Distributed Patterns of Brain Activity Underlying Real-Time fMRI Neurofeedback Training

Rotem Kopel*, Kirsten Emmert, Frank Scharnowski, Sven Haller, and Dimitri Van De Ville, *Senior Member, IEEE*

***Abstract*—Neurofeedback (NF) based on real-time functional magnetic resonance imaging (rt-fMRI) is an exciting neuroimaging application. In most rt-fMRI NF studies, the activity level of a single region of interest (ROI) is provided as a feedback signal and the participants are trained to up or down regulate the feedback signal. NF training effects are typically analyzed using a confirmatory univariate approach, i.e., changes in the target ROI are explained by a univariate linear modulation. However, learning to self-regulate the ROI activity through NF is mediated by distributed changes across the brain. Here, we deploy a multivariate decoding model for assessing NF training effects across the whole brain. Specifically, we first explain the NF training effect by a posthoc multivariate model that leads to a pattern of coactivation based on 90 functional atlas regions. We then use cross validation to reveal the set of brain regions with the best fit. This novel approach was applied to the data from a rt-fMRI NF study where the participants learned to down regulate the auditory cortex. We found that the optimal model consisted of 16 brain regions whose coactivation patterns best described the training effect over the NF training days. Cross validation of the multivariate model showed that it generalized across the participants. Interestingly, the participants could be clustered into two groups with distinct patterns of coactivation, potentially reflecting different NF learning strategies. Overall, our findings revealed that multiple brain regions are involved in learning to regulate an activity in a single ROI, and thus leading to a better understanding of the mechanisms underlying NF training.**

***Index Terms*—Basal ganglia, coactivation, functional networks, insula, neurofeedback (NF), real-time functional magnetic resonance imaging (rt-fMRI), thalamus.**

I. Introduction

NEUROFEEDBACK (NF) based on real-time functional magnetic resonance imaging (rt-fMRI) is an emerging technique that allows us to train the participants voluntary control over their own brain activity [25], [26]. In the majority of rt-fMRI NF experiments, the feedback signal reflects neuronal activity within a single region-of-interest (ROI) and the participants are taught to up or down regulate the feedback signal. Previous rt-fMRI NF training studies have demonstrated that healthy participants can indeed gain control over localized brain activity, and that such training affects behavior. For example, training of the parahippocampal cortex modulated memory function [77] of the right anterior cingulate cortex (ACC) reduced pain perception [22], [62], of the precentral gyrus speeded up motor responses [31], of the inferior frontal gyrus improved linguistic performance [1], of the insula modulated emotions [5], of the occipital cortex improved visual perception [29], [33], and of the right auditory cortex modulated auditory perception [9], [19]. Recent studies have also demonstrated clinical relevance of rt-fMRI-based NF training. For example, chronic pain patients were trained to regulate the ACC [62], chronic tinnitus patients learned control over the auditory cortex [20], Parkinson's disease patients learned control over the supplementary motor area [10], major depression patients learned to increase activity in brain regions involved in positive emotions [28], chronic stroke patients learned control over the ventral premotor cortex [28], nicotine addicts learned control over the ACC and ROIs in the prefrontal cortex [34], and schizophrenia patients learned control over the insular cortex [15].

To shed light on the neural underpinnings of successful self-regulation, training-related changes in brain networks have been investigated posthoc for some of the above-mentioned ROI-based NF studies. For example, Rota *et al.* analyzed functional connectivity (FC) by using a seed correlation approach that revealed FC changes with the NF target ROI (i.e., the inferior frontal gyrus) as a function of NF training [1], [16]. Their FC analyses revealed changes in pairwise correlations between the NF target ROI and other brain regions, but FC changes between brain regions other than the NF target ROI could not be detected with the seed correlation approach. Ruiz *et al.*

Manuscript received November 11, 2015; revised July 19, 2016; accepted July 23, 2016. Date of publication August 10, 2016; date of current version May 15, 2017. This work was supported in part by the Swiss National Science Foundation (Marie Heim-Vögtlin) under Grant PMCDP2-145442, Grant 320030_127079\1, Grant 320030_147126\1, and Grant BSSG10-155915, in part by the Boninchi Foundation, in part by the Wyss Center Geneva, and in part by the Center for Biomedical Imaging. *Asterisk indicates corresponding author.*

*R. Kopel is with the Faculty of Medicine, University of Geneva, Geneva CH-1211, Switzerl, and also with the Institute of Bioengineering, École Polytechnique Fédérale de Lausanne, Lausanne CH-1015, Switzerland (e-mail: rotem.kopel@epfl.ch).

K. Emmert is with the Faculty of Medicine, University of Geneva.

F. Scharnowski is with the Psychiatric University Hospital, the Neuroscience Center Zürich, and the Zürich Center for Integrative Human Physiology, University of Zürich.

S. Haller is with the Affidea Centre Diagnostique Radiologique de Carouge, with the Department of Surgical Sciences, Radiology, Uppsala University, and also with the Department of Neuroradiology, University Hospital Freiburg.

D. Van De Ville is with the Faculty of Medicine, University of Geneva, and also with the Institute of Bioengineering, École Polytechnique Fédérale de Lausanne.

