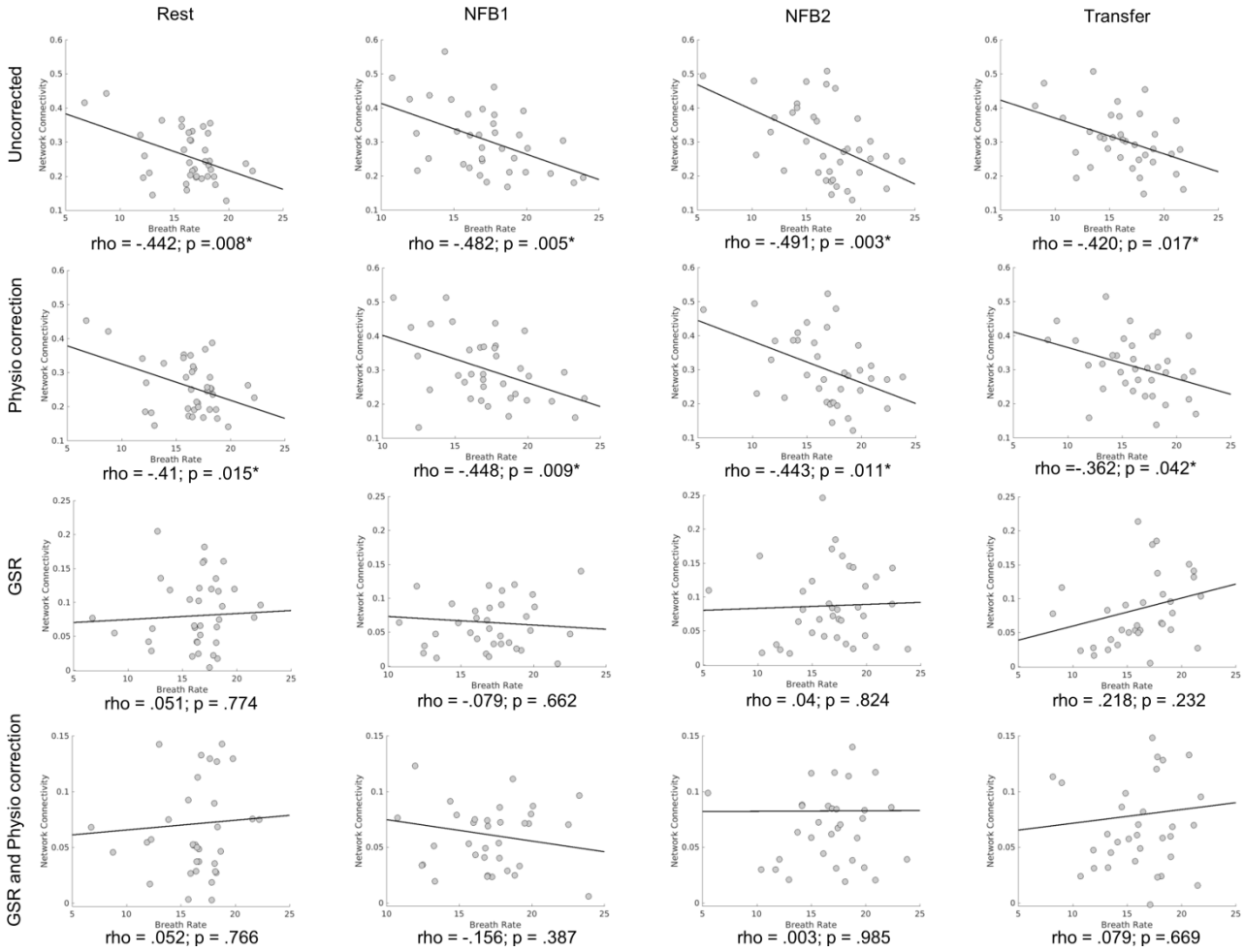


Figure 4. Association of respiratory parameters with network connectivity across all runs. Correlations of the respiratory parameters Breath Rate (peaks/breaths per minute) with target network connectivity for the different physiology corrections. Physio: model-based physiology correction; GSR: global signal regression. Please note that, as expected, GSR overall shifted connectivity estimates towards 0.

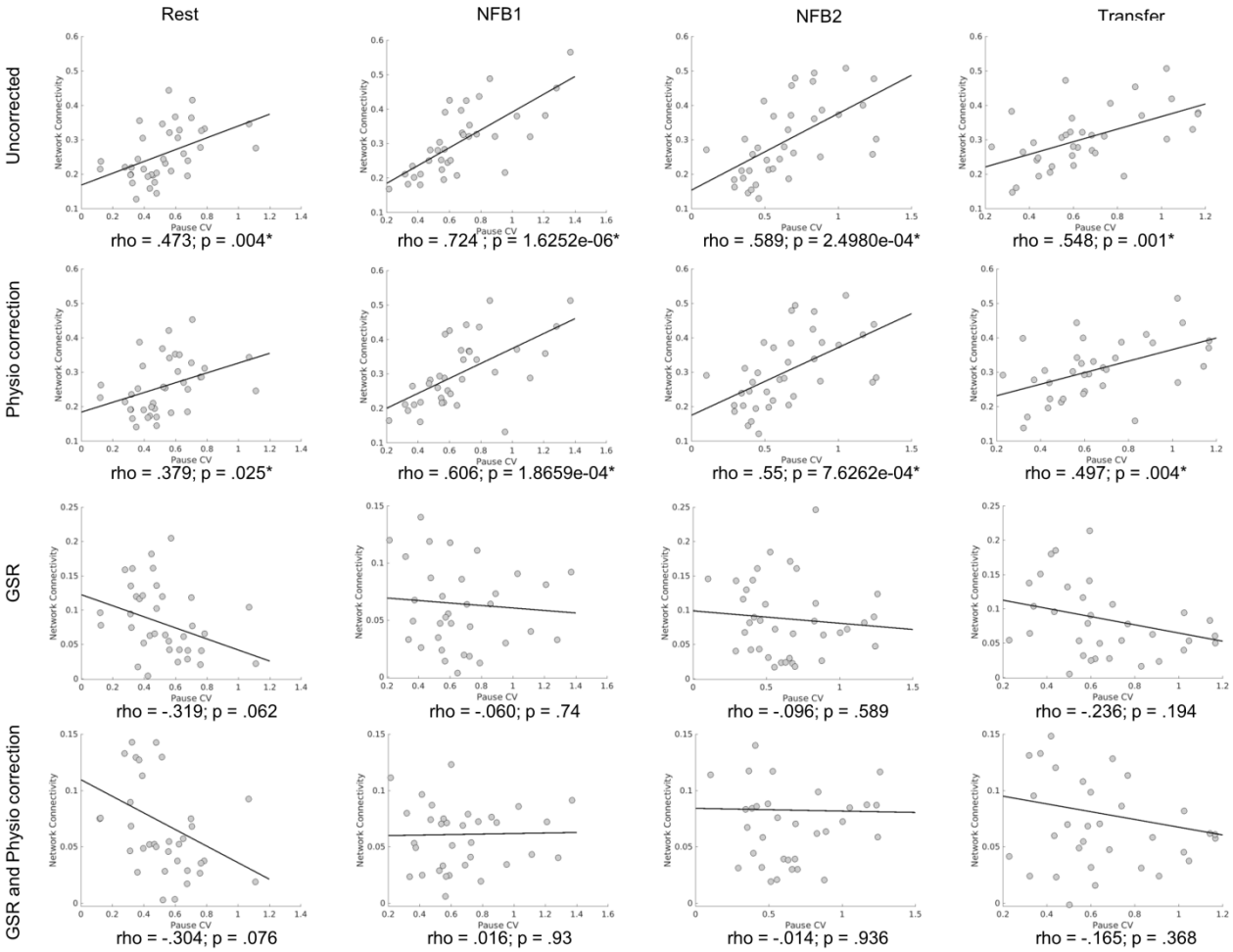


Figure 5. Association of respiratory parameters with network connectivity across all runs. Correlations of the respiratory parameters Pause CV (coefficient of variance of pause duration) with target network connectivity for the different physiology corrections. Please see Supplementary Figure 13 for an illustration how pause durations and the Pause CV parameter were estimated. Physio: model-based physiology correction; GSR: global signal regression. Please note that, as expected, GSR overall shifted connectivity estimates towards 0.

