- -

Predictors of real-time fMRI neurofeedback performance and improvement – A machine learning mega-analysis

- -

Haugg, Amelie; Renz, Fabian M.; Nicholson, Andrew A.; Lor, Cindy; Götzendorfer, Sebastian J.; Sladky, Ronald; Skouras, Stavros; McDonald, Amalia; Craddock, Cameron; Hellrung, Lydia; Kirschner, Matthias; Herdener, Marcus; Koush, Yury; Papoutsi, Marina [**and 34 more**]

Predictors of real-time fMRI neurofeedback performance and improvement – A machine learning mega-analysis

Amelie Haugg [a,b,*], Fabian M. Renz [b], Andrew A. Nicholson [b], Cindy Lor [b],
Sebastian J. Götzendorfer [b], Ronald Sladky [b], Stavros Skouras [c], Amalia McDonald [d],
Cameron Craddock [e], Lydia Hellrung [f], Matthias Kirschner [a,g], Marcus Herdener [a], Yury Koush [h],
Marina Papoutsi [i,j], Jackob Keynan [k], Talma Hendler [k], Kathrin Cohen Kadosh [l], Catharina Zich [m],
Simon H. Kohl [n], Manfred Hallschmid [o,p,q], Jeff MacInnes [r], R. Alison Adcock [s,t],
Kathryn C. Dickerson [t], Nan-Kuei Chen [u], Kymberly Young [v], Jerzy Bodurka [w,ao],
Michael Marxen [x], Shuxia Yao [y], Benjamin Becker [y], Tibor Auer [l], Renate Schweizer [z],
Gustavo Pamplona [aa], Ruth A. Lanius [ab], Kirsten Emmert [ac], Sven Haller [ad], Dimitri Van De
Ville [ae,af], Dong-Youl Kim [ag], Jong-Hwan Lee [ag], Theo Marins [ah], Fukuda Megumi [ai],
Bettina Sorger [aj], Tabea Kamp [aj], Sook-Lei Liew [ak], Ralf Veit [o,p,q,al], Maartje Spetter [am],
Nikolaus Weiskopf [an,ap], Frank Scharnowski [a,b,#], David Steyrl [a,b,#]

[a] Department of Psychiatry, Psychotherapy and Psychosomatics, Psychiatric University Hospital, University of Zurich, Switzerland
[b] Faculty of Psychology, University of Vienna, Austria
[c] Department of Biological and Medical Psychology, University of Bergen, Norway
[d] Department of Psychology, University of Virginia, United States
[e] Department of Diagnostic Medicine, The University of Texas at Austin Dell Medical School, United States
[f] Zurich Center for Neuroeconomics, Department of Economics, University of Zurich, Switzerland
[g] McConnell Brain Imaging Centre, Montréal Neurological Institute, McGill University, Canada
[h] Department of Radiology and Biomedical Imaging, Yale University, United States
[i] UCL Huntington's Disease Centre, Institute of Neurology, University College London, United Kingdom
[j] IXICO plc, United Kingdom
[k] Functional Brain Center, Wohl Institute for Advanced Imaging, Tel-Aviv Sourasky Medical Center, Tel-Aviv University, Israel
[l] School of Psychology, University of Surrey, United Kingdom
[m] Nuffiled Department of Clinical Neurosciences, University of Oxford, United Kingdom
[n] JARA-Institute Molecular Neuroscience and Neuroimaging (INM-11), Jülich Research Centre, Germany
[o] Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Germany
[p] Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Center Munich at the University of Tübingen, Germany
[q] German Center for Diabetes Research (DZD), Germany
[r] Institute for Learning and Brain Sciences, University of Washington, United States
[s] Duke Institute for Brain Sciences, Duke University, United States
[t] Department of Psychiatry and Behavioral Sciences, Duke University, United States
[u] Department of Biomedical Engineering, University of Arizona, United States
[v] Department of Psychiatry, School of Medicine, University of Pittsburgh, United States
[w] Laureate Institute for Brain Research, United States
[x] Department of Psychiatry, Technische Universität Dresden, Germany
[y] Clinical Hospital of the Chengdu Brain Science Institute, MOE Key Laboratory for Neuroinformation, University of Electronic Science and Technology of China, China
[z] Functional Imaging Laboratory, German Primate Center, Germany
[aa] Department of Ophthalmology, University of Lausanne and Fondation Asile des Aveugles, Switzerland
[ab] Department of Psychiatry, University of Western Ontario, Canada
[ac] Department of Neurology, University Medical Center Schleswig-Holstein, Kiel University, Germany
[ad] Department of Surgical Sciences, Radiology, Uppsala University, Sweden
[ae] Center for Neuroprosthetics, Ecole polytechnique féderale de Lausanne, Switzerland
[af] Faculty of Medicine, University of Geneva, Switzerland
[ag] Department of Brain and Cognitive Engineering, Korea University, Korea
[ah] D'Or Institute for Research and Education, Brazil

* Corresponding author.
  E-mail address: amelie.haugg@uzh.ch (A. Haugg).
# Contributed equally.

ai *Center for Brain Science, RIKEN, Japan*
aj *Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, The Netherlands*
ak *University of Southern California, United States*
al *High-Field Magnetic Resonance Center, Max Planck Institute for Biological Cybernetics, Germany*
am *School of Psychology, University of Birmingham, United Kingdom*
an *Max Planck Institute for Human Cognitive and Brain Sciences, Germany*
ao *Stephenson School of Biomedical Engineering, University of Oklahoma, United States*
ap *Felix Bloch Institute for Solid State Physics, Faculty of Physics and Earth Sciences, Leipzig University, Germany*

ARTICLE INFO

ABSTRACT

Real-time fMRI neurofeedback is an increasingly popular neuroimaging technique that allows an individual to gain control over his/her own brain signals, which can lead to improvements in behavior in healthy participants as well as to improvements of clinical symptoms in patient populations. However, a considerably large ratio of participants undergoing neurofeedback training do not learn to control their own brain signals and, consequently, do not benefit from neurofeedback interventions, which limits clinical efficacy of neurofeedback interventions. As neurofeedback success varies between studies and participants, it is important to identify factors that might influence neurofeedback success. Here, for the first time, we employed a big data machine learning approach to investigate the influence of 20 different design-specific (e.g. activity vs. connectivity feedback), region of interest-specific (e.g. cortical vs. subcortical) and subject-specific factors (e.g. age) on neurofeedback performance and improvement in 608 participants from 28 independent experiments.

With a classification accuracy of 60% (considerably different from chance level), we identified two factors that significantly influenced neurofeedback performance: Both the inclusion of a pre-training no-feedback run before neurofeedback training and neurofeedback training of patients as compared to healthy participants were associated with better neurofeedback performance. The positive effect of pre-training no-feedback runs on neurofeedback performance might be due to the familiarization of participants with the neurofeedback setup and the mental imagery task before neurofeedback training runs. Better performance of patients as compared to healthy participants might be driven by higher motivation of patients, higher ranges for the regulation of dysfunctional brain signals, or a more extensive piloting of clinical experimental paradigms. Due to the large heterogeneity of our dataset, these findings likely generalize across neurofeedback studies, thus providing guidance for designing more efficient neurofeedback studies specifically for improving clinical neurofeedback-based interventions. To facilitate the development of data-driven recommendations for specific design details and subpopulations the field would benefit from stronger engagement in open science research practices and data sharing.

1. Introduction

Real-time functional magnetic resonance (fMRI) neurofeedback is a non-invasive technique that enables healthy individuals and patients to voluntarily regulate neural signals. In the last decades, this method has gained growing popularity in the neuroimaging community and, to date, a wide range of real-time fMRI neurofeedback studies have collectively demonstrated the feasibility of volitional regulation through real-time fMRI neurofeedback (see Thibault et al. (2018)). Further, many of these studies have also shown behavioral changes in healthy individuals, as well as clinical improvements in patient populations after neurofeedback training. In healthy participants, real-time fMRI neurofeedback training has been specifically linked to improvements in attention (e.g. DeBettencourt et al., 2015; Pamplona et al., 2020), emotion regulation (Koush et al., 2015; Paret and Hendler, 2020; Zich et al., 2020), memory (e.g. Scharnowski et al., 2015; Sherwood et al., 2016; Zhang et al., 2013), motivation (e.g. Zhi et al., 2018), motor performance (e.g. Bray et al., 2007; Scharnowski et al., 2015; Sitaram et al., 2012; Zhao et al., 2013), speech performance (Rota et al., 2009), and visual perception (e.g. Scharnowski et al., 2012; Shibata et al., 2011). In clinical populations, real-time fMRI neurofeedback training has been shown to both improve clinical measures and normalize pathological neural characteristics in patients suffering from a wide range of disorders, such as alcohol and nicotine addiction (Canterberry et al., 2013; Hanlon et al., 2013; Hartwell et al., 2016; Karch et al., 2015; Kim et al., 2015; X. Li et al., 2013), anxiety (Morgenroth et al., 2020), borderline personality disorder (Paret et al., 2016), depression (Linden et al., 2012; Mehler et al., 2018; Quevedo et al., 2020; Young et al., 2017, 2014), obsessive compulsive disorder (Buyukturkoglu et al., 2015), phobia (Zilverstand et al., 2015), post-traumatic stress disorder (Gerin et al., 2016; Nicholson et al., 2017), schizophrenia

(Bauer et al., 2020), Tourette syndrome (Sukhodolsky et al., 2020), chronic pain (deCharms et al., 2005; Guan et al., 2014), Huntington's disease (Papoutsi et al., 2018), obesity (Frank et al., 2012; Kohl et al., 2019), Parkinson's disease (Buyukturkoglu et al., 2013; Subramanian et al., 2011), stroke rehabilitation (Mehler et al., 2020), tinnitus (Emmert et al., 2017; Haller et al., 2010), and visuo-spatial neglect (Fabien Robineau et al., 2019).