Digital Object Identifier 10.1109/TBME.2016.2598818

studied changes in effective connectivity that were associated with the NF training of the insular cortex using Granger causality modeling (GCM), which is a connectivity analysis approach in which time series from preselected brain regions are used to predict time courses of another region [3], [12], [13], [15]. They found enhanced effective connectivity of the NF target ROI with other brain areas but similar as that with thebr seed-based approach; Granger causality limits the number of brain regions whose connectivity changes were analyzed. Scharnowski *et al.* found posthoc connectivity changes associated with the NF training of the visual cortex using a psychophysiological interaction (PPI) analysis and dynamic causal modeling (DCM) [33], [69]. Whereas PPI is exploratory and allows us to identify task-related correlation changes with the NF target ROI [70], DCM is hypothesis driven and uses the Bayesian model comparison to compare which network architecture explains the data best [72]. However, similar to the seed-based approach, PPI allows us to analyze only FC changes with the NF target ROI, and similar to the Granger causality, DCM allows us to analyze effective connectivity of only a few predefined ROIs. Finally, Haller *et al.* applied independent component analysis to reveal network changes associated with training the auditory cortex [9], [19]. Their analysis revealed functionally relevant independent components (ICs), including the auditory network that contained the NF target ROI, the default mode network (DMN), and the executive control network. Even though this study used a data-driven multivariate approach for the decomposition into ICs, the FC analysis was done using a univariate approach (i.e., changes in pairwise correlations between IC time courses over training days were analyzed separately). Finally, a recent study by Harmelech *et al.* demonstrated that even a single session of ROI-based NF training induced lasting changes of FC within the DMN [27]. Overall, these posthoc analyses of global changes associated with ROI-based NF training indicate that the effects of NF training extend beyond the target ROI. However, these investigations focused on either hypothesis-driven multivariate analyses observing connectivity changes between the target ROI and a limited number of regions, or on whole brain univariate analyses observing changes in pairwise connections. Here, we extend these previous approaches by proposing a posthoc multivariate decoding method to identify changes in brain activity across the whole brain that are associated with the NF training. This decoding analysis includes activity signals from all brain regions, which are used to explain the NF training effects in a multivariate model. Specifically, we

1) used a linear model but with a multivariate decoding setup, where time courses of all brain regions as regressors, and the interaction between the self-regulation paradigm and a linear improvement across NF training as the target signal;

2) ranked all brain regions according to their consistency over subjects;

3) tested this hypothesis that also includes an estimation of the optimal model size by implementing a leave-one-participant-out (LOO) cross validation and backward elimination methods.

Our new multivariate approach will decode and identify coactivation maps of brain regions involved in improved self-regulation skills, and will thus reveal the coactivation patterns underlying ROI-based NF training. To validate the method, we applied this novel approach to data from a previous ROI-based NF study, in which 12 participants learned down regulation of the right auditory cortex activity while being presented with acoustic stimulation [9], [19].

II. Materials and Methods

A. Data Description

Details about the participants, task procedure, and data acquisition can be found in [9] and [19]. For completeness, the main parameters are summarized here.

1) Participants: Twelve healthy right-handed volunteers with normal audition took part in the study. Mean age of participants was 28.37 years (range 24–33 years). The study was approved by the local ethics committee, and all participants gave written informed consent. Before the experiment, volunteers received written instructions explaining that they will learn to down regulate their primary auditory cortex activity with the help of NF.

2) Task Procedure: Each participant had four days (sessions) of NF training with approximately 1 week intervals between them. At each day, participants had four runs, which led to a total of 16 runs per participant. Before each training session, the participants primary auditory area was identified using a standard fMRI auditory block-design paradigm, consisting of 20 s ON and 20 s OFF bilateral auditory stimulation using a 1000 Hz pulsating sine tone, repeated five times. Following the localizer, four rt-fMRI NF training runs were performed, the right localized auditory region was the target ROI. Each run was composed of five 30 s baseline blocks, interleaved with 60 s down-regulation blocks. The same pulsating sine tone of 1000 Hz was provided as an auditory stimulation during down-regulation blocks. The signal from the target ROI was provided as a visual feedback during the entire run as a moving line graph. The participants were informed about the data processing delay of about 1 s and of the intrinsic physiological hemodynamic response delay of about 6 s during the down-regulation periods. No specific regulation strategy was recommended to the participants, but it was emphasized that they should find an individual strategy that worked best for them. For online data analysis and feedback presentation we used the Turbo BrainVoyager software package (Brain Innovation, Maastricht, The Netherland) in combination with in-house MATLAB (Mathworks Inc., Natick, MA, USA) scripts.

3) Data Acquisition: The experiment was performed on a 3 T whole-body MR sc anner with a standard 12-channel head coil (Siemens Magnetom Verio, Siemens Erlangen, Germany). Functional data were acquired with an echo-planar imaging sequence (echo time 40 ms, repetition time 2000 ms, matrix size 64×64, voxel size $3 \times 3 \times 3$ mm^3, and 19 repetitions). Additionally, we acquired an anatomical T1-weighted whole brain image using a magnetization prepared rapid gradient echo sequence (matrix size 256×256, 176 partitions, 1 mm^3 isotropic voxels, 26 slices with 1 mm thickness).

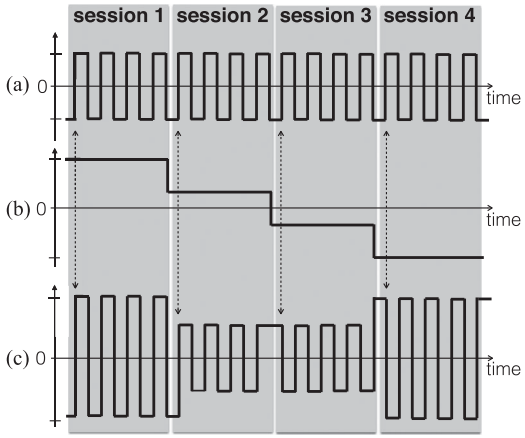

Fig. 1. Construction of the Z_{target} signal. (a) Z_{block} block paradigm of all runs concatenated in time, indicating the regulation and rest periods. (b) Z_{training} a linear descending training signal over four training days. (c) Z_{target} element-wise multiplication between Z_{training} and Z_{block}.