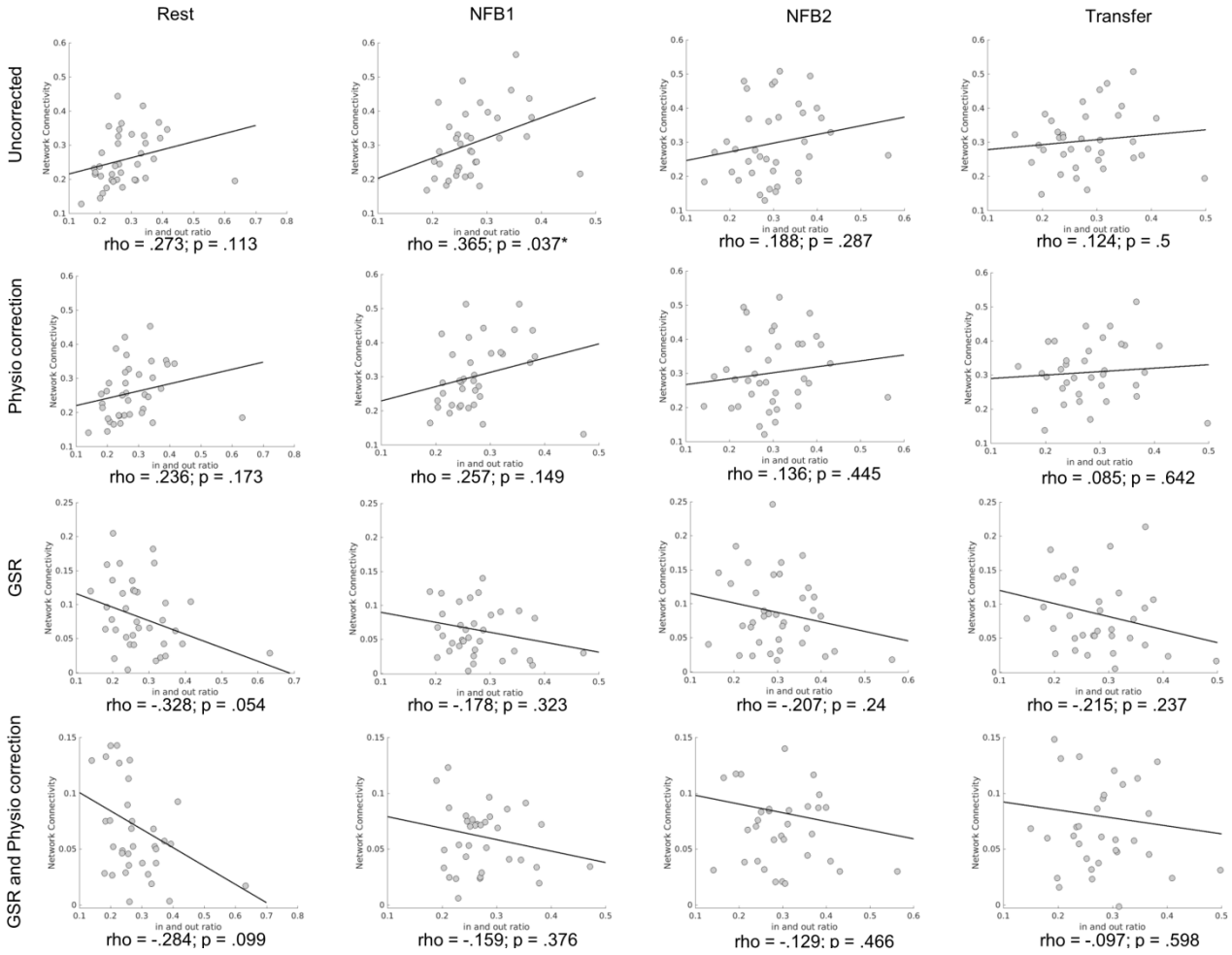


Figure 6. Association of respiratory parameters with network connectivity across all runs. Correlations of the respiratory parameter in and out ratio (expiration-to-inspiration time ratio) (see Zamoscik et al., 2018 for further details) with target network connectivity for the different physiology corrections. Expiration was defined as starting at each maximum peak of the respiratory signal and ending at the lowest local minimum before the next maximum peak, and correspondingly, inspiration was defined as the opposite. Physio: model-based physiology correction; GSR: global signal regression. Please note that, as expected, GSR overall shifted connectivity estimates towards 0.

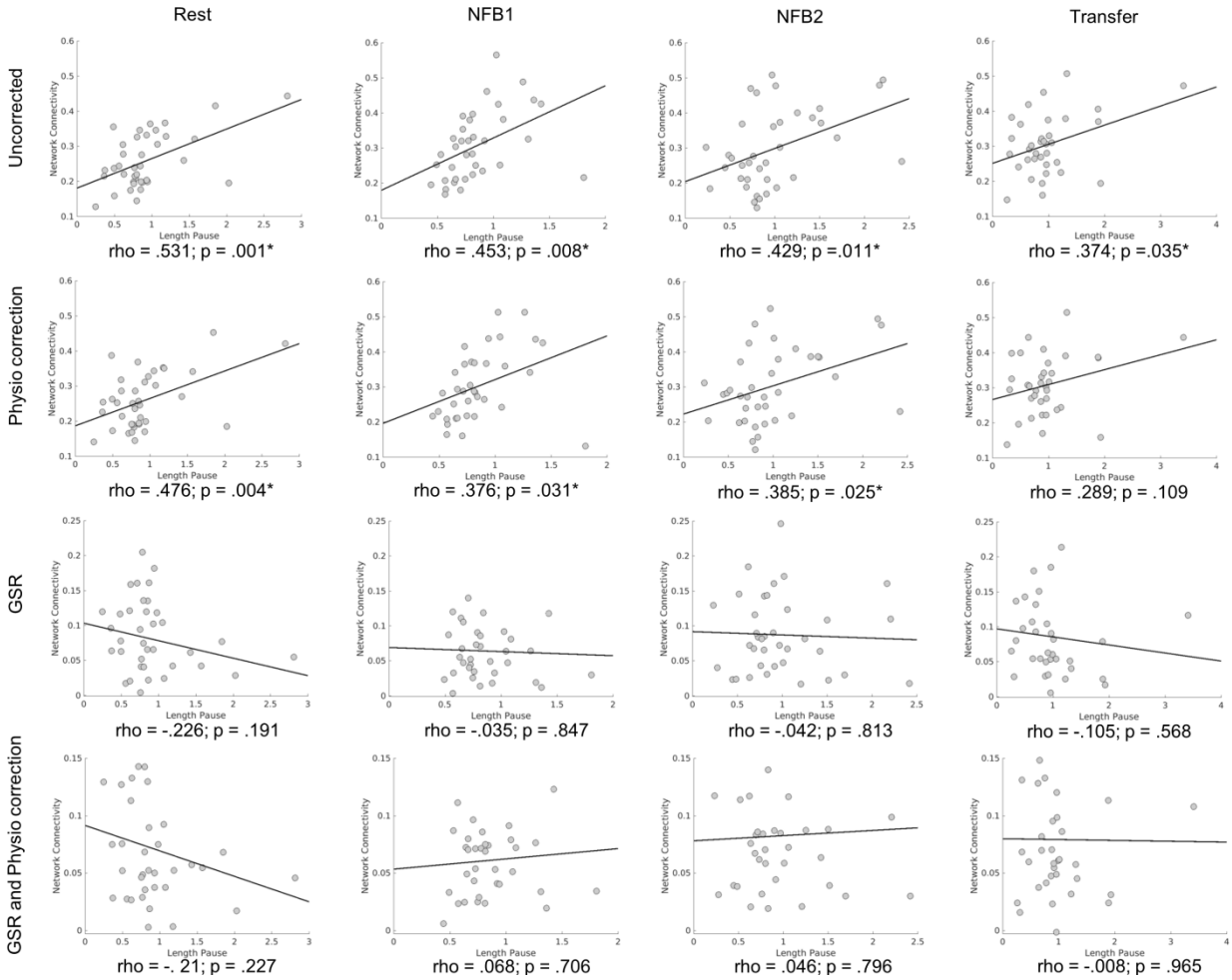


Figure 7. Association of respiratory parameters with network connectivity across all runs. Correlations of the respiratory parameters Length Pause (length of pauses between inspiration cycles). (see Zamoscik et al., 2018 for further details) with target network connectivity for the different physiology corrections. Please see Supplementary Figure 13 for an illustration how pause durations were determined. Physio: model-based physiology correction; GSR: global signal regression. Please note that, as expected, GSR overall shifted connectivity estimates towards 0.