Critically however, not all participants undergoing real-time fMRI neurofeedback training optimally benefit from the aforementioned improvements on behavioral and clinical measures, due to variations in their success on acquiring neural control. Previous real-time fMRI neurofeedback studies have reported relatively high rates of non-responders, i.e., participants who fail to regulate their brain signals in the desired direction (Bray et al., 2007; Chiew et al., 2012; deCharms et al., 2005; Johnson et al., 2012; Ramot et al., 2016; F. Robineau et al., 2014; Scharnowski et al., 2012; Yoo et al., 2008). Averaging across these studies, the non-responder rate of real-time fMRI neurofeedback studies is estimated to lie around 38% (Haugg et al., 2020). Here, it should be noted that, to date, no standard thresholds for identifying non-responders are available and definitions of non-responders often vary between studies. Generally, even real-time fMRI neurofeedback participants who were eventually able to gain control over their own brain signals still showed large variability in their neurofeedback regulation performance (Haugg et al., 2020). Similar estimations and observations have also been reported in the electroencephalogram (EEG) neurofeedback literature, where the so-called "neurofeedback inefficacy problem" refers to the variability in neurofeedback success and comprises a well-known issue (Alkoby et al., 2017). Therefore, the fields of both EEG- and fMRI-based neurofeedback would greatly benefit from methodologically advanced investigations that can reveal the factors responsible for the unexplained variability of neurofeedback success.

Interestingly, previous studies have demonstrated that the proportion of responders varies between different neurofeedback studies. Of importance, this suggests that some neurofeedback study-specific parameters might be more beneficial for neurofeedback success than others. Previously, few empirical studies have investigated the influence of neurofeedback design parameters on neurofeedback success. Specifically, two independent studies found that using an intermittent feedback display was superior over using a continuous feedback display (Hellrung et al., 2018; Johnson et al., 2012), while conversely, a third study reported this effect only for a single session of neurofeedback, but not for multiple neurofeedback sessions (Emmert et al., 2017). In another study, Papoutsi and colleagues investigated the influence of activity- versus connectivity-based neurofeedback on neurofeedback success, but did not find a significant difference between activity- and connectivity-based neurofeedback (Papoutsi et al., 2020). Interestingly, Kim et al. reported increased neurofeedback efficacy when combining connectivity-based information with activity-based neurofeedback (Kim et al., 2015). Focusing on subject-specific psychological factors in a systematic review, Cohen Kadosh and colleagues observed that attention and motivation might be important factors for determining neurofeedback success (Cohen and Staunton, 2019). However, an empirical validation of these suggestions is still needed. Other empirical studies observed a relationship between subject-specific questionnaires and neurofeedback success, yet these questionnaires were highly specific for the trained target region and participant population, and therefore do not generalize to other neurofeedback studies (Emmert et al., 2017; Koush et al., 2015).

Taken together, these empirical studies contribute invaluable information regarding the optimal design of neurofeedback studies. However, many critical factors that might influence neurofeedback success have not been investigated yet. For instance, it is not known whether a large number of neurofeedback training runs is beneficial for neurofeedback success, an essential question in the field of fMRI-based neurofeedback due to the high cost of scanning hours. This also includes the question of whether neurofeedback training should be performed across several training days to facilitate neurofeedback learning through sleep consolidation. Other important factors are the inclusion of reinforcers such as monetary rewards (Sepulveda et al., 2016) and social rewards (Mathiak et al., 2010), or the highly debated question of whether participants should receive precise or more open instructions regarding regulation strategies (Sitaram et al., 2016). Ultimately, the number of possible factors that might influence neurofeedback performance and the number of conceivable interactions between these factors are immense and it would not be feasible to untangle them and optimize design empirically. Further, statistical power and generalizability across different study designs are limited in original empirical studies.

On balance, 'big data' approaches encompassing a wide range of neurofeedback participants and studies constitute an unprecedented opportunity that can be used to investigate a large number of factors that might influence neurofeedback success. In addition, big data methods allow correcting for possible interactions and usually result in relatively generalizable findings. To date, however, big data investigations encompassing a large number of participants are still rare in the field of real-time fMRI neurofeedback. Several reviews have descriptively summarized the field (Heunis et al., 2020; Thibault et al., 2018) or studies focusing on a specific application (Dudek and Dodell-Feder, 2021; Trambaiolli et al., 2021; Tursic et al., 2020; Wang et al., 2018). Other big data investigations investigated the influence of pre-training brain activation levels on neurofeedback success (Haugg et al., 2020) and the relationship between brain structures and neurofeedback success (Zhao et al., 2021).

Here, for the first time, we employ exploratory machine learning methods to compute the influence of a wide range of different subject- and study-specific factors on real-time fMRI neurofeedback success. In particular, we investigated the influence of 20 different factors on neurofeedback success in 608 participants undergoing neurofeedback training across 28 independent studies. The investigated factors included three subject-specific factors, six region of interest (ROI)-based factors, and eleven paradigm-specific factors.

Identifying factors that influence neurofeedback success can help to design more effective neurofeedback studies in the future. This can improve neurofeedback studies investigating healthy participants and, more importantly, it can, further, improve clinical neurofeedback interventions. Future designs with increased effectiveness will allow participants to train their target brain regions more efficiently, thus reducing cognitive exhaustion and overall costs. Critically, increasing the effectiveness of neurofeedback designs is an important step towards the alleviation of clinical symptoms, by enabling the development of advanced, personalized treatments for psychiatric and neurological disorders. Taken together, our research aim is to utilize big data approaches in an effort to guide future empirical investigations that utilize real-time fMRI neurofeedback.

2. Material and methods

2.1. Included studies

Data for this mega-analysis could not be gathered from publications alone as single subject data were needed. Therefore, we contacted corresponding authors from real-time fMRI neurofeedback studies via i) the mailing list of the real-time functional neuroimaging community (https://utlists.utexas.edu/sympa/subscribe/rtfin), ii) neuroimaging conferences, and iii) direct email communication, in order to ask for data contributions. To ensure generalizability and to generate a dataset sufficiently large for machine learning analyses, we included fMRI-based neurofeedback studies of any training type (activity- as well as connectivity-based neurofeedback), any target brain region(s), and any participant populations. We received data contributions from authors of 28 independent studies (Auer et al., 2015; Emmert et al., 2017; Hellrung et al., 2018; Kim et al., 2015; Kirschner et al., 2018; Kohl et al., 2019; MacInnes et al., 2016; Marins et al., 2015; Marxen et al., 2016; McDonald et al., 2017; Megumi et al., 2015; Nicholson et al., 2017; Pamplona et al., 2020; Papoutsi et al., 2020, 2018; Scharnowski et al., 2012, 2015; Sorger et al., 2018; Spetter et al., 2017; Yao et al., 2016; Young et al., 2017; Zich et al., 2020), covering a wide range of trained brain regions, different study designs, and participant populations. Table 1 provides an overview of all studies that contributed data to this mega-analysis. In total, we collected data from 608 participants, including 229 patients and 379 healthy participants. All studies included were approved by the respective local ethics committees.

2.2. Data availability

Due to ethical restrictions, the majority of the included data cannot be made publicly available, but will be shared upon request. Data from the study by McDonald et al. is publicly available as part of the Nathan Kline Institute-Rockland Sample (http://fcon_1000.projects.nitrc.org/indi/enhanced/). Data from the study by Kohl et al. has been made publicly available on the Open Science Framework (https://osf.io/3agvn/).