B. Data Preprocessing

Preprocessing was performed using an SPM8 software (Wellcome Department of Imaging Neuroscience, London, UK). All functional volumes were spatially realigned to the first volume of each run and normalized into MNI space (Montreal Neurological Institute, resampled voxel size: $2 \times 2 \times 2$ mm^3) by using cubic B-spline interpolation. Image series were then parcelated into $V = 90$ regions based on the Greicius functional atlas [11], and time courses were regionally averaged, demeaned and linearly detrended using MATLAB standard functions. For each participant n, data were concatenated into a single matrix for all 16 runs: $\overline{Y}^{(n)} = [Y_1^{(n)}\ Y_2^{(n)} \dots\ Y_{16}^{(n)}]$, where $Y_k^{(n)}$ is the spatio-temporal matrix of run k.

C. Within-Subject Model

We deploy a linear model to reveal the main effect of regulation. For each participant n time courses $\overline{Y}^{(n)}$ are used to explain the block paradigm Z_{block}. Z_{block} is a vector of length $16 \cdot 195$, constructed by 16 concatenated block designs, normalized to zero mean and unit variance (see Fig. 1). This linear model is defined for each participant n as follows:

$$Z_{\text{block}} = \overline{Y}^{(n)^T} \cdot \beta_{\text{main}}^{(n)} + \epsilon_{\text{main}}^{(n)} \tag{1}$$

where $\beta_{\text{main}}^{(n)}$ is the parameter vector and $\epsilon_{\text{main}}^{(n)}$ is the noise term. The optimal parameter estimates $\widehat{\beta}_{\text{main}}^{(n)}$ are found by minimizing the sums of squares of the residuals $\widehat{e}_{\text{main}}^{(n)} = Z_{\text{block}} - \overline{Y}^{(n)^T} \times \widehat{\beta}_{\text{main}}^{(n)}$ between the predicted and the fitted models, which is optimal assuming $\epsilon_{\text{main}}^{(n)}$ is independent identically distributed normally distributed.

Next, we deployed another linear model to reveal the training effect over days, here, the time courses $\overline{Y}^{(n)}$ are used to explain the interaction between the linear improvement across training days and the block-design paradigm. For this purpose, we defined the improvement signal as

$$Z_{\text{target}} = Z_{\text{training}} \odot Z_{\text{block}} \tag{2}$$

where $Z_{\text{training}}^T = [\underbrace{3 \dots 3}_{4\cdot195}\ \underbrace{1 \dots 1}_{4\cdot195}\ \underbrace{-1 \dots -1}_{4\cdot195}\ \underbrace{-3 \dots -3}_{4\cdot195}]$. Z_{training} is also normalized to unit variance, and then multiplied element wise ($\odot$) with Z_{block} to generate the improvement signal Z_{target} (see Fig. 1). The linear model for each participant n was defined as follows:

$$Z_{\text{target}} = \overline{Y}^{(n)^T} \cdot \beta^{(n)} + \epsilon^{(n)} \tag{3}$$

where $\beta^{(n)}$ is the parameter vector and $\epsilon^{(n)}$ is the noise term.

To test the null hypothesis so that the improvement signal is explained better than by chance, we phase-randomized the matrix $\overline{Y}^{(n)}$ (along its rows) in the temporal Fourier domain. The null is rejected if the estimation error for the real data is significantly better than for surrogate data. Specifically, generating 19 surrogate datasets allows significance to be established at 5%. Participants that fit the model well are those with real-data residual sum of squares (RSS) values lower than all 19 surrogates RSS values (confidence intervals of 95%). For future use, we defined a $195 \times |\mathcal{N}|$ matrix based on a general set $\mathcal{N}$: $\widehat{\beta}_{\mathcal{N}}^{+} = [\cdots \widehat{\beta}^{j} \cdots],\ j \in \mathcal{N}$. According to the RSS single-participant performances, we defined $\mathcal{M}$ as the set of "good" participants that successfully fitted the model, and the estimated parameter matrix $\widehat{\beta}_{\mathcal{M}}^{+}$ accordingly.

D. Regions Involved in Improvement of Self-Regulation

In order to investigate the key regions involved in the improvement of self-regulation, we analyzed $\widehat{\beta}_{\mathcal{M}}^{+}$, by performing a two-sided one-sample *t*-test on each row (i.e., per brain region across participants). Based on the results, we ranked the regions according to their *t*-values. We also included the anatomical location of the functional network regions (see Table I). Next, the generalizability of the model across participants was tested with a LOO cross-validation scheme (see Fig. 2). In this approach, the fitted model for $|\mathcal{M}| - 1$ training participants was used to establish the corresponding $\widehat{\beta}$, which was then applied to the left-out participant. In each LOO iteration, we selected a left-out participant n, defined the training set $\mathcal{M}^{(n)} = \mathcal{M} \setminus \{n\}$, and updated matrix $\widehat{\beta}_{\mathcal{M}^{(n)}}^{+}$ accordingly. Due to the high interparticipant variability, we clustered $\widehat{\beta}_{\mathcal{M}^{(n)}}^{+}$ into two groups using *k*-means clustering that used the cosine distance measure with 10 000 replicates, resulting in two sets of participants: $\mathcal{M}_i^{(n)} = \{j : j \in \mathcal{M}^{(n)} | j \in \text{ group } i\},\ i = 1, 2$. To avoid double dipping, *k*-mean clustering was calculated within each fold and the number of clusters was consistently set to two based on the Calinski–Harabasz criterion [71]. For each group i, we adjusted the fitted data matrix $\widehat{\beta}_{\mathcal{M}_i^{(n)}}^{+}$, based on which we calculated the training response

$$\overline{\beta}_{\mathcal{M}_i^{(n)}} = 1/(|\mathcal{M}_i^{(n)}|) \cdot \widehat{\beta}_{\mathcal{M}_i^{(n)}}^{+} \cdot \mathbf{1} \tag{4}$$