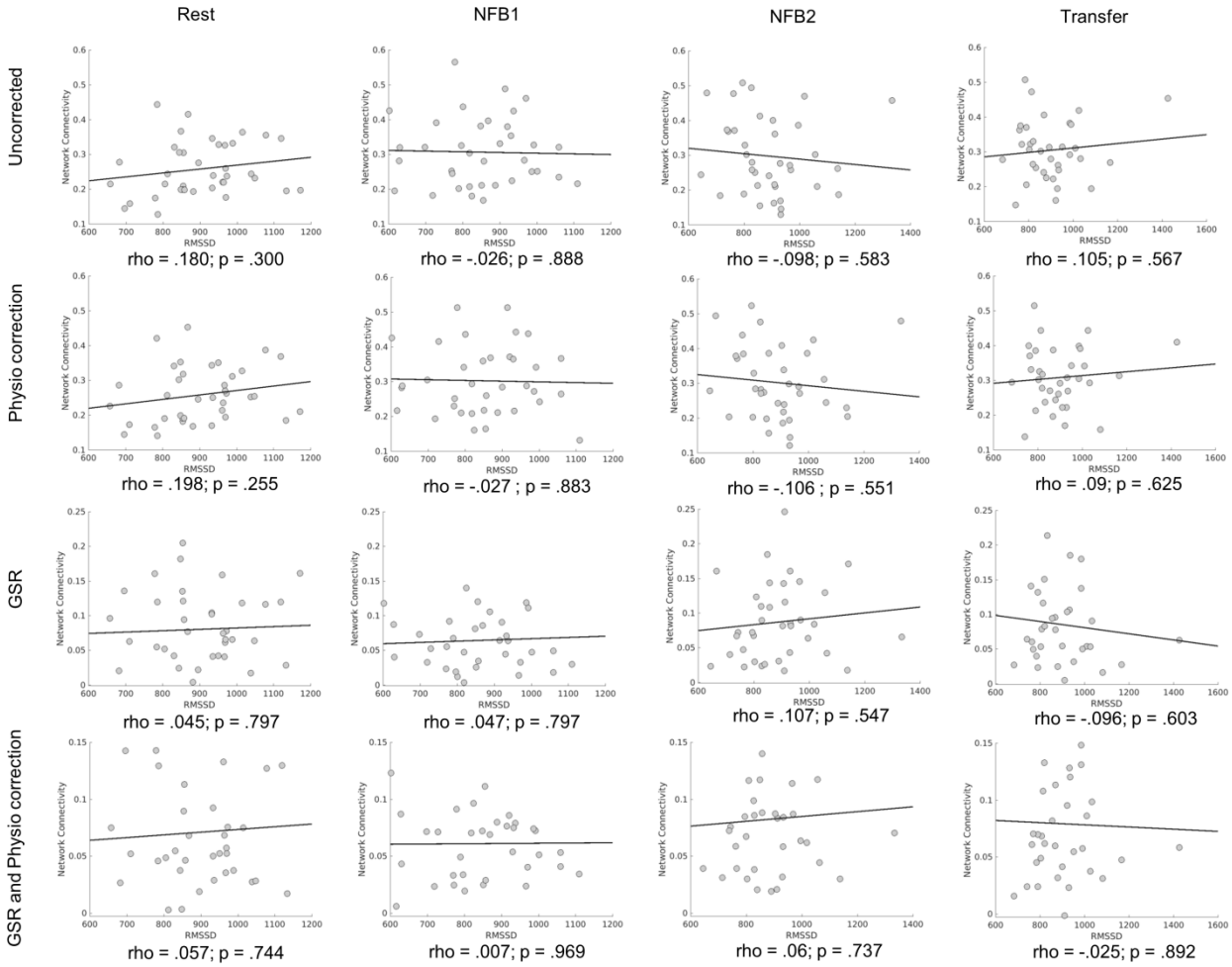


Figure 8. Association of cardiac parameters with network connectivity across all runs. Correlations of the cardiac parameter RMSSD (root mean square of the successive differences [ms]) with target network connectivity for the different physiology corrections. Physio: model-based physiology correction; GSR: global signal regression. Please note that, as expected, GSR overall shifted connectivity estimates towards 0.

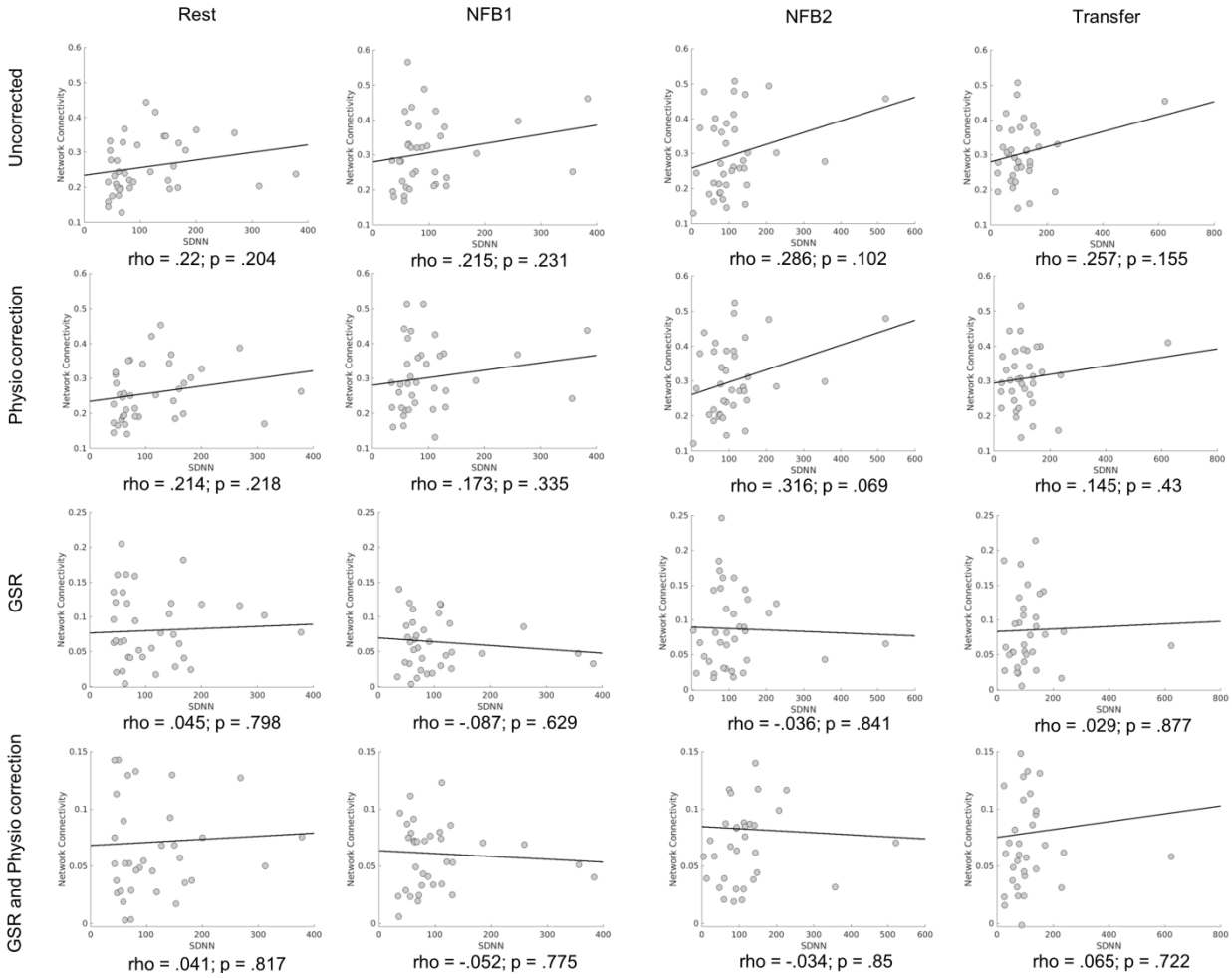


Figure 9. Association of cardiac parameters with network connectivity in NFB across all runs. Correlations of the cardiac parameter SDNN (standard deviation of all beat intervals [ms]) with target network connectivity for the different physiology corrections. Physio: model-based physiology correction; GSR: global signal regression. Please note that, as expected, GSR overall shifted connectivity estimates towards 0.