2.3. Neurofeedback success measures

To assess neurofeedback success, we asked authors to provide the average feedback value for each neurofeedback training run. Feedback values were defined as the measures that determined the feedback given to the participants during neurofeedback training. Consequently, the type of feedback values varied between different neurofeedback studies (e.g. percent signal change values, beta values, Bayes factors, correlations values etc.). Based on these feedback values, we then defined two general measures for neurofeedback success that would allow for compar-

Table 1

Overview of studies included in the mega-analysis. We received data from 28 independent neurofeedback studies, including 608 participants (229 patients and 379 healthy participants). 24 studies used activity-based neurofeedback, 6 studies used connectivity-based neurofeedback. Abbreviations: ACC – Anterior Cingulate Cortex, dlPFC – dorsolateral Prefrontal Cortex, mPFC – medial Prefrontal Cortex, M1 – Primary Motor Cortex, OFC – Orbitofrontal Cortex, PCC – Posterior Cingulate Cortex, PMC – Pre-Motor Cortex, PHC – Parahippocampal Cortex, SMA – Supplementary Motor Cortex, SMC – Somatomotor Cortex, SPL – Superior Parietal Lobe, VTA – Ventral Tegmental Area.

Author (year)	ROI(s)	participants	neurofeedback type
Auer et al. (2015)	SMC	healthy ($N = 16$)	activity
Emmert et al. (2017)	auditory cortex	tinnitus ($N = 11$)	activity
Hellrung et al. (2018)	amygdala	healthy ($N = 34$)	activity
Hellrung et al. (in prep)	amygdala	healthy ($N = 16$)	activity
Hellrung et al. (in prep)	insula	healthy ($N = 11$)	activity
Keynan et al. (in prep)	amygdala	healthy ($N = 33$)	activity
Kim et al. (2015)	ACC, mPFC, OFC, PCC, precuneus	tobacco use disorder ($N = 14$)	connectivity, activity
Kirschner et al. (2018)	VTA	healthy ($N = 27$), cocaine use disorder ($N = 24$)	activity
Kirschner et al. (in prep)	VTA	schizophrenia ($N = 14$)	activity
Kohl et al. (2019)	dlPFC	overweight ($N = 16$)	activity
Kohl (pilot data)	dlPFC	overweight ($N = 9$)	activity
Liew et al. (in prep)	left PMC, left SMA	healthy ($N = 10$)	connectivity
MacInnes et al. (2016)	VTA	healthy ($N = 19$)	activity
Marins et al. (2015)	left PMC	healthy ($N = 14$)	activity
Marxen et al. (2016)	amygdala	healthy ($N = 32$)	activity
McDonald et al. (2017)	default mode network	healthy ($N = 68$), psychiatric disorders ($N = 72$)	activity
Megumi et al. (2015)	left lateral parietal, left M1	healthy ($N = 12$)	connectivity
Nicholson et al. (2017)	amygdala	PTSD ($N = 14$)	activity
Pamplona et al. (2020)	default mode network, sustained attention network	healthy ($N = 15$)	activity
Papoutsi et al. (2018)	SMA	Huntington's disease ($N = 10$)	activity
Papoutsi et al. (2020)	SMA, left striatum	Huntington's disease ($N = 16$)	connectivity, activity
Scharnowski et al. (2015)	SMA, PHC	healthy ($N = 7$)	activity
Scharnowski et al. (2012)	visual cortex	healthy ($N = 10$)	activity
(Sorger et al., 2018)	individually different	healthy ($N = 10$)	activity
Spetter et al. (2017)	dlPFC, vmPFC	obesity ($N = 8$)	connectivity
Yao et al. (2016)	anterior insula	healthy ($N = 18$)	activity
Young et al. (2017)	amygdala	depression ($N = 18$)	activity
Zich et al. (2020)	amygdala, dlPFC	adolescents ($N = 27$)	connectivity

isons between participants of different studies and, more importantly, for pooling all participants together:

- *Neurofeedback performance*: General neurofeedback performance for each participant was calculated based on the ratio of successful neurofeedback training runs as compared to unsuccessful neurofeedback training runs. Successful neurofeedback training runs were defined as runs showing feedback values with positive signs for up-regulation and negative signs for down-regulation. For the classification analyses, participants who showed more than 50% of successful neurofeedback training runs were labelled as successful, the others as unsuccessful.
- *Neurofeedback improvement*: Neurofeedback improvement of each participant was calculated based on the slope of the neurofeedback learning curve, i.e. the slope of the regression line over the feedback values of all neurofeedback training runs. For classification analyses, successful participants were then defined as participants with a slope greater than 0, non-successful participants showed a slope smaller or equal 0.

2.4. Investigated factors influencing neurofeedback performance and neurofeedback improvement

We investigated the influence of 20 different factors on neurofeedback success. All factors needed to be applicable to every study (i.e. we did not include measures which were only assessed in part of the included studies) and needed to demonstrate different values in at least 10% of the data. As a consequence, ROIs were clustered together in categories to ensure enough data diversity for each ROI-related factor. The continuous and categorical factors included:

- Three subject-specific factors: (1) age of the participant in years, (2) sex of the participant, (3) health status of the participant (healthy participant or patient);

- Six region of interest (ROI)-based factors: (1) ROI(s) is/are cortical or subcortical, (2) ROI(s) is/are a sensory brain region, (3) ROI(s) is/are part of the default mode network (DMN), (4) ROI(s) is/are part of the salience network, (5) ROI(s) is/are part of the motor network, (6) ROI(s) consist(s) of one brain region or more brain regions;
- Eleven experimental design-specific factors: (1) use of connectivity- vs activity-based measure for feedback computation, (2) use of continuous vs intermittent feedback presentation, (3) use vs no use of functional localizer for defining the trained ROI(s), (4) up- vs down-regulation, (5) use of precise strategy suggestions vs no or open strategy suggestions, (6) use of external (monetary) reward vs no external reward given, (7) use of pre-training no-feedback run (functional runs prior to NFB training, where participants are already asked to modulate their brain signals, however, no feedback over regulation performance is provided) vs no pre-training no-feedback run, (8) length of a single neurofeedback training run in seconds, (9) length of a single neurofeedback regulation block in seconds), (10) number of performed neurofeedback training runs, (11) neurofeedback training on one day vs across several days

2.5. Multivariable predictions of neurofeedback performance and neurofeedback improvement

Individual machine learning analyses were performed in Python (v3.8.3) to identify factors that predict participant-specific neurofeedback performance as well as neurofeedback improvement, using multivariable classification models. For the machine learning models, an *Extra Trees* (ExtraTreesClassifier, scikit-learn library v0.23.1; Pedregosa et al., 2011) approach was used, which is a computationally efficient non-linear classification method. *Extra Trees* implements an ensemble of *Extremely randomized trees* (Geurts et al., 2006). Ensemble methods improve the performance of base predictors, e.g. decision trees, by accumulating the predictions of the base predictors via, e.g.,

Table 2

Extra Trees prediction accuracy for the neurofeedback performance and the neurofeedback improvement target.

	Average of weighted accuracy ± standard deviation	Average of weighted accuracy at chance level	p value of smaller or equal chance level
1. Predicting neurofeedback performance	60.3% ± 12.3	51.0%	$p < 0.001$
2. Predicting neurofeedback improvement	48.1% ± 9.0	48.4%	$p = 0.614$

majority voting. To obtain diverse predictions from the same base predictors processes that introduce randomness are applied when building the base predictors.

The model performance – the prediction accuracy – was estimated using a nested cross-validation (CV) procedure (Cawley and Talbot, 2010). In the main CV loop, a shuffle-split data partitioning with 10% of the studies in the testing-set and 90% of the studies in the training-set was repeated 100 times, resulting in 100 Extra Trees models (300 trees per model). Splitting data along studies allows for assessing robust model performance that generalizes between studies. Feature scaling (z-scoring) and hyper-parameter tuning was carried out within the main CV loop, using the training-set data of the current CV loop only. Hyper-parameter tuning was implemented in an inner (nested) CV procedure, so a separate CV was carried out for each repetition of the outer CV loop. The inner CV loops used, again, a shuffle-split partitioning scheme with 10% of the studies in the inner testing set and 50 repetitions. To control model complexity, we restricted the maximum number of possible interactions of a decision tree in the *Extra Trees* ensembles by controlling the number of maximum leaf nodes per tree. The candidate maximum number of leave nodes was randomly drawn between 2 and 32 (50 random draws, RandomizedSearchCV, scikit-learn, v0.23.1). The maximum number of leave nodes that led to the lowest squared error was subsequently used in the outer CV loop.

After hyper-parameter tuning, an Extra Trees model was trained in the main (outer) CV loop using the obtained hyper-parameter and 300 trees with no maximum features. Further, minimum samples split was set to 2, minimum samples leaf to 1, and minimum weight fraction leaf to 0.0. No maximum depth and no maximum samples were chosen, minimum impurity decrease was 0.0, ccp alpha was 0.0, and the class weight was computed from training data.

The obtained model was then tested on the respective hold-out set of the main CV loop. The hold-out set (10% of the studies) was explicitly not used in the inner CV loop. In each repetition of the main CV loop, model prediction accuracy was computed. To counter unbalanced classes (more samples in one class than in the other) weighted accuracy was used (Hastie et al., 2001). For that purpose an additional model was trained and tested on a shuffled version of the data in each CV loop.