TABLE I
BRAIN REGIONS RANKED ACCORDING TO THEIR CONSISTENCY OVER SUBJECTS WITHIN EACH GROUP

Group 1 ranking	t	β	anatomical location
1	-6.115	-0.030	Right Middle Occipital Gyrus, Superior Occipital Gyrus
2	-2.889	-0.024	Right Angular Gyrus. Middle Occipital Gyrus
3	-2.888	-0.023	Right Insula
4	2.796	0.012	Left Crus I
5	-2.764	-0.022	Right Caudate
6	-2.753	-0.015	Right Inferior Frontal Gyrus
7	-2.738	-0.006	Right Thalamus Caudate, Putamen
8	-2.625	-0.026	Left Crus l. Crus ll. Lobule Vl
9	2.601	0.021	Right Middle Frontal Gyrus
10	-2.559	-0.032	Right Midcingulate Cortex
11	-2.408	-0.015	Left Frontal Operculum, Inferior Frontal Gyrus
12	-2.082	-0.032	Left Precentral Gyrus. Postcentral Gyrus
13	-2.033	-0.016	Left Parahippocampal Gyrus
14	1.897	0.010	Left Thalamus
15	-1.888	-0.019	Right Superior Frontal Gyrus
16	1.809	0.016	Right Inferior Parietal Gyrus. Supramarginal Gyrus. Angular Gyrus
17	1.761	0.019	Left Supramarginal Gyrus. Inferior Parietal Gyrus
18	1.695	0.039	Right Superior Frontal Gyrus, Middle Frontal Gyrus
19	-1.689	-0.024	Right Superior Frontal Gyrus
20	1.611	0.041	Left Middle Frontal Gyrus. Superior Frontal Gyrus. Precentral Gyrus
21	1.607	0.019	Right Frontal Operculum. Inferior Frontal Gyrus
22	1.595	0.012	Right Lobule VI. Crus I
23	-1.551	-0.020	Precuneus
24	-1.498	-0.015	Right Angular Gyrus
25	1.493	0.037	Left Angular Gyrus
26	1.470	0.030	Left Middle Frontal Gyros. Superior Frontal Gyros
27	-1.463	-0.041	Right Superior Parietal Gyrus. Precuneus
28	1.430	0.015	Right Retrosplenial Cortex. Posterior Cingulate Cortex
29	-1.419	-0.037	Right Superior Frontal Gyrus
30	1.414	0.010	Inferior Frontal Gyrus
31	1.358	0.023	Precuneus
32	1.344	0.003	Pons
33	-1.324	-0.020	Right Crus I
34	-1.286	-0.018	Right Middle Frontal Gyrus. Right Superior Frontal Gyrus
35	1.282	0.015	Left Middle Occipital Gyrus
36	1.171	0.015	Anterior Cingulate Cortex. Medial Prefrontal Cortex. Supplementary Motor Area
37	1.089	0.007	Right Supramarginal Gyrus. Superior Temporal Gyrus. Middle Temporal Gyros
38	-1.087	-0.001	Left Middle Temporal Gyrus
39	-1.085	-0.013	Left Middle Temporal Gyrus. Superior Temporal Gyrus. Supramarginal Gyrus. Angular Gyrus
40	-1.039	-0.019	Left Inferior Parietal Sulcus
41	-1.020	-0.015	Right Middle Frontal Gyrus
42	0.949	0.012	Right Lobule VI. Crus I
43	0.949	0.008	Left Middle Frontal Gyrus.
44	0.919	0.046	Right Supramarginal Gyrus. Inferior Parietal Gyrus
45	-0.909	-0.010	Right Thalamus
46	-0.876	-0.029	Left Superior Parietal Gyrus. Inferior Parietal Gyrus. Precuneus. Angular Gyrus
47	-0.848	-0.006	Midcingulate Cortex. Posterior Cingulate Cortex
48	0.834	0.010	Left and Right Thalamus
49	0.823	0.018	Left Superior Temporal Gyrus Heschrs Gyrus
50	0.789	0.003	Left Inferior Frontal Gyrus
51	-0.764	-0.007	Left inferior Temporal Gyrus. Middle Temporal Gyrus
52	0.763	0.033	Right Supplementary Motor Area
53	-0.761	-0.014	Right Middle Frontal Gyrus
54	0.761	0.012	Right Angular Gyrus
55	0.753	0.004	Right Superior Temporal Gyrus
56	-0.732	-0.011	Right Middle Temporal Gyrus
57	0.700	0.001	Left Angular Gyrus
58	0.692	0.004	Right Thalamus
59	0.664	0.030	Left Retrosplenial Cortex. Posterior Cingulate Cortex
60	-0.649	-0.006	Right Lobule Vlll. Lobule Vllb
61	-0.645	-0.020	Medial Prefrontal Cortex Anterior Cingulate Cortex. Orbitofrontal Cortex
62	0.586	0.006	Calcarine Sulcus
63	0.558	0.008	Right Thalamus
64	-0.532	-0.017	Left Thalamus
65	-0.522	-0.005	Left Middle Frontal Gyrus
66	-0.436	-0.001	Left Thalamus
67	0.434	0.007	Right Parahippocampal Gyrus
68	0.427	0.014	Left Thalamus, Caudate
69	-0.422	-0.010	Right Hippocampus
70	0.392	0.003	Left Inferior Temporal Gyrus
71	-0.371	0.000	Right Posterior Insula
72	0.358	0.009	Right Inferior Parietal Lobule
73	0.357	0.001	Left Middle Occipital Gyrus. Superior Occipital Gyrus
74	-0.348	-0.005	Left Precuneus
75	-0.332	-0.002	Left Lobule Vlll, Lobule Vllb
76	-0.276	-0.005	Left Middle Frontal Gyrus
77	0.271	0.001	Right Lobule IX
78	-0.233	0.000	Left Thalamus
79	0.233	-0.002	Lobule VI
80	0.227	0.001	Left Insula
81	-0.214	-0.003	Left Middle Temporal Gyros. Angular Gyrus
82	0.206	0.002	Left Lobule VI. Crus I
83	-0.187	0.004	Bilateral Lobule IV, Lobule V, Lobule VI
84	-0.182	0.000	Left Hippocampus
85	0.177	-0.002	Right Inferior Frontal Gyrus
86	-0.169	-0.002	Midcingulate Cortex
87	0.162	-0.009	Left Posterior Insula. Putamen
88	0.155	0.005	Right Precentral Gyrus. Postcentral Gyrus
89	-0.134	0.003	Lobule VI
90	0.042	0.006	Left Inferior Frontal Gyrus. Orbitofrontal Gyrus