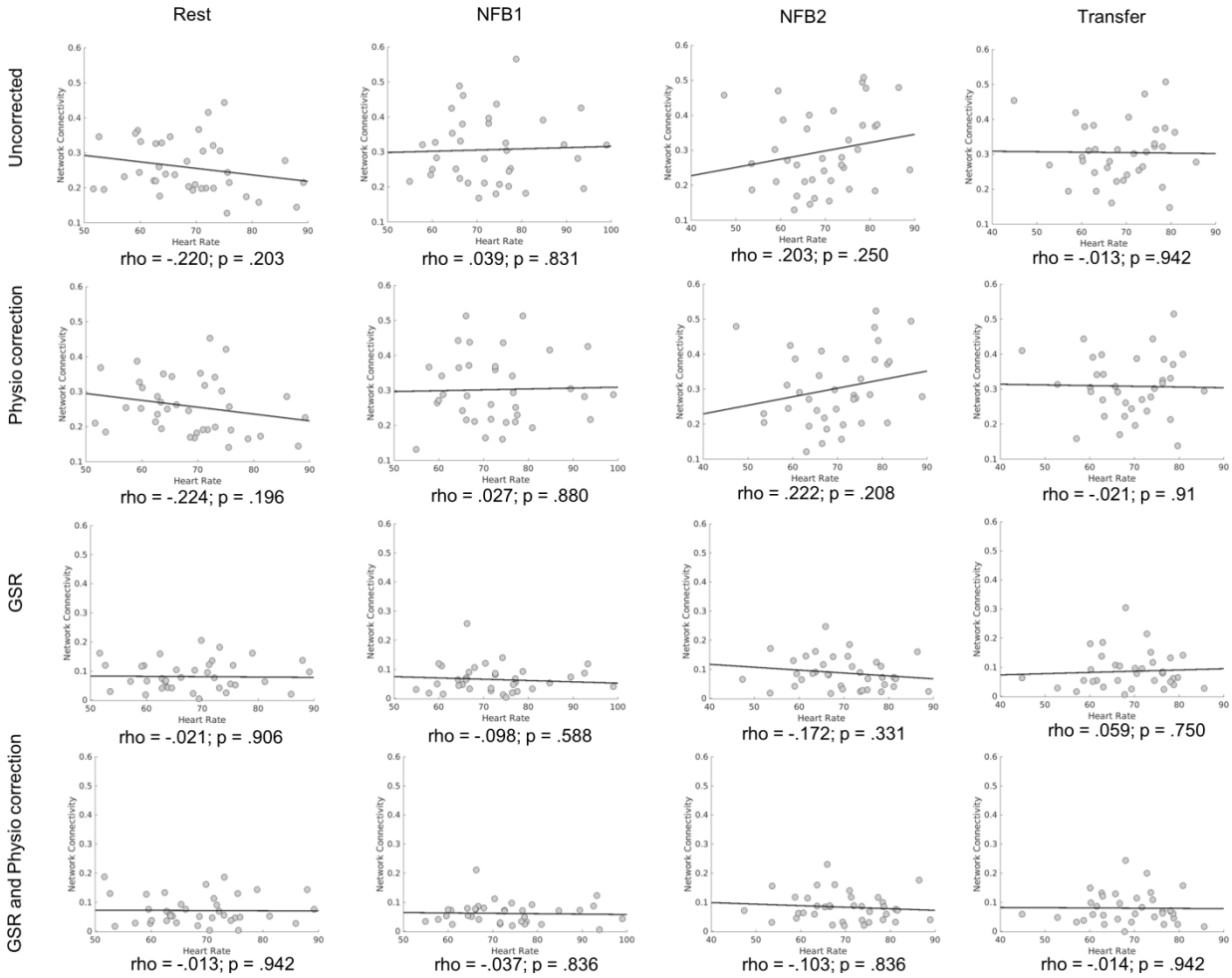


Figure 10. Association of cardiac parameters with network connectivity across all runs. Correlations of the cardiac parameter Heart Rate (see Zamoscik et al., 2018 for further details) with target network connectivity for the different physiology corrections. Physio: model-based physiology correction; GSR: global signal regression. Please note that, as expected, GSR overall shifted connectivity estimates towards 0.

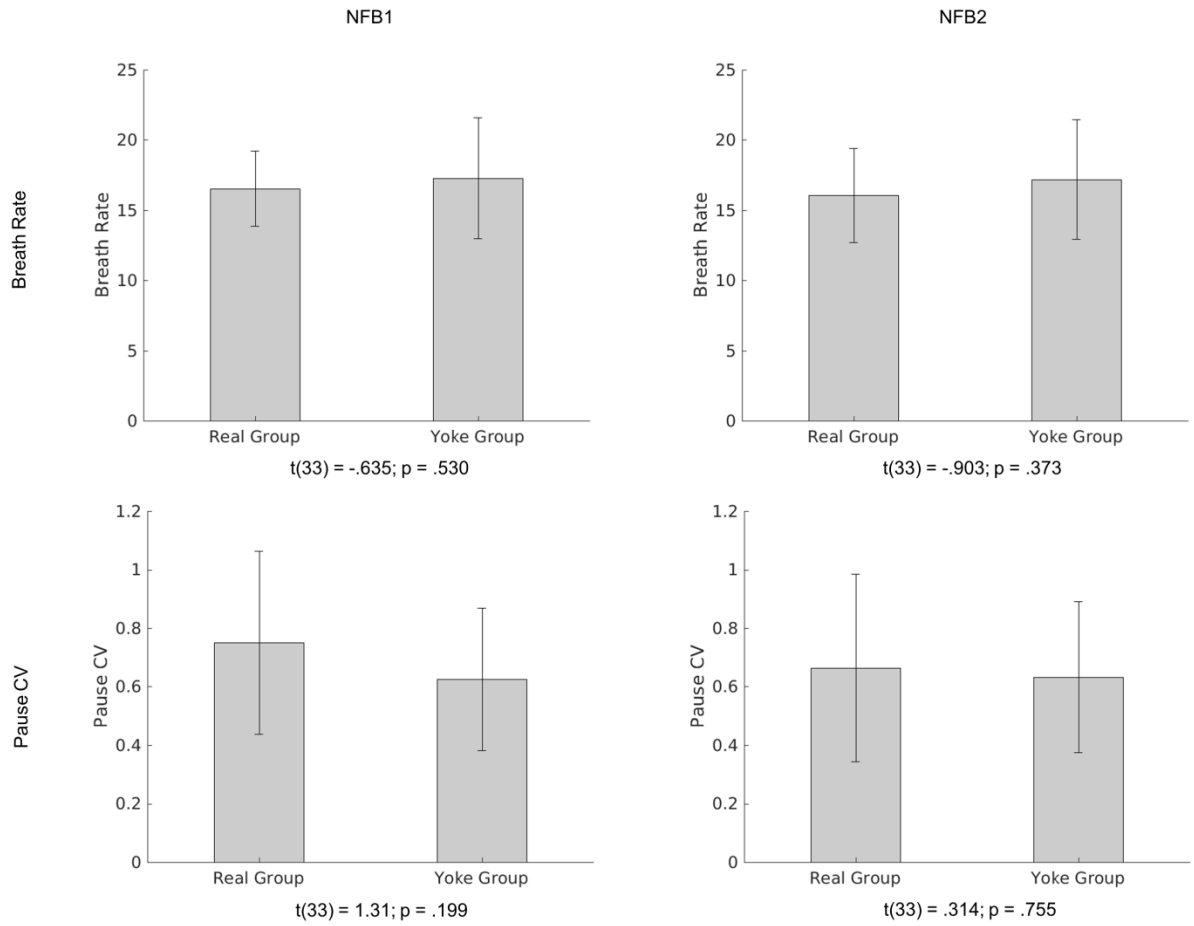


Figure 11. Comparison of the respiratory parameters Breath Rate (peaks/breaths per minute) and Pause CV (coefficient of variance of pause duration) between the groups for NFB1 and NFB2.