After obtaining the results of the 100 repetitions of the outer CV loop, we assessed whether the models performed statistically significantly better than chance level by applying a bootstrap test (100,000 bootstrap samples; Efron, 1979). For this test, the null-hypothesis was that the difference between accuracy and chance level is on average smaller or equal to zero (Table 2).

Further, we analyzed the importance of each factor for the overall model performance. In specific, the factor importance was estimated by summing up contributions per factor, over the decision tree splits. The total importance of a feature was then computed as the normalized importance of that feature averaged over the trees in the ensemble (Hastie et al., 2001). Correlation of features leads to a split of this importance measure among these features (see Figure S1 in Supplementary Material for correlation map). To determine whether a feature's contribution was statistically significant, we tested that feature's importance against the feature importance obtained by a model that was trained with the same parameters, but shuffled data. The null-hypothesis tested per feature was that the median difference in feature importance is smaller or equal to zero. The null-hypothesis was tested with a bootstrap test (100,000 bootstrap samples per feature;

Efron, 1979). Obtained p-values were Bonferroni-corrected for multiple comparisons.

The entire analysis (computing the models and the contributions of factors) was carried out two times. First, to predict neurofeedback performance and a second time to predict neurofeedback improvement.

3. Results

3.1. Neurofeedback success

When investigating neurofeedback performance, we observed that 69.41% of all participants were labelled as successful, meaning that for them, more than 50% of all neurofeedback training runs were successful. Only 9.70% of participants were characterized by 25% or less successful runs. On average, participants presented 72.36% successful neurofeedback runs. For neurofeedback improvement over runs, we observed an average slope of 0.09 across all participants. Here, 59.70% of the participants showed a positive slope and, therefore, were able to improve their neurofeedback performance over time (see Fig. 1).

3.2. Prediction accuracy of neurofeedback performance and neurofeedback improvement

The *Extra Trees* machine learning model was able to predict neurofeedback performance from the investigated factors with an average accuracy of 60.3%, which is significantly better than the average accuracy at chance level with 51% ($p<0.001$). However, no prediction better than chance was revealed for neurofeedback improvement (Table 2).

As only the neurofeedback performance measure could be predicted with a better than chance accuracy, only the influence of factors on neurofeedback performance, but not neurofeedback improvement, are valid to be interpreted. Consequently, normalized model-based feature importance was only calculated for the neurofeedback performance target, but not for the neurofeedback improvement target (see Fig. 2). Two factors contributed significantly to the prediction result: whether a study included a pre-training no-feedback run (median relative importance 59.3%; Fig. 2) and whether a participant was a patient or a healthy participant (median relative importance 31.1%; Fig. 2). More specifically, including a pre-training no-feedback run, as well as performing a study with patients increases the chance for a successful neurofeedback run. Patients did not show a higher ratio of pre-training no-feedback runs (17.47%) than healthy individuals (23.22%).

4. Discussion

In this exploratory study, we investigated the influence of 20 different factors on neurofeedback performance and improvement, including three subject-specific factors, six ROI-based factors, and eleven paradigm-specific factors. When targeting neurofeedback performance, our classification model achieved an accuracy of 60.3%, which was significantly better than chance level. In contrast, classification for the neurofeedback improvement target did not reach an accuracy level above chance level. This difference in classification accuracy between neurofeedback performance and improvement is driven by the different types of success which are captured by the two success measures: Neurofeedback performance describes if an individual is able to regulate their brain signals to a level above (or below if downregulation is performed) baseline. In contrast, neurofeedback improvement describes whether an
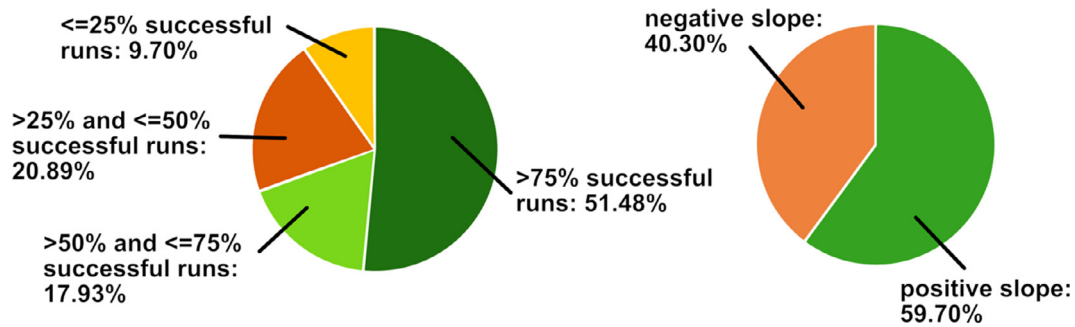
Fig. 1. Distribution of neurofeedback success. Left: More than half (51.48%) of all participants undergoing neurofeedback training performed 75% or more of successful runs. Only 9.70% of the participants performed 25% successful runs or fewer. **Right:** 59.70% of all participants undergoing neurofeedback training show positive slopes of their learning curves, indicating an improvement over time.
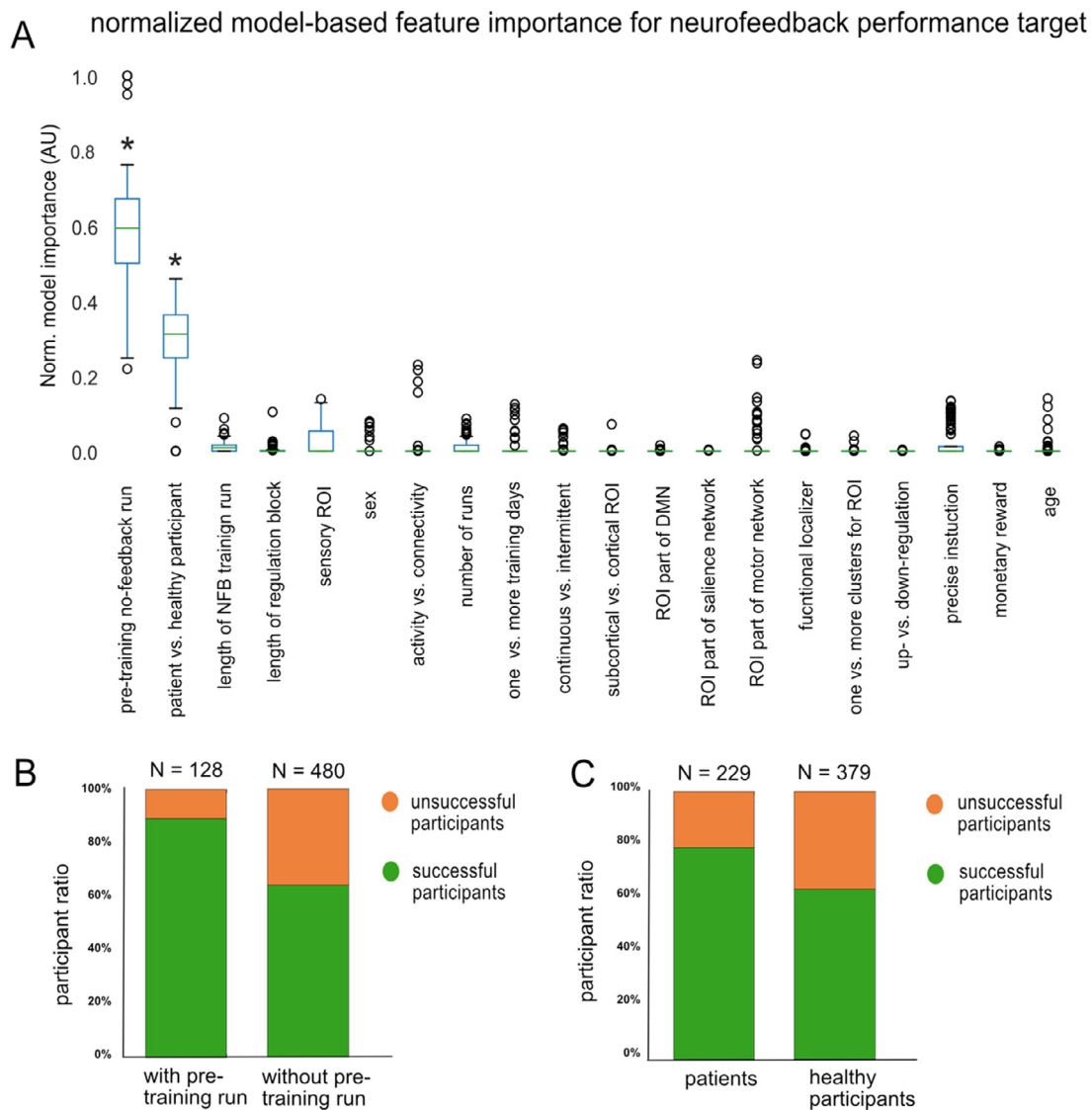

Fig. 2. Distribution of feature importance for predicting neurofeedback performance. A 20 different features were investigated with respect to their importance for the prediction result. Significance was determined using a bootstrapping approach comparing each feature's importance for the model against the feature's importance for the same model with shuffled data. Significant normalized model-based feature importance was observed for the feature pre-training no-feedback run and for the feature patient versus healthy participant. **B** Participants who performed a pre-training no-feedback run were more successful during neurofeedback than participants without a pre-training no-feedback run. **C** Patients were more successful than healthy participants.

individual is able to increase (or decrease in downregulation is performed) their neurofeedback signal over time. Consequently, some individuals might show high levels of performance, but low levels of improvement, or vice versa.