Group 2 ranking	t	β	anatomical location
1	4.815	0.012	Midcingulate Cortex
2	-3.316	-0.018	Right Angular Gyrus
3	2.860	0.030	Right Middle Frontal Gyrus
4	-2.725	-0.063	Right Inferior Parietal Lobule
5	-2.650	-0.029	Left Inferior Frontal Gyrus. Orbitofrontal Gyrus
6	2.589	0.044	Right Supramarginal Gyrus. Inferior Parietal Gyrus
7	-2.283	-0.025	Left inferior Temporal Gyrus. Middle Temporal Gyrus
8	2.146	0.037	Right Superior Temporal Gyrus
9	-1.925	-0.018	Right Supplementary Motor Area
10	1.888	0.018	Left Middle Temporal Gyrus
11	-1.887	-0.029	Left Retrosplenial Cortex. Posterior Cingulate Cortex
12	-1.828	-0.037	Left Angular Gyrus
13	-1.803	-0.051	Right Middle Occipital Gyrus, Superior Occipital Gyrus
14	1.609	0.028	Left Middle Temporal Gyros. Angular Gyrus
15	1.446	0.010	Midcingulate Cortex. Posterior Cingulate Cortex
16	-1.385	-0.029	Left Middle Frontal Gyrus. Superior Frontal Gyrus. Precentral Gyrus
17	1.382	0.035	Right Retrosplenial Cortex. Posterior Cingulate Cortex
18	1.342	0.038	Left Superior Parietal Gyrus. Inferior Parietal Gyrus. Precuneus. Angular Gyrus
19	1.321	0.020	Right Inferior Parietal Gyrus. Supramarginal Gyrus. Angular Gyrus
20	-1.296	-0.040	Left Superior Temporal Gyrus Heschrs Gyrus
21	-1.279	-0.025	Left Supramarginal Gyrus. Inferior Parietal Gyrus
22	-1.269	-0.021	Right Supramarginal Gyrus. Superior Temporal Gyrus. Middle Temporal Gyros
23	1.253	0.018	Left Middle Occipital Gyrus
24	1.206	0.026	Calcarine Sulcus
25	1.189	0.005	Left Thalamus
26	1.168	0.051	Right Middle Frontal Gyrus. Right Superior Frontal Gyrus
27	1.149	0.015	Left Middle Frontal Gyros. Superior Frontal Gyros
28	-1.129	-0.009	Left Inferior Frontal Gyrus
29	-1.069	-0.012	Right Thalamus Caudate, Putamen
30	1.057	0.014	Left Inferior Parietal Sulcus
31	-1.056	-0.013	Left Hippocampus
32	1.049	0.020	Right Middle Frontal Gyrus
33	-1.048	-0.026	Right Frontal Operculum. Inferior Frontal Gyrus
34	-1.046	-0.007	Left Thalamus
35	1.045	0.007	Right Hippocampus
36	1.026	0.008	Inferior Frontal Gyrus
37	-1.025	-0.016	Right Superior Frontal Gyrus, Middle Frontal Gyrus
38	0.979	0.022	Left Frontal Operculum, Inferior Frontal Gyrus
39	-0.961	-0.007	Right Parahippocampal Gyrus
40	-0.900	-0.010	Right Insula
41	0.888	0.010	Left Precuneus
42	0.885	0.011	Right Posterior Insula
43	-0.882	-0.005	Right Caudate
44	0.854	0.006	Right Lobule Vlll. Lobule Vllb
45	0.827	0.019	Right Superior Parietal Gyrus. Precuneus
46	-0.806	-0.022	Medial Prefrontal Cortex Anterior Cingulate Cortex. Orbitofrontal Cortex
47	-0.806	-0.020	Precuneus
48	0.804	0.011	Left Angular Gyrus
49	0.801	0.012	Left Middle Frontal Gyrus
50	0.764	0.013	Right Thalamus
51	0.728	0.004	Right Thalamus
52	0.722	0.006	Right Midcingulate Cortex
53	-0.674	0.002	Right Middle Frontal Gyrus
54	0.671	0.002	Right Lobule VI. Crus I
55	-0.627	-0.004	Lobule VI
56	0.602	0.009	Left Middle Occipital Gyrus. Superior Occipital Gyrus
57	0.588	0.006	Left Posterior Insula. Putamen
58	-0.584	-0.005	Left Crus I
59	-0.545	-0.003	Left Middle Frontal Gyrus.
60	0.539	0.008	Left Precentral Gyrus. Postcentral Gyrus
61	-0.532	-0.001	Right Lobule IX
62	0.506	0.004	Right Lobule VI. Crus I
63	-0.498	-0.005	Right Precentral Gyrus. Postcentral Gyrus
64	-0.476	-0.007	Lobule VI
65	0.474	0.004	Right Inferior Frontal Gyrus
66	-0.466	-0.010	Left Parahippocampal Gyrus
67	-0.447	-0.010	Right Crus I
68	-0.437	-0.002	Left Lobule VI. Crus I
69	0.401	0.006	Right Angular Gyrus. Middle Occipital Gyrus
70	0.396	0.004	Precuneus
71	-0.388	-0.007	Right Angular Gyrus
72	-0.370	-0.002	Bilateral Lobule IV, Lobule V, Lobule VI
73	-0.341	-0.001	Left Crus l. Crus ll. Lobule Vl
74	0.338	0.000	Right Superior Frontal Gyrus
75	-0.332	-0.016	Anterior Cingulate Cortex. Medial Prefrontal Cortex. Supplementary Motor Area
76	-0.318	-0.001	Left Inferior Temporal Gyrus
77	-0.314	-0.004	Left Middle Frontal Gyrus
78	0.312	0.007	Right Inferior Frontal Gyrus
79	0.310	0.002	Left Thalamus
80	0.308	0.004	Left Thalamus
81	0.295	0.009	Left Insula
82	0.260	0.007	Right Superior Frontal Gyrus
83	-0.245	-0.005	Right Superior Frontal Gyrus
84	0.241	-0.002	Right Middle Temporal Gyrus
85	0.240	-0.003	Left Thalamus, Caudate
86	0.201	0.006	Left and Right Thalamus
87	-0.172	-0.003	Left Middle Temporal Gyrus. Superior Temporal Gyrus. Supramarginal Gyrus. Angular Gyrus
88	0.102	0.001	Pons
89	0.034	-0.002	Left Lobule Vlll, Lobule Vllb
90	-0.019	-0.005	Right Thalamus