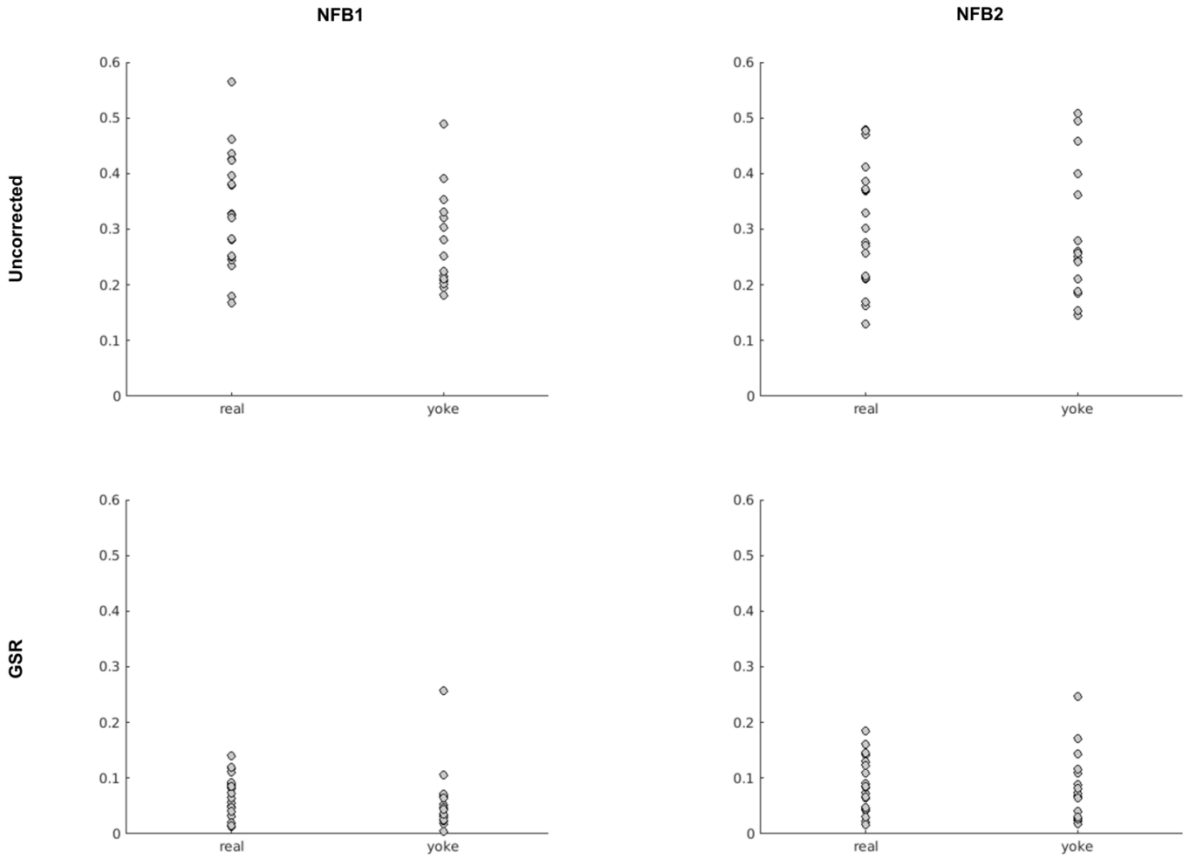


Figure 12. Individual connectivity values of the groups in the two neurofeedback runs. Upper row: Group comparison of network connectivity between the real NFB and yoke NFB group uncorrected for physiology. Lower row: Group comparison of network connectivity between the real NFB and yoke NFB group in data with global signal (GSR) correction. A closer inspection of the connectivity values in NFB1 and NFB2 shows that some subjects from the yoke group increased connectivity, which is still apparent after GSR correction, and might explain the diminished NFB effect in NFB2. We can only speculate, but a possible explanation might be that these subjects tried harder to control the yoked signal they had no control of, which increased connectivity in our target executive frontostriatal network.

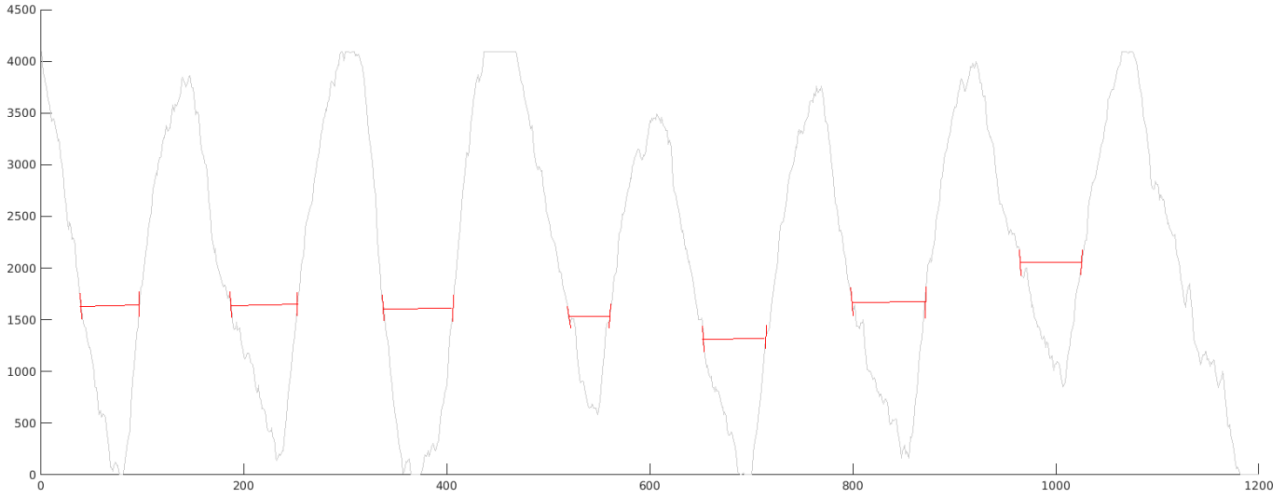


Figure 13. Respiration pause estimation depicted on the respiration time course. Data of a not specifically selected participant (first participant) is shown. Pauses are determined by calculating the slope of the respiration curve at a window size of five samples with a cut-off of 2. Clusters of minimum peaks were used to determine coarse temporal markers for a provisional pause onset. This was then recursively extended into both directions based on the slope parameters to determine pause onset and offset. Red colour indicates how the pause parameter is derived from the signal. Please note that estimation is based on the locally determined points where the slope changes. Therefore, the location of the points can differ between breaths. The Pause CV parameter is the coefficient of variance (the standard deviation divided by the mean) of the estimated pause durations.