Overall, we observed considerably high neurofeedback success rates across all 28 studies, with around 60% of all participants showing positive slopes and around 70% of all participants showing more than 50% successful neurofeedback runs. Our results revealed two factors that showed high model importance for the neurofeedback performance classification, suggesting that they may significantly influence neurofeedback performance.

4.1. Factors that influence neurofeedback performance

The first significant factor influencing neurofeedback performance is the presence or absence of a pre-training no-feedback run. Here, significantly higher ratios of successful neurofeedback runs were found for studies that included a pre-training no-feedback run in their study design. Pre-training no-feedback runs are functional runs prior to neurofeedback training, where participants are already asked to modulate their brain signals, however, no feedback regarding regulation performance is provided (e.g. see Kim et al., 2015; Kirschner et al., 2018; MacInnes et al., 2016; Young et al., 2017). These no-feedback runs can serve several purposes, for instance, helping participants familiarize themselves with the neurofeedback paradigm and scanning environment where the following runs will take place. Importantly, they serve as a baseline run for comparisons with subsequent neurofeedback training runs and transfer no-feedback runs after neurofeedback training (Auer et al., 2015; MacInnes et al., 2016). One reason for our finding that pre-training no-feedback runs can benefit neurofeedback performance might be that prior familiarization with the neurofeedback setup and an additional run to practice one's brain regulation strategies will make it easier for the participants to perform well.

The second factor that demonstrated significant model importance for neurofeedback performance classification was whether a healthy participant or a clinical patient was undergoing neurofeedback training. Specifically, we found that patients showed higher ratios of successful neurofeedback runs than healthy participants. Similar results have already been reported in an empirical neurofeedback study where the authors observed significantly higher default mode network (DMN) upregulation performance in a heterogeneous group of patients, compared to healthy controls (Skouras and Scharnowski, 2019). The authors argued that this finding might be linked to higher observed scores in DMN eigenvector centrality in the patient group than in the control group, i.e. in the patient group the DMN was more strongly connected to the rest of the brain. This is in line with a recent suggestion by Bassett and Khambhati who argue that areas which are strongly functionally connected within the brain (such as it is the case for the DMN) might be easier to be trained with neurofeedback (Bassett and Khambhati, 2017). Further, it is also possible that patients show better performance in neurofeedback regulation due to more dysfunctional brain patterns as compared to healthy subjects, leaving more room for regulation and making ceiling effects less likely. Here, it should be noted that neurofeedback performance might still differ significantly between different patient populations, due to differences in cognitive deficits which might attenuate attention in general and neurofeedback regulation performance in specific (Heeren et al., 2014; Li et al., 2010; Lussier and Stip, 2001). Further, the observed differences in neurofeedback performance between patients and healthy participants might also be driven by differences in the experimental design. Neurofeedback paradigms in clinical populations have oftentimes been piloted more thoroughly, and sometimes even follow a series of several neurofeedback studies in healthy populations which serve as pilots or templates for implementing the optimized final neurofeedback patient studies. For instance, Kirschner et al., (2018) trained participants with cocaine use disorder to regulate their dopaminergic midbrain using a paradigm that had been

previously successfully applied to healthy participants (Sulzer et al., 2013). Consequently, high risk studies that are more likely to show a high percentage of unsuccessful neurofeedback runs, e.g. studies using a novel analysis method or an ultra-high-field MRI scanner, might be less often performed with patient populations. Finally, also a difference in the participants' motivation might influence the better performance of patients as compared to healthy participants. Many patients undergo neurofeedback training in the hope to improve their clinical symptoms while healthy participants mainly participate out of generic interest or in order to receive a monetary compensation. Therefore, it is likely that patients put more effort into the neurofeedback regulation task than healthy participants.

Taken together, our results indicate that it would be beneficial to include a pre-training no-feedback run in order to improve neurofeedback performance. Furthermore, our results demonstrate better neurofeedback performance of patients as compared to healthy participants. While the participant sample is primarily defined by the biological/clinical question under investigation and, thus, does not constitute an open parameter regarding design optimization, this finding nevertheless has strong implications for the design of future neurofeedback studies. Further, our findings emphasize the clinical potential of neurofeedback interventions: Even in cases where only small or moderate effects have been observed in neurofeedback studies on healthy participants, effects in patients might be nonetheless considerably stronger and clinically relevant, based on the same neurofeedback paradigm.

4.2. Features that do not predict neurofeedback performance

Most of the features included in the machine learning analysis did not play an important role with regards to the classification of participants, neither for neurofeedback performance nor neurofeedback improvement analyses. One reason for this might be that the majority of our included features were based on parameters specific for each study's design, such as information on the paradigm or chosen ROI(s), rather than subject-specific features. These design-specific features were deliberately chosen for our analysis to identify parameters that could be easily modified when designing future neurofeedback studies. However, neurofeedback success also varied considerably within single neurofeedback studies (Bray et al., 2007; Chiew et al., 2012; deCharms et al., 2005; Haugg et al., 2020; Johnson et al., 2012; Ramot et al., 2016; F. Robineau et al., 2014; Scharnowski et al., 2012; Yoo et al., 2008), despite all design-specific parameters being identical for the participants of a study. This indicates that subject-specific factors such as biological measures (e.g. heart rate, pulse, stress level), personality traits and cognitive measures, intelligence, the ability to perform mental imagery, or the subject's attention and motivation (see (Cohen and Staunton, 2019) for a systematic review) might be important factors for successful neurofeedback training. Further, also individual brain-based measures, such as functional connectivity (Scheinost et al., 2014), eigenvector centrality (Skouras and Scharnowski, 2019), gray matter volume (Zhao et al., 2021), or the connectivity of the trained brain region to other higher-order cognitive areas (Bassett and Khambhati, 2017) have been previously discussed as possible factors that might influence neurofeedback success. Due to such information not being available for our data, we were not able to assess the effect that these parameters might have on neurofeedback success. In the future, more harmonization efforts in assessing subject-specific data across different neurofeedback studies will therefore be necessary.

A complementary reason why many features included in our analysis were not predictive of neurofeedback success was the heterogeneity of the dataset. As we aimed at finding generalizable factors that influence neurofeedback success across a wide range of different neurofeedback studies, we purposely included diverse studies training different ROIs, different participant populations, and using a variety of experimental designs and methods, thus making predictions very difficult. It is possible that by investigating more homogeneous subsets of the data, cer-

tain additional factors might become predictive even though they were not predictive when pooling all studies together. However, establishing more homogeneous subsets for solid machine learning analyses will require more data than is currently available. For future investigations on this topic, we therefore want to further encourage data sharing and open science practices and, in particular, emphasize the importance of common practices for data sharing and the description of datasets. This includes use of standardized data structures such as the Brain Imaging Data Structure (BIDS) (Gorgolewski et al., 2016) in general and the reporting of standardized neurofeedback design aspects, such as described in CRED-NF (Ros et al., 2020), in specific.

Finally, it should be noted that the selection of papers who contribute data to mega-analyses and meta-analyses could be biased in a sense that successful results might be more likely to be shared by authors than unsuccessful results. In the future, the use of study registrations and preregistered reports will be an important step to reducing bias in secondary analyses (Allen and Mehler, 2019), not only in the field of neurofeedback, but across all scientific disciplines.

4.3. Neurofeedback success target measures

Our results were most likely not only driven by the included features, but also by the chosen target measures for neurofeedback success. To date, no commonly accepted measure for neurofeedback success has been established and measures vary between different studies (Haugg et al., 2020; Paret et al., 2019; Thibault et al., 2018). For instance, neurofeedback feedback values during a single neurofeedback regulation block or run can be assessed with a wide variety of different methods, such as percent signal change, beta values, or connectivity values. The heterogeneity of feedback values complicates machine learning approaches that require a common target feature. Even if we had access to the raw imaging data, post-hoc re-analyses with an identical analysis pipeline for all studies would not solve this problem, because such a measure would not reflect the feedback that was provided to the participants during training. Choosing neurofeedback performance and neurofeedback improvement as targets for this mega-analysis allowed for pooling this large set of heterogeneous studies, thus, increasing statistical power and generalizability. In addition, by using a dichotomous classification approach (e.g. positive vs. negative slope), we could, further, account for some of the heterogeneity of our data. For instance, when the slope of a neurofeedback learning curve is computed based on only two runs, the resulting values are more likely to be actual outliers, as compared to when the slope of a neurofeedback learning curve based on 20 runs is calculated (Kwak and Kim, 2017). We avoided this problem by using a classification-based instead of a regression-based machine learning approach.