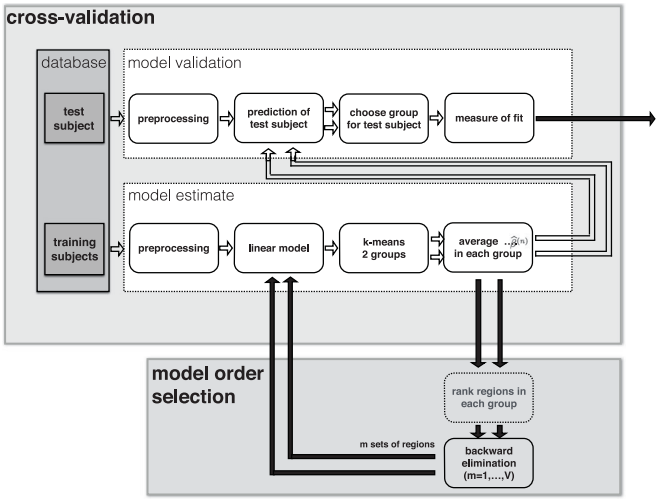

Fig. 2. Overview of the processing pipe line with cross validation. The model estimation was done for all participants but one (test participant). The model validation for test participant used the model estimation results. Backward elimination for brain regions repeats this procedure for all possible model sizes. Regions were eliminated based on the ranking of each set which was calculated once using the full model (i.e., $m = V$). Results are the measure of fit calculated for each possible model order size.

where $\mathbf{1}$ is a one vector of length $|\mathcal{M}_i^{(n)}|$. The prediction for the left-out participant n was defined as follows:

$$\overline{Z}_i^{(n)} = \overline{Y}^{(n)^T} \cdot \overline{\beta}_{\mathcal{M}_i^{(n)}}. \tag{5}$$

Furthermore, we defined the measure of fit as the correlation between the improvement paradigm Z_{target} and the prediction signal

$$\overline{F}_i^{(n)} = \text{corr}\left(Z_{\text{target}}, \overline{Z}_i^{(n)}\right). \tag{6}$$

The group level coactivation maps were calculated separately for each of the two groups using the following steps:

1) rankings $\overline{\beta}_{\mathcal{M}_i^{(n)}}$ according to one-sample *t*-test were calculated for each group in each fold;

2) the optimal model (top 16 regions, see Section II-E) in each fold was identified;

3) a list counting the times each region was a part of the optimal model was created;

4) coactivations maps were defined based on the list created in step (3), regions that were ranked within the top 16 regions for more than 6 out of 11 folds were selected.

E. Model Order Selection

Region ranking for optimal model size analysis was calculated once for each LOO fold. A one-sample *t*-test was calculated for the training set matrix $\overline{\beta}_{\mathcal{M}_i^{(n)}}$ that included all regions (i.e., full model $m = V$) , results were sorted in descending order that sets the ranking of regions. The search for the optimal model order was done by using the backward elimination approach [43] that was executed separately for each fold, where at the initial step all regions are included and in following steps regions are eliminated based on the ranking that was determined. The performances analysis of all possible model sizes was done separately for each fold by exploring the model estimation of the training set (i.e., two group clustering and $\overline{\beta}_{\mathcal{M}_i^{(n)}}$ in each of the groups) and model validation [i.e., the measure of fit $\overline{F}_{i,m}^{(n)}$, (6)]. Based on those results, we could assign the left-out participant n to one of the two clusters $\mathcal{M}_i^{(n)}$, $i = 1, 2$ by maximizing the mean measure of fit between the two clusters: $\{n \in \mathcal{M}_k^{(n)} \mid \sum_{j=1}^{V} \overline{F}_{k,j}^{(n)} > \sum_{j=1}^{V} \overline{F}_{t,j}^{(n)}\}$. The performance results of all the folds were consolidated by averaging the measure of fit over all folds, resulting in one measure for each possible model size. The selection of the optimal model size was based on comparing that average measure of fit to the performances of 1000 surrogate datasets, and choosing the order size that maximizes the data performance and exceed the 98th percentile of the surrogate distribution.