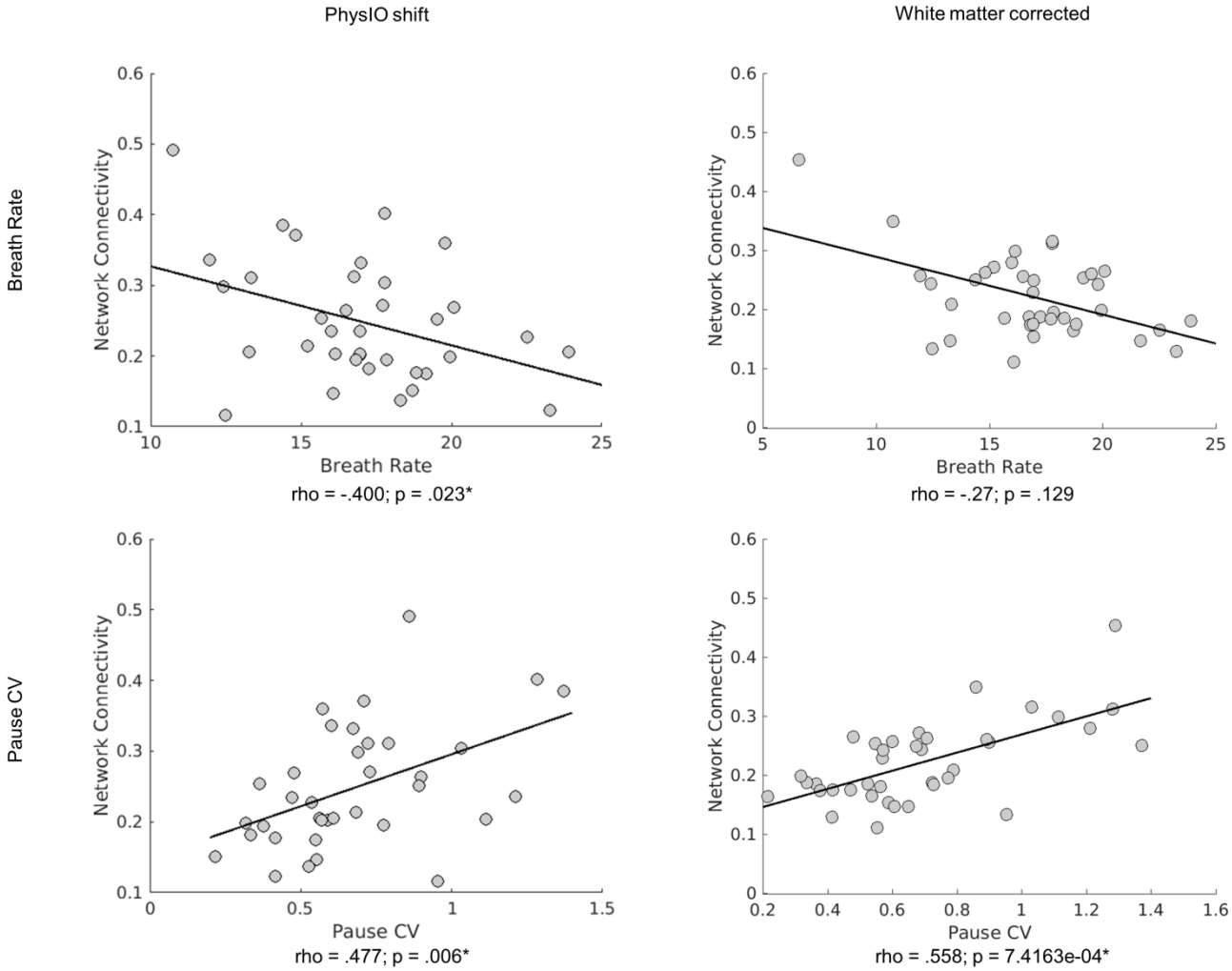


Figure 14. Association of respiratory parameters with network connectivity during NFB1. Correlations of the respiratory parameters Pause CV (coefficient of variance of pause duration) and Breath Rate (peaks/breaths per minute) with target network connectivity for the different physiology corrections. Please see Supplementary Figure 13 for an illustration how pause durations and the Pause CV parameter were estimated. PhysIO shift: model-based physiology correction with temporally shifted versions of the respiratory response function before convolution with RVT (shifts from -24 s to 18 s in 6 s steps and additional shits of -3 -1 1 3 s); White matter corrected: white matter regression.

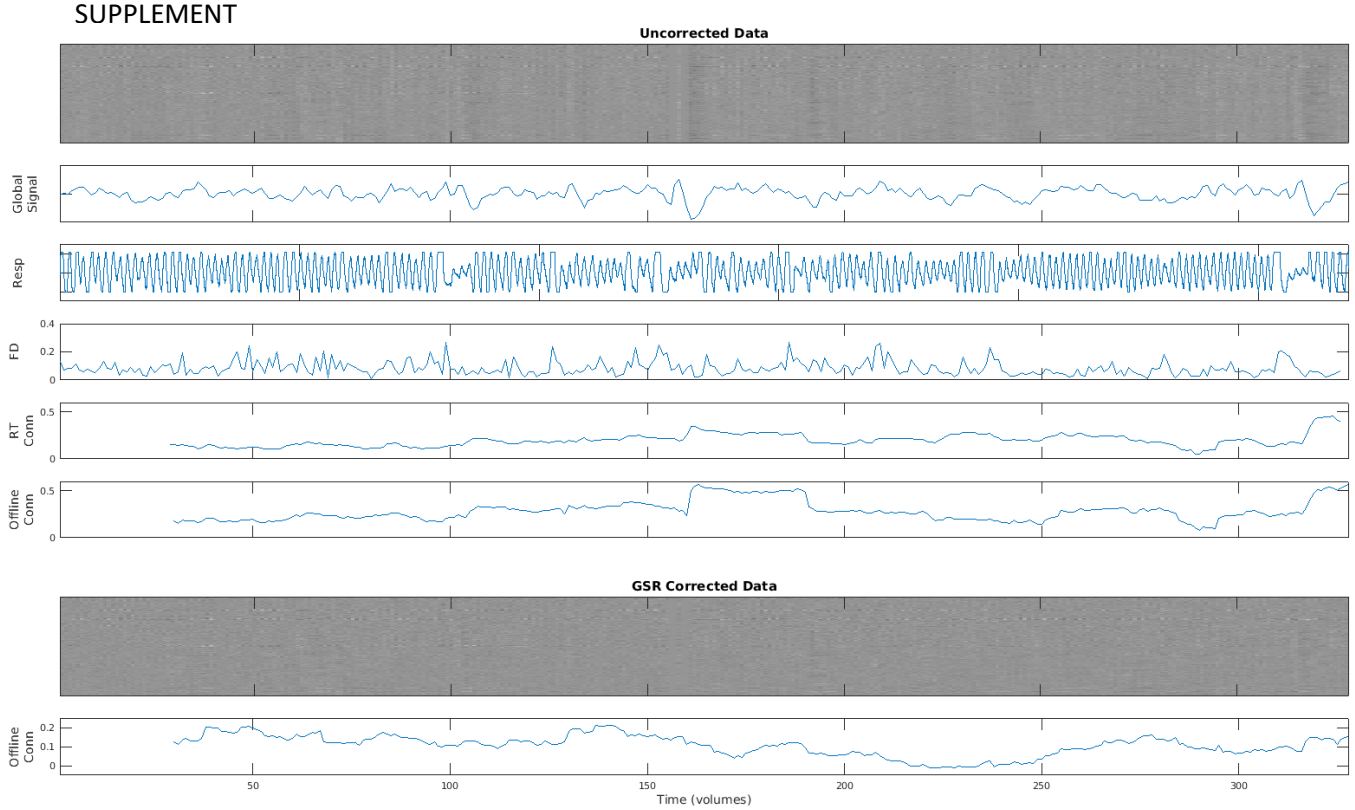


Figure 15. Illustration of time courses covariation of relevant signals. Exemplary data from a not specifically selected participant (first participant). From top to bottom: Carpet plot of gray matter voxel intensities over time not corrected for physiology (uncorrected data), global signal, respiration (Resp), framewise displacement (FD) of the head in mm, real-time dynamic functional connectivity (RT conn) of the target network used for feedback, offline dynamic functional connectivity (Offline Conn) of the target network without physiology correction, Carpet plot of gray matter voxel intensities over time after GSR correction, offline dynamic functional connectivity (Offline Conn) of the target network after GSR correction.

11.2.2 Supplementary figures NeCoSchi II

Supplementary Table 1. Group comparisons of functional connectivity per run corrected for age and gender as covariates. Functional connectivity was not corrected for baseline resting state activity.

Runs	**Group Differences Functional Connectivity**
Day1_NF1	t(32) = -0.0082 p = 0.49675
Day1_NF2	t(30) = -0.2884 p = 0.38755
Day1_transfer	t(31) = -0.2995 p = 0.3833
Day2_NF1	t(36) = -0.5687 p = 0.28655
Day2_NF2	t(31) = -0.4512 p = 0.3275
Day2_NF3	t(33) = -0.2920 p = 0.38605
Day3_NF1	t(35) = 0.4108 p = 0.34185
Day3_NF2	t(33) = 1.2299 p = 0.1137
Day3_transfer	t(33) = 0.8042 p = 0.2135

Supplementary Table 2. Associations of offline functional connectivity with respiratory measures displayed per group.

	Real		**Yoke**	
	Respiratory measures		**Respiratory measures**	
Runs	**Pause CV**	**Breath Rate**	**Pause CV**	**Breath Rate**
Day1_NF1	rho = -0.1885 p = 0.4537	rho = -0.2536 p = 0.3100	rho = 0.2253 p = 0.3847	rho = 0.0578 p = 0.8256
Day1_NF2	rho = -0.4650 p = 0.0939	rho = -0.3912 p = 0.1666	rho = -0.2721 p = 0.2597	rho = 0.1207 p = 0.6226
Day1_transfer	rho = -0.5911 p = 0.0203*	rho = -0.0976 p = 0.7293	rho = -0.1508 p = 0.5503	rho = 0.4324 p = 0.0731
Day2_NF1	rho = -0.0280 p = 0.9093	rho = -0.1887 p = 0.4839	rho = 0.3652 p = 0.1241	rho = -0.3439 p = 0.1494
Day2_NF2	rho = -0.2716 p = 0.3089	rho = -0.0239 p = 0.9301	rho = 0.6572 p = 0.0030*	rho = -0.0879 p = 0.7288
Day2_NF3	rho = -0.2254 p = 0.4013	rho = -0.3271 p = 0.2162	rho = -0.3212 p = 0.1799	rho = 0.2630 p = 0.2767
Day3_NF1	rho = -0.4378 p = 0.0608	rho = 0.3303 p = 0.1673	rho = -0.1536 p = 0.5301	rho = 0.1542 p = 0.5285
Day3_NF2	rho = -0.2707 p = 0.2933	rho = -0.1840 p = 0.4797	rho = -0.0182 p = 0.9412	rho = 0.2095 p = 0.3893
Day3_transfer	rho = -0.3791 p = 0.1475	rho = 0.2949 p = 0.2675	rho = -0.1079 p = 0.6701	rho = 0.3807 p = 0.1191

Supplementary Table 3. Associations of online functional connectivity with respiratory measures displayed per group.