However, the chosen success measures still show limitations. For instance, the neurofeedback performance measure does not control for the fact that the performance of a cognitive task alone can already lead to activation of a specific target region. Currently, only few studies include control groups which perform the exactly same cognitive tasks and undergo MRI scanning, so more data will be necessary to perform machine learning analyses which are explicitly corrected for cognitive task activation. However, while we cannot exclude that neurofeedback learning in studies without suitable control conditions might in fact be due to cognitive task activations, previous studies indicate that neurofeedback learning goes beyond cognitive task activation alone. For instance, Haugg and colleagues observed that brain activity levels during pre-training no-feedback runs did not correlate with neurofeedback success, not even with neurofeedback success within the first neurofeedback run right after the no-feedback run (Haugg et al., 2020). This implies that the provided feedback might have a stronger effect on neurofeedback training success than the actual task which is performed in the MR scanner.

Furthermore improvement regarding the heterogeneity of the neurofeedback success measures might be expected from developing and establishing a commonly accepted model of neurofeedback learning. To date, the underlying mechanisms of neurofeedback have not been fully determined (Cohen and Staunton, 2019; Sitaram et al., 2016), making it difficult to identify the most important attributes of neurofeedback learning, towards creating a comprehensive neurofeedback success measure. With more neurofeedback data becoming publicly available thanks to the open science movement, another solution might be to only consider studies that used exactly the same feedback success measure while still finding enough data to carry out similar analyses. A larger amount of neurofeedback data will also allow future studies to investigate target measures with particular clinical relevance, such as clinical improvement after neurofeedback training, behavioral and/or cognitive changes, or specific neurophysiological responses.

5. Conclusion

With 59.70% of all participants showing positive slopes and 69.41% of all participants having more than 50% of successful neurofeedback runs, our data indicate that neurofeedback training is overall successful, although with large room for improvement. Using machine learning on the largest neurofeedback data set to date, we were able to identify two measures that might influence neurofeedback success and, thus, could lead to improvements in the efficacy of neurofeedback interventions: Participants who performed a pre-training no-feedback run prior to neurofeedback training and participants who were patients generally performed better. Nevertheless, the medium overall predictability of our analyses indicates that further studies based on larger datasets and including more features are needed. In the future, our mega-analysis machine learning approach combined with increased data availability from homogeneous studies might allow for identifying more crucial factors, designing more efficient neurofeedback studies, improving clinical neurofeedback-based interventions, and understanding better how learning with neurofeedback is accomplished.

Credit author statement

Amelie Haugg: conceptualization, methodology, formal analysis, data curation, writing of original draft, funding acquisition, project administration, visualization

Fabian M. Renz: conceptualization, data curation, writing: review and editing

Andrew A. Nicholson: conceptualization, writing: review and editing

Cindy Lor: conceptualization, data curation, writing: review and editing

Sebastian J. Götzendorfer: conceptualization, data curation, writing: review and editing

Ronald Sladky: conceptualization, writing: review and editing

Stavros Skouras: writing: review and editing, data curation, methodology

Frank Scharnowski: conceptualization, methodology, resources, writing of original draft, funding acquisition, supervision

David Steyrl: conceptualization, methodology, software, validation, formal analysis, writing of original draft, visualization

All other authors: writing: review and editing, data curation, formal analysis

Data availability

Due to ethical restrictions, the majority of the included data cannot be made publicly available, but will be shared upon request. Data from the study by McDonald et al. is publicly available as part of the Nathan Kline Institute-Rockland Sample

(http://fcon_1000.projects.nitrc.org/indi/enhanced/). Data from the study by Kohl et al. has been made publicly available on the Open Science Framework (https://osf.io/3agvn/).

Declaration of Competing Interest

SHK receives payment from Mendi Innovations AB, Stockholm, Sweden.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.118207.

References

Alkoby, O., Abu-Rmileh, A., Shriki, O., Todder, D., 2017. Can we predict who will respond to neurofeedback? A review of the inefficacy problem and existing predictors for successful EEG neurofeedback learning. Neuroscience doi:10.1016/j.neuroscience.2016.12.050.

Allen, C., Mehler, D.M.A, 2019. Open science challenges, benefits and tips in early acreer and beyond. PLoS Biol. 17 (5), e3000246.

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 (October), 547. doi:10.3389/fnhum.2015.00547.

Bassett, D.S., Khambhati, A.N., 2017. A network engineering perspective on probing and perturbing cognition with neurofeedback. Ann. N. Y. Acad. Sci. 1396, 126–143. doi:10.1111/nyas.13338.

Bauer, C.C.C., Okano, K., Gosh, S.S., Lee, Y.J., Melero, H., Angeles, C.de los, …, 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 (January). doi:10.1016/j.psychres.2020.112770.

Bray, S., Shimojo, S., O'Doherty, J.P, 2007. Direct instrumental conditioning of neural activity using functional magnetic resonance imaging-derived reward feedback. J. Neurosci. 27 (28), 7498–7507. doi:10.1523/JNEUROSCI.2118-07.2007.

Buyukturkoglu, K., Rana, M., Ruiz, S., Hackley, S.A., Soekadar, S.R., Birbaumer, N., Sitaram, R, 2013. Volitional regulation of the supplementary motor area with fMRI-BCI neurofeedback in Parkinson's disease: a pilot study. 2013 6th International IEEE/EMBS Conference on Neural Engineering (NER) doi:10.1109/NER.2013.6696025.

Buyukturkoglu, Korhan, Roettgers, H., Sommer, J., Rana, M., Dietzsch, L., Arikan, E.B., …, Ruiz, S., 2015. Self-regulation of anterior insula with real-time fMRI and its behavioral effects in obsessive-compulsive disorder: a feasibility study. PLoS ONE 10 (8), 1–26. doi:10.1371/journal.pone.0135872.

Canterberry, M., Hanlon, C.A., Hartwell, K.J., Li, X., Owens, M., LeMatty, T., …, George, M.S., 2013. Sustained reduction of nicotine craving with real-time neurofeedback: exploring the role of severity of dependence. Nicotine Tobacco Res. 15 (12), 2120–2124. doi:10.1093/ntr/ntt122.

Cawley, G.C., Talbot, N.L.C, 2010. On over-fitting in model selection and subsequent selection bias in performance. Evaluation 11, 2079–2107.

Chiew, M., LaConte, S.M., Graham, S.J., 2012. Investigation of fMRI neurofeedback of differential primary motor cortex activity using kinesthetic motor imagery. Neuroimage 61 (1), 21–31. doi:10.1016/j.neuroimage.2012.02.053.

Cohen, K., Staunton, G., 2019. NeuroImage a systematic review of the psychological factors that in fl uence neurofeedback learning outcomes. Neuroimage 185, 545–555. doi:10.1016/j.neuroimage.2018.10.021.

DeBettencourt, M.T., Cohen, J.D., Lee, R.F., Norman, K.A., Turk-Browne, N.B., 2015. Closed-loop training of attention with real-time brain imaging. Nat. Neurosci. 18 (3), 470–478. doi:10.1038/nn.3940.

deCharms, R.C., Maeda, F., Glover, G.H., Ludlow, D., Pauly, J.M., Soneji, D., …, Mackey, S.C., 2005. Control over brain activation and pain learned by using real-time functional MRI. Proc. Natl. Acad. Sci. 102 (51), 18626–18631. doi: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. doi:10.1016/j.neubiorev.2020.12.020.

Efron, B., 1979. Bootstrap methods: another look at the jackknife. Ann. Stat. 7 (1), 1–26.

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

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

Frank, S., Lee, S., Preissl, H., Schultes, B., Birbaumer, N., Veit, R., 2012. The obese brain athlete: self-regulation of the anterior insula in adiposity. PLoS ONE 7 (8), 3–8. doi:10.1371/journal.pone.0042570.

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 (JUN), 1–11. doi:10.3389/fpsyt.2016.00111.

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Mach. Learn 63 (1), 3–42. doi:10.1007/s10994-006-6226-1.

Gorgolewski, K.J., Auer, Tibor, Calhoun, V.D., Craddock, R.C., Das, S., Duff, E.P., …, Poldrack, R.A., 2016. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci. Data, 3 doi:10.1007/978-1-4020-6754-9_1720.