III. Results

A. Learned Down Regulation of Primary Auditory Cortex

As reported previously, over the course of four days of NF training, participants learned to down-regulate activity in the NF target region, i.e., the right auditory cortex [9], [19].

During the experiment, feedback was only based on the activity level of a single subject-specific target single ROI that was determined by using a functional localizer. Here, we investigated training effects based on the combined activity from multiple brain regions that were taken from a predefined functional atlas.

B. Within-Subject Discriminative Model

In the analysis of the main effect, we found positive bilateral activation in the auditory regions and frontal medial cortex, contralateral thalamic activity, caudate and middle occipital lobe (left regions with activations while right regions with deactivation); and bilateral deactivation of insula, medial superior frontal gyrus, and calcarine. Next, for the training effect analysis we found that all except one participant showed significant training effect as function of training days (see Fig. 3) $\mathcal{M} = \{1, 2 \ldots 9, 11, 12\}$. For those who did, the linear combination of activity in the V regions was related to changes in regulation strength across NF days. Table I shows the brain regions ranked according to the *t*-values and also reports the average linear combination values (i.e., β values).

C. Regions Involved in Improving Self-Regulation

A complete view of the learning effect needs to consider two results, the first is the main effect that corresponds to the average activation level during down-regulation session and the second is the change of activation over training sessions. The multivariate analysis of the training shows the coactivated regions that explain a linear change in activation over sessions. The clustering analysis revealed two distinct groups of participants. Cross validation identified distinct sets of brain regions that were involved in training (see Fig. 4). Here, we closely examine the two groups: the optimal model for group 1 includes left crus l, right

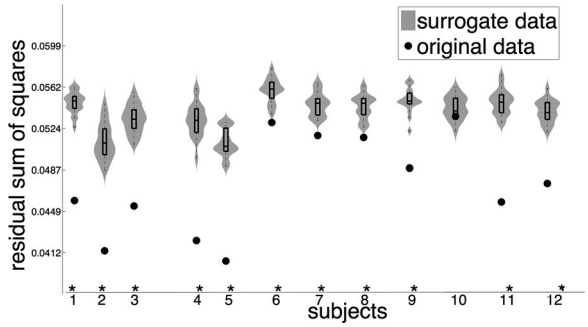

Fig. 3. RSS for all participants. RSS of the original data (circles) was significantly lower than the RSS of the surrogate data ($\diamond$) for eight participants (1, 2, 3, 4, 5, 9, 11, 12; indicated by $*$). Participants 6, 7, and 8 present low performances but still within the 95% significant levels (also indicated by $*$). One participant (number 10) did not show a significant learning effect across NF training.

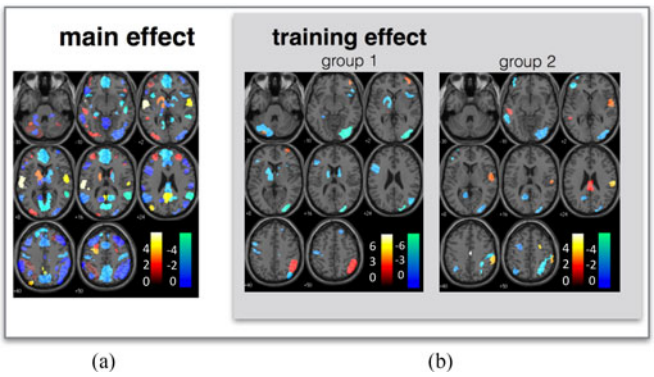

Fig. 4. Coactivation maps shown as axial slices in neurological convention. (a) *T*-value of the main effect analysis for all brain regions. (b) Optimal model for the rt-fMRI training effect over four sessions. Shown here are the positive and negative *t*-values averaged across the within-subject multivariate models.

insula, caudate, right midcingulate cortex, right angular gyrus, right thalamus, and putamen areas that are associated with negative beta values (an increase of activation over the session), while left frontal operculum, left crus I, and right middle frontal gyrus are associated with positive beta values (a decrease of activation over the session). In the optimal model for group 2 it was found that right inferior parietal lobule, right angular gyrus, left inferior temporal gyrus, and left inferior frontal gyrus are associated with negative beta values and an increase of activation, while midcingulate cortex, left middle temporal gyrus, right superior temporal gyrus (i.e., the target auditory ROI), and right supramarginal gyrus are associated with positive beta values and thus a decrease of activation.