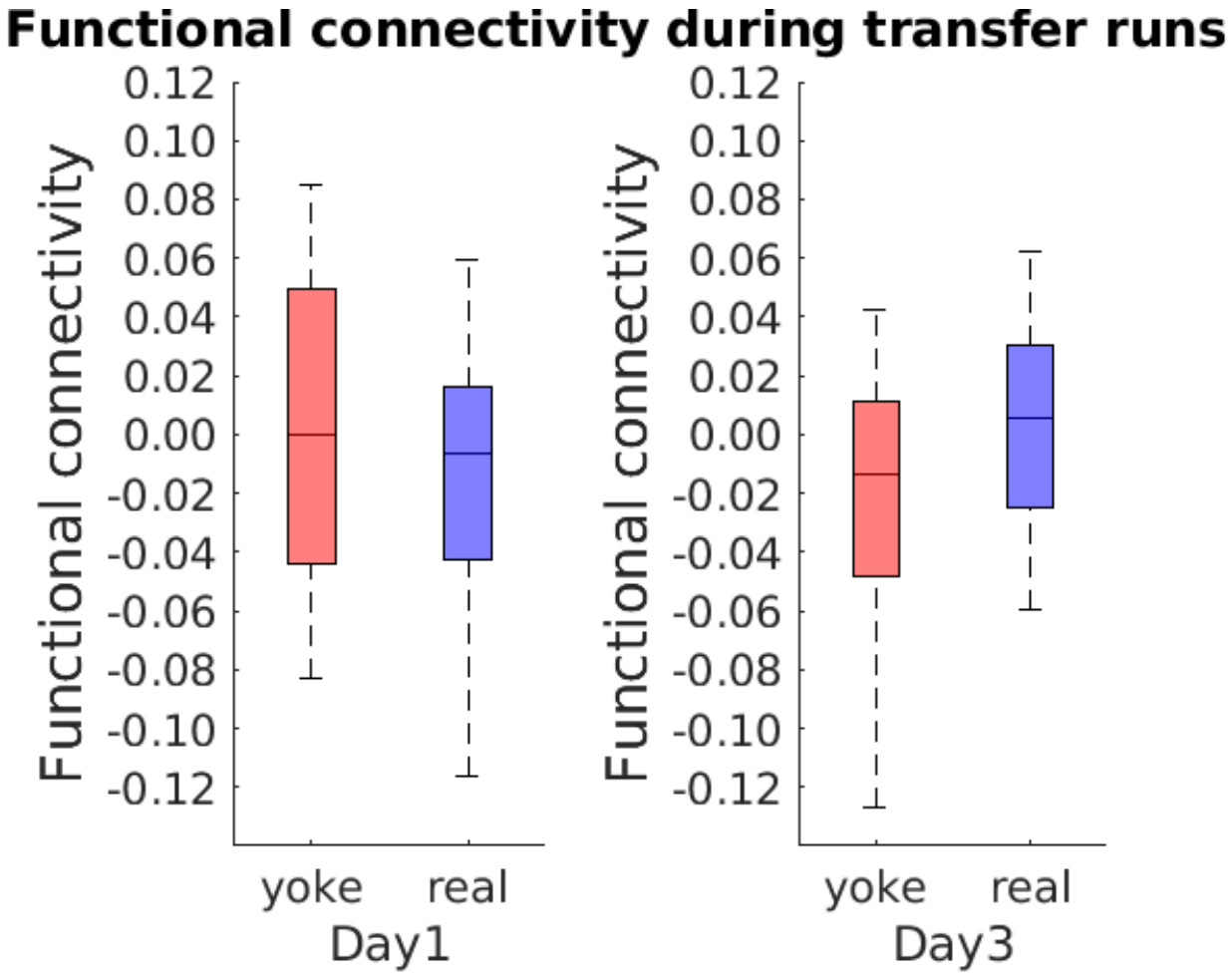
	Real		**Yoke**	
	Respiratory measures		**Respiratory measures**	
Runs	**Pause CV**	**Breath Rate**	**Pause CV**	**Breath Rate**
Day1_NF1	rho = 0.6498 p = 0.0035*	rho = 0.1172 p = 0.6432	rho = 0.4511 p = 0.0692	rho = -0.3838 p = 0.1283
Day1_NF2	rho = -0.1008 p = 0.7317	rho = -0.2795 p = 0.3332	rho = -0.1328 p = 0.5877	rho = -0.2012 p = 0.4089
Day1_transfer	-	-	-	-
Day2_NF1	rho = -0.0359 p = 0.8840	rho = -0.3755 p = 0.1132	rho = 0.4821 p = 0.0366*	rho = -0.5795 p = 0.0093*
Day2_NF2	rho = 0.0558 p = 0.8373	rho = -0.3524 p = 0.1806	rho = 0.3124 p = 0.2069	rho = -0.3479 p = 0.1571
Day2_NF3	rho = 0.3974 p = 0.1275	rho = -0.5630 p = 0.0232*	rho = -0.2574 p = 0.3024	rho = -0.3408 p = 0.1664
Day3_NF1	rho = 0.1403 p = 0.5669	rho = -0.3422 p = 0.1515	rho = 0.4510 p = 0.0526	rho = -0.4309 p = 0.0655
Day3_NF2	rho = 0.7444 p = 6.1052e-04*	rho = -0.7031 p = 0.0016*	rho = 0.5000 p = 0.0293*	rho = -0.6471 p = 0.0027*
Day3_transfer	-	-	-	-

Supplementary Table 4. Associations of offline functional connectivity with standard pre-processing with respiratory measures displayed per group.

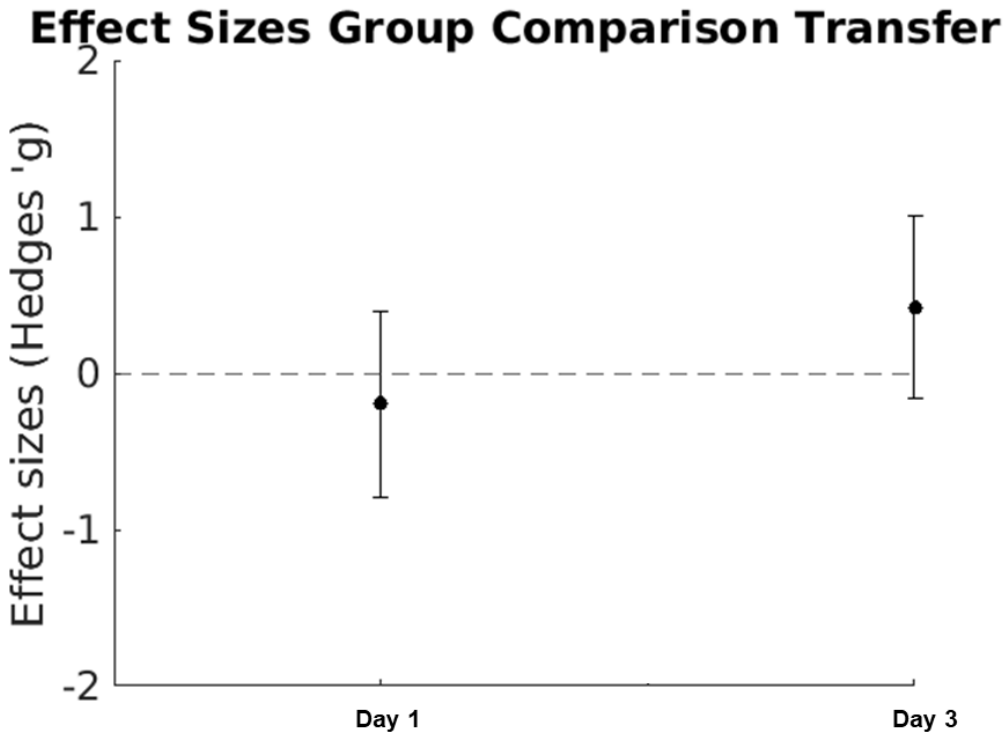
	Real		**Yoke**	
	Respiratory measures		**Respiratory measures**	
Runs	**Pause CV**	**Breath Rate**	**Pause CV**	**Breath Rate**
Day1_NF1	rho = 0.7636 p = 2.2615e-04*	rho = -0.0746 p = 0.7686	rho = 0.6176 p = 0.0082*	rho = -0.5231 p = 0.0312*
Day1_NF2	rho = 0.7277 p = 0.0032*	rho = 0.1915 p = 0.5120	rho = 0.4607 p = 0.0472*	rho = -0.6268 p = 0.0041*
Day1_transfer	rho = 0.4715 p = 0.0761	rho = -0.4244 p = 0.1148	rho = 0.6327 p = 0.0036*	rho = -0.4354 p = 0.0624
Day2_NF1	rho = 0.7473 p = 2.3544e-04*	rho = -0.2061 p = 0.3974	rho = 0.8777 p = 7.9500e-07*	rho = -0.8073 p = 2.9321e-05*
Day2_NF2	rho = 0.5822 p = 0.0142*	rho = -0.1632 p = 0.5315	rho = 0.7990 p = 4.0792e-05*	rho = -0.7416 p = 2.7910e-04*
Day2_NF3	rho = 0.7840 p = 3.2488e-04*	rho = -0.1934 p = 0.4729	rho = 0.0282 p = 0.9088	rho = -0.6168 p = 0.0049*
Day3_NF1	rho = 0.5763 p = 0.0098*	rho = -0.5674 p = 0.0113*	rho = 0.6262 p = 0.0041*	rho = -0.5786 p = 0.0095*
Day3_NF2	rho = 0.8355 p = 2.9782e-05*	rho = -0.4874 p = 0.0472*	rho = 0.4082 p = 0.0827	rho = -0.6110 p = 0.0055*
Day3_transfer	rho = 0.7328 p = 0.0012*	rho = -0.4997 p = 0.0487*	rho = 0.4096 p = 0.0816	rho = -0.6258 p = 0.0042*