Guan, M., Ma, L., Li, L., Tong, L., Zhang, Y., Zheng, D., …, Shi, D., 2014. Self-regulation of rACC activation in patients with postherpetic neuralgia: a preliminary study using real-time fMRI neurofeedback. Ismrm 22, 5889. doi:10.7910/DVN/27368.

Haller, S., Birbaumer, N., Veit, R., 2010. Real-time fMRI feedback training may improve chronic tinnitus. Eur. Radiol. 20 (3), 696–703. doi:10.1007/s00330-009-1595-z.

Hanlon, C.A., Hartwell, K.J., Canterberry, M., Li, X., Owens, M., LeMatty, T., …, George, M.S., 2013. Reduction of cue-induced craving through realtime neurofeedback in nicotine users: the role of region of interest selection and multiple visits. Psychiatry Res. Neuroimag. 213 (1), 79–81. doi:10.1016/j.pscychresns.2013.03.003.

Hartwell, K.J., Hanlon, C.A., Li, X., Borckardt, J.J., Canterberry, M., Prisciandaro, J.J., …, Brady, K.T., 2016. Individualized real-time fMRI neurofeedback to attenuate craving in nicotine-dependent smokers. J. Psychiatry Neurosci. 41 (1), 48–55. doi:10.1503/jpn.140200.

Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning. Springer doi:10.1111/j.1532-5415.1984.tb02220.x.

Haugg, A., Sladky, R., Skouras, S., McDonald, A., Craddock, C., Kirschner, M., …, Scharnowski, F., 2020. Can we predict real-time fMRI neurofeedback learning success from pretraining brain activity? Hum. Brain Mapp. 41 (14), 3839–3854. doi:10.1002/hbm.25089.

Heeren, A., Maurage, P., Perrot, H., De Volder, A., Renier, L., Araneda, R., …, Philippot, P., 2014. Tinnitus specifically alters the top-down executive control sub-component of attention: evidence from the Attention Network Task. Behav. Brain Res. 269, 147–154. doi:10.1016/j.bbr.2014.04.043.

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

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. 1–29. doi:10.1002/hbm.25010.

Johnson, K., Hartwell, K.J., Lematty, T., Borckardt, J., Morgan, P., Govindarajan, K., …, George, M.S., 2012. Intermittent "Real-time" fMRI feedback is superior to continuous presentation for a motor imagery task: a pilot study. J. Neuroimag. 22 (1), 58–66. doi:10.1111/j.1552-6569.2010.00529.x.

Karch, S., Keeser, D., Hümmer, S., Paolini, M., Kirsch, V., Karali, T., …, Pogarell, O., 2015. Modulation of craving related brain responses using real-time fMRI in patients with alcohol use disorder. PLoS ONE 10 (7). doi:10.1371/journal.pone.0133034.

Kim, D.-.Y., Yoo, S.-.S., Tegethoff, M., Meinlschmidt, G., Lee, J.-.H., 2015. The inclusion of functional connectivity information into fMRI-based neurofeedback improves its efficacy in the reduction of cigarette cravings. J. Cogn. Neurosci. 27 (8), 1552–1572. doi:10.1162/jocn.

Kirschner, M., Sladky, R., Haugg, A., Stämpfli, P., Jehli, E., Hodel, M., Herdener, M., 2018. Self-regulation of the dopaminergic reward circuit in cocaine users with mental imagery and neurofeedback. EBioMedicine 37, 489–498. https://doi.org/10.1016/j.ebiom.2018.10.052.

Kohl, S.H., Veit, R., Spetter, M.S., Günther, A., Rina, A., Lührs, M., …, Hallschmid, M., 2019. NeuroImage 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. doi:10.1016/j.neuroimage.2019.02.033.

Koush, Y., Meskaldji, D.-.E., Pichon, S., Rey, G., Rieger, S.W., Linden, D.E.J., …, Scharnowski, F., 2015. Learning control over emotion networks through connectivity-based neurofeedback. Cereb. Cortex doi:10.1093/cercor/bhv311.

Kwak, S.G., Kim, J.H., 2017. Central limit theorem: the cornerstone of modern statistics. Korean J. Anesthesiol. 70 (2), 144–156. doi:10.4097/kjae.2017.70.2.144.

Li, C.T., Lin, C.P., Chou, K.H., Chen, I.Y., Hsieh, J.C., Wu, C.L., …, Su, T.P., 2010. Structural and cognitive deficits in remitting and non-remitting recurrent depression: a voxel-based morphometry study. Neuroimage 50 (1), 347–356. doi:10.1016/j.neuroimage.2009.11.021.

Li, X., Hartwell, K.J., Borckardt, J., Prisciandaro, J.J., Saladin, M.E., Morgan, P., …, George, M.S., 2013. Volitional reduction of anterior cingulate cortex activity produces

decreased cue craving in smoking cessation: a preliminary real-time fMRI study. Addict. Biol. 18 (4), 739–748. doi:10.1111/j.1369-1600.2012.00449.x.Volitional.

Linden, D.E.J., Habes, I., Johnston, S.J., Linden, S., Tatineni, R., Subramanian, L., …, Goebel, R., 2012. Real-time self-regulation of emotion networks in patients with depression. PLoS ONE 7 (6). doi:10.1371/journal.pone.0038115.

Lussier, I., Stip, E., 2001. Memory and attention deficits in drug naive patients with schizophrenia. Schizophr. Res. 48 (1), 45–55. doi:10.1016/S0920-9964(00)00102-X.

MacInnes, J.J., Dickerson, K.C., Chen, N.kuei, Adcock, R.A., 2016. Cognitive neurostimulation: learning to volitionally sustain ventral tegmental area activation. Neuron 89 (6), 1331–1342. doi:10.1016/j.neuron.2016.02.002.

Marins, T.F., Rodrigues, E.C., Engel, A., Hoefle, S., Basílio, R., Lent, R., …, Tovar-Moll, F., 2015. Enhancing motor network activity using real-time functional MRI neurofeedback of left premotor cortex. Front. Behav. Neurosci. 9 (December), 1–12. doi:10.3389/fnbeh.2015.00341.

Marxen, M., Jacob, M.J., Müller, D.K., Posse, S., Ackley, E., Hellrung, L., …, Smolka, M.N., 2016. Amygdala regulation following fMRI-neurofeedback without instructed strategies. Front. Hum. Neurosci. 10, 1–14. doi:10.3389/fnhum.2016.00183, APR2016.

Mathiak, K.A., Koush, Y., Dyck, M., Gaber, T.J., Alawi, E., Zepf, F.D., …, Mathiak, K., 2010. Social reinforcement can regulate localized brain activity. Eur. Arch. Psychiatry Clin. Neurosci. 260 (2), 132–136. doi:10.1007/s00406-010-0135-9.

McDonald, A.R., Muraskin, J., Dam, N.T.V., Froehlich, B., Puccio, B., Pellman, J., …, Craddock, R.C., 2017. The real-time fMRI neurofeedback based stratification of Default Network Regulation Neuroimaging data repository. Neuroimage 146, 157–170. doi:10.1016/j.neuroimage.2016.10.048.

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 (March). doi:10.3389/fnhum.2015.00160.

Mehler, D.M.A., Sokunbi, M.O., Habes, I., Barawi, K., Subramanian, L., Range, M., …, 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. doi:10.1038/s41386-018-0126-5.

Mehler, D.M.A., Williams, A.N., Whittaker, J.R., Krause, F., Lührs, M., Kunas, S., …, Linden, D.E.J., 2020. Graded fMRI neurofeedback training of motor imagery in middle cerebral artery stroke patients: a preregistered proof-of-concept study. Front. Hum. Neurosci. 14 (July), 1–17. doi:10.3389/fnhum.2020.00226.

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

Nicholson, A.A., Rabellino, D., Densmore, M., Frewen, P.A., Paret, C., Kluetsch, 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. doi:10.1002/hbm.23402.

Pamplona, G.S.P., Heldner, J., Langner, R., Koush, Y., Michels, L., Ionta, S., …, Salmon, C.E.G., 2020. Network-based fMRI-neurofeedback training of sustained attention. Neuroimage 221 (July), 117194. doi:10.1016/j.neuroimage.2020.117194.

Papoutsi, M., Magerkurth, J., Josephs, O., Pépés, S.E., Ibitoye, T., Reilmann, R., …, Tabrizi, S.J., 2020. Activity or connectivity? A randomized controlled feasibility study evaluating neurofeedback training in Huntington's disease. Brain Commun. (1) doi:10.1093/braincomms/fcaa049.

Papoutsi, M., Weiskopf, N., Langbehn, D., Reilmann, R., Rees, G., Tabrizi, S.J., 2018. Stimulating neural plasticity with real-time fMRI neurofeedback in Huntington's disease: a proof of concept study. Hum. Brain Mapp. 39 (3), 1339–1353. doi:10.1002/hbm.23921.

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

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. doi:10.1016/j.neuroimage.2019.116107.

Paret, C., Hendler, T., 2020. Live from the "regulating brain": harnessing the brain to change emotion. Emotion 20 (1), 126–131. doi:10.1037/emo0000674.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., …, Cournapeau, D., 2011. Scikit-learn: machine learning in python. J. Mane Learn. Res. Learn. Res. 12 (1), 2825–2830. doi:10.1145/2786984.2786995.

Quevedo, K., Yuan Teoh, J., Engstrom, M., Wedan, R., Santana-Gonzalez, C., Zewde, B., …, Cohen Kadosh, K., 2020. Amygdala circuitry during neurofeedback training and symptoms' change in adolescents with varying depression. Front. Behav. Neurosci. 14 (July), 1–16. doi:10.3389/fnbeh.2020.00110.

Ramot, M., Grossman, S., Friedman, D., Malach, R., 2016. Covert neurofeedback without awareness shapes cortical network spontaneous connectivity. Proc. Natl. Acad. Sci. 113 (17), E2413–E2420. doi:10.1073/pnas.1516857113.

Robineau, F., Rieger, S.W., Mermoud, C., Pichon, S., Koush, Y., Van De Ville, D., …, Scharnowski, F., 2014. Self-regulation of inter-hemispheric visual cortex balance through real-time fMRI neurofeedback training. Neuroimage 100, 1–14. doi:10.1016/j.neuroimage.2014.05.072.

Robineau, Fabien, Saj, A., Neveu, R., Van De Ville, D., Scharnowski, F., Vuilleumier, P., 2019. Using real-time fMRI neurofeedback to restore right occipital cortex activity in patients with left visuo-spatial neglect: proof-of-principle and preliminary results. Neuropsychol. Rehabil. 29 (3), 339–360. doi:10.1080/09602011.2017.1301262.

Ros, T., Enriquez-Geppert, S., Zotev, V., Young, K.D., Wood, G., Whitfield-Gabrieli, S., …, 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. doi:10.1093/brain/awaa009.

Rota, G., Sitaram, R., Veit, R., Erb, M., Weiskopf, N., Dogil, G., Birbaumer, N., 2009. Self-regulation of regional cortical activity using real-time fmri: the right inferior frontal gyrus and linguistic processing. Hum. Brain Mapp. 30 (5), 1605–1614. doi:10.1002/hbm.20621.

Scharnowski, F., Hutton, C., Josephs, O., Weiskopf, N., Rees, G., 2012. Improving visual perception through neurofeedback. J. Neurosci. 32 (49), 17830–17841. doi:10.1523/JNEUROSCI.6334-11.2012.

Scharnowski, F., Veit, R., Zopf, R., Studer, P., Bock, S., Diedrichsen, J., …, Weiskopf, N., 2015. Manipulating motor performance and memory through real-time fMRI neurofeedback. Biol. Psychol. 108, 85–97. doi:10.1016/j.biopsycho.2015.03.009.

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 (September), 338. doi:10.3389/fnbeh.2014.00338.

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. Human Brain Mapping 37 (9), 3153–3171. doi:10.1002/hbm.23228. https://doi.org/10.1002/hbm.23228.

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

Shibata, K., Watanabe, T., Sasaki, Y., Kawato, M., 2011. Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. Science 334 (6061), 1413–1415. doi:10.1126/science.1212003.Perceptual.

Sitaram, R., Ros, T., Stoeckel, L.E., Haller, S., Scharnowski, F., Lewis-Peacock, J., …, Sulzer, J., 2016. Closed-loop brain training: the science of neurofeedback. Nat. Neurosci. doi:10.1038/nrn.2016.164.

Sitaram, R., Veit, R., Stevens, B., Caria, A., Gerloff, C., Birbaumer, N., Hummel, F., 2012. Acquired control of ventral premotor cortex activity by feedback training: an exploratory real-time fMRI and TMS study. Neurorehabil. Neural. Repair 26 (3), 256–265. doi:10.1177/1545968311418345.

Skouras, S., Scharnowski, F., 2019. The effects of psychiatric history and age on self-regulation of the default mode network. Neuroimage 198, 150–159. doi:10.1016/j.neuroimage.2019.05.008.

Sorger, B., Kamp, T., Weiskopf, N., Peters, J.C., Goebel, R., 2018. When the brain takes 'BOLD' steps: real-time fMRI neurofeedback can further enhance the ability to gradually self-regulate regional brain activation. Neuroscience 378, 71–88. doi:10.1016/j.neuroscience.2016.09.026.

Spetter, M.S., Malekshahi, R., Birbaumer, N., Lührs, M., van der Veer, A.H., Scheffler, K., …, Hallschmid, M., 2017. Volitional regulation of brain responses to food stimuli in overweight and obese subjects: a real-time fMRI feedback study. Appetite 112, 188–195. doi:10.1016/j.appet.2017.01.032.

Subramanian, L., Hindle, J.V., Johnston, S., Roberts, M.V., Husain, M., Goebel, R., Linden, D., 2011. Real-time functional magnetic resonance imaging neurofeedback for treatment of parkinson's disease. J. Neurosci. 31 (45), 16309–16317. doi:10.1523/JNEUROSCI.3498-11.2011.

Sukhodolsky, D.G., Walsh, C., Koller, W.N., Eilbott, J., Rance, M., Fulbright, R.K., …, 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. doi:10.1016/j.biopsych.2019.07.035.

Sulzer, J., Sitaram, R., Blefari, M.L., Kollias, S., Birbaumer, N., Stephan, K.E., Gassert, R., 2013. Neurofeedback-mediated self-regulation of the dopaminergic midbrain. Neuroimage 75, 176–184. doi:10.1016/B978-0-12-384719-5.00424-X.

Thibault, R.T., MacPherson, A., Lifshitz, M., Roth, R.R., Raz, A., 2018. Neurofeedback with fMRI: a critical systematic review. Neuroimage 172, 786–807. doi:10.1016/j.neuroimage.2017.12.071.

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. doi:10.1016/j.neubiorev.2021.02.015.

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

Wang, T., Mantini, D., Gillebert, C.R., 2018. The potential of real-time fMRI neurofeedback for stroke rehabilitation: a systematic review. Cortex 107, 148–165. doi:10.1016/j.cortex.2017.09.006.

Yao, S., Becker, B., Geng, Y., Zhao, Z., Xu, X., Zhao, W., …, Kendrick, K.M., 2016. Voluntary control of anterior insula and its functional connections is feedback-independent and increases pain empathy. Neuroimage 130, 230–240. doi:10.1016/j.neuroimage.2016.02.035.

Yoo, S.-S., Lee, J.-.H., O'Leary, H., Panych, L., Jolesz, F.A, 2008. Neurofeedback fMRI-mediated learning and consolidation of regional brain activation during motor imagery. Int. J. Imag. Syst. Technol. 18 (1), 69–78. doi:10.1002/ima.20139.Neurofeedback.

Young, K.D., Siegle, G.J., Zotev, V., Phillips, R., Misaki, M., Yuan, H., …, 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 (20) doi: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). doi:10.1371/journal.pone.0088785.

Zhang, G., Yao, L., Zhang, H., Long, Z., Zhao, X., 2013. Improved working memory performance through self-regulation of dorsal lateral prefrontal cortex activation using real-time fMRI. PLoS ONE 8 (8), 1–9. doi:10.1371/journal.pone.0073735.

Zhao, X., Zhang, H., Song, S., Ye, Q., Guo, J., Yao, L., 2013. Causal interaction following the alteration of target region activation during motor imagery training using real-time fMRI. Front. Hum. Neurosci. 7 (DEC), 1–8. doi:10.3389/fnhum.2013.00866.

Zhao, Z., Yao, S., Zweerings, J., Zhou, X., Zhou, F., Kendrick, K.M., …, Becker, B., 2021. Putamen volume predicts real-time fMRI neurofeedback learning success across paradigms and neurofeedback target regions. Hum. Brain Mapp. 1879–1887. doi:10.1002/hbm.25336.

Zhi, L., Zhang, C., Huang, J., Wang, Y., Yan, C., Li, K., …, Chan, R.C.K., 2018. Improving motivation through real-time fMRI-based self-regulation of the nucleus accumbens this. Neuropsychology 32 (6), 764–776. doi:10.15199/48.2015.01.45.

Zich, C., Johnstone, N., Lührs, M., Lisk, S., Haller, S.P., Lipp, A., …, Kadosh, K.C., 2020. Modulatory effects of dynamic fMRI-based neurofeedback on emotion regulation networks in adolescent females. Neuroimage 220 (June), 117053. doi:10.1016/j.neuroimage.2020.117053.

Zilverstand, A., Sorger, B., Sarkheil, P., Goebel, R., 2015. fMRI neurofeedback facilitates anxiety regulation in females with spider phobia. Front. Behav. Neurosci. 9 (June), 1–12. doi:10.3389/fnbeh.2015.00148.