Supplementary Table 5. Associations of offline functional connectivity with GSR only with respiratory measures displayed per group.

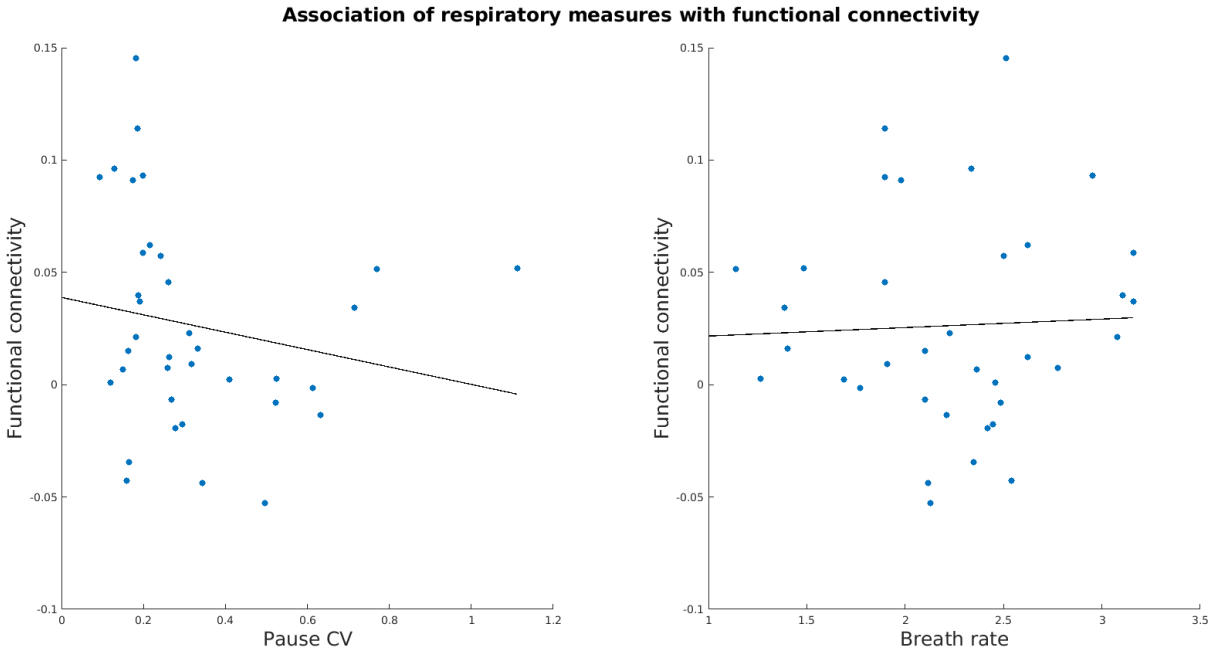
	Real		**Yoke**	
	Respiratory measures		**Respiratory measures**	
Runs	**Pause CV**	**Breath Rate**	**Pause CV**	**Breath Rate**
Day1_NF1	rho = -0.1348 p = 0.5940	rho = -0.3541 p = 0.1494	rho = 0.1376 p = 0.5986	rho = 0.1571 p = 0.5470
Day1_NF2	rho = -0.4013 p = 0.1549	rho = -0.4605 p = 0.0975	rho = -0.3021 p = 0.2087	rho = 0.1660 p = 0.4971
Day1_transfer	rho = -0.5911 p = 0.0203*	rho = -0.1170 p = 0.6779	rho = -0.2034 p = 0.4035	rho = 0.4568 p = 0.0493*
Day2_NF1	rho = -0.0163 p = 0.9473	rho = -0.2119 p = 0.3838	rho = 0.4550 p = 0.0503	rho = -0.3952 p = 0.0940
Day2_NF2	rho = -0.3889 p = 0.1229	rho = 0.0316 p = 0.9042	rho = 0.5811 p = 0.0091*	rho = -0.0468 p = 0.8491
Day2_NF3	rho = -0.1128 p = 0.6775	rho = -0.3295 p = 0.2126	rho = -0.3035 p = 0.2066	rho = 0.2410 p = 0.3202
Day3_NF1	rho = -0.4167 p = 0.0760	rho = 0.3425 p = 0.1512	rho = -0.1684 p = 0.4908	rho = 0.1130 p = 0.6450
Day3_NF2	rho = -0.2151 p = 0.4069	rho = -0.3508 p = 0.1674	rho = -0.0335 p = 0.8917	rho = 0.2514 p = 0.2992
Day3_transfer	rho = -0.3326 p = 0.2082	rho = 0.0356 p = 0.8959	rho = -0.0696 p = 0.7771	rho = 0.3701 p = 0.1188



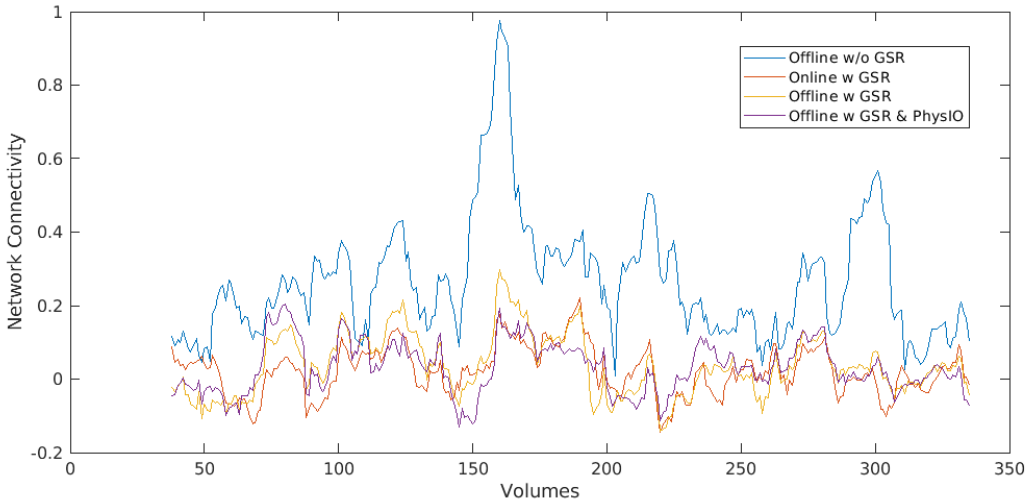
Supplementary Figure 1. Group comparisons of Functional Connectivity (FC) during transfer. Transfer runs were conducted on the first ($t(31) = -0.5548$, $p = 0.2915$) and third training day ($t(32) = 1.2275$, $p = 0.1143$). FC is normalized by initial resting state FC of the respective day and corrected for age and gender.



Supplementary Figure 2. Effect sizes of the group comparisons during transfer runs. Transfer runs were conducted on the first (g = -0.1960) and third training day (g = 0.4199).



Supplementary Figure 3. Relationship of respiratory measures with functional connectivity (FC) in the second NF run of the third training day. The left coordinate system displays FC in relation to Pause CV (rho = -.0182, p = 0.287). On the right, the association with Breath rate is shown (rho = 0.043, p = 0.803).



Supplementary Figure 4. Exemplary target network connectivity time course of a single run in a single participant with the different applied processing strategies. The online time course (red) was estimated during training while the other three were estimated with the offline processing pipeline. The strong influence of GSR on the network connectivity estimate is clearly visible. The differences between online and offline analyses and the additional effect of the PhysIO toolbox appear small in comparison.