D. Model Order Selection

The performance of the proposed model was evaluated using the average measure of fit, i.e., correlation between the improvement paradigm and the prediction for the test participant. Each test participant n was assigned to its best fitted group. In Fig. 5, we present these cross-validation results for the real and surrogate data. The distribution of the test statistic under the null

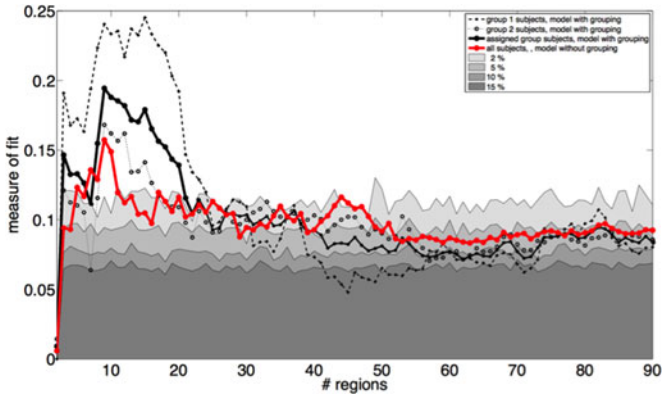

Fig. 5. Performance of cross validation over different model sizes. Measure of fit indicate average group level correlation between Z_{target} and test participants prediction. Dashed line represents the results for group 1, dotted line for group 2; solid line for assigned group, i.e., each test participant was assigned to the group with the best fit; red line without subdividing into two groups, i.e., for each left-out subject, the model is estimated based on all other subjects. Gray scale represents the percentile of the surrogate distribution.

hypothesis indicated that 16 regions can be considered as an optimal model order. An examination of the optimal models across the folds revealed that the *t*-values used for the ranking of the top 16 regions where significant (two-tailed uncorrected t-value > 2, $p < 0.05$) It is important to note that for the optimal model size the cross-validation results indicated that the two groups analysis with real training data outperformed the analysis including all subjects with real data (i.e., without subdividing subjects in two groups) as well as with surrogate data (see Fig. 5). Despite the high interparticipant variability, this further confirmed that the top-ranked regions can be generalized to out-of-fold participants.

IV. Discussion

Voluntary control over brain activity in a single ROI can be learned using rt-fMRI NF. Here, we deployed a multivariate data-driven model to reveal how the coactivation of multiple brain regions explains the successful NF training. For evaluating the consistency of the activated regions across participants, we used cross validation to determine the most economical and generalizable model, which consisted of 16 brain regions. Our results show that

1) NF training of a single ROI caused distributed changes across the whole brain.

2) A multivariate model of coactivated brain regions that generalizes across participants can be identified.

3) Participants can be clustered into two distinct groups who each coactivated different sets of brain regions.

A. Posthoc Analyses of Functional Network Reorganization

Training brain activity in a single ROI using rt-fMRI NF does not only affect the NF target ROI, but also other regions across the brain. A better characterization of these changes is important for understanding the neural underpinnings of NF training, thus potentially improving its efficacy. As presented

in Section I, the methods that were previously used to investigate global distributed changes include seed-based correlation, whole-brain pairwise correlation, GCM, DCM, and PPI. Each of these approaches has its advantages and disadvantages. For example, GCM and DCM allow us to determine the directionality of connectivity changes, and DCM allows for modeling of effective connectivity at the neuronal level [72]. On one hand, these multivariate approaches are limited to analyzing connectivity changes of only a limited number of brain regions that have to be defined *a priori*. On the other hand, seed-based correlation, whole-brain pairwise correlation, and PPI can handle more brain regions, but they can consider only pairwise connections, and therefore cannot detect interactions between multiple brain regions. Our proposed multivariate approach is complementary in that it allows for investigating changes related to NF training by activity traces across the entire brain. This does not require any prior assumption about how brain regions interact. Such multivariate interactions that are characterized by coactivation patterns and that are specific to the improvement of self-regulation across NF training days could not be revealed by the previously used methods. Therefore, our approach extends previous investigations of changes related to NF training.

Previous analyses were all carried out on the group level, even though NF learning success, strategies, and its neural underpinning vary substantially between participants [25]. We examined generalization and consistency of the results using cross validation, as it was implemented in our approach that allows ensuring and quantifying how the coactivation maps generalize across participants. To account for the high interparticipant variability of learning in NF experiments, we applied clustering procedures where participants were clustered according to similarities in their coactivation maps. Our analysis revealed two distinct and consistent groups of participants, which showed differences and commonalities in NF learning across participants.

B. Self-Regulation of Activity in Auditory Cortex Implicates Distributed Set of Regions

When applying the proposed analysis method to data from an NF training study where participants learned to down regulate the ROI auditory cortex, we found evidence of changes in a distributed set of brain regions, which was associated with NF learning. Specifically, we found that a coactivation model with 16 brain regions best explained the NF training effect. The clustering analysis had revealed two groups, each with a distinct coactivation pattern. In the first group, the coactivation pattern consisted of brain regions mainly related to self-awareness (e.g., precuneus, insula, angular gyrus) [73]–[75], cognitive control (e.g., frontal operculum) [76], and skill learning (e.g., caudate nucleus, putamen) [85]. Especially the latter is interesting, because they have frequently been reported to be involved in NF experiments [22], [58], [59], [77], [86], and it has recently been proposed that NF learning is linked to skill learning [78]. In contrast, in the second group, the coactivation pattern consisted of brain regions related to the auditory/language pathway (e.g., right superior temporal gyrus, which contains the target ROI, as well left inferior frontal gyrus, and supramarginal gyrus) [79], [80], sensory information processing (e.g., inferior parietal lobe) [81], [82], and reward-related learning (e.g., midcingulate gyrus) [48], [54].

To support the validity of the two group clustering, we have compared the group performances against the one obtained when not separating the subjects into two groups, i.e., for each left-out subject, the model was estimated based on all other subjects. The results suggest that the models for the subgroups were not only different, but also lead to a more accurate model than the one obtained jointly from all subjects. Although interpreting the differences in coactivation patterns between the two subgroups would require extensive meta-analytic profiling, their functional differences already suggest different learning strategies. For example, participants in group 1 might have adopted an explicit skill learning strategy, whereas participants in group 2 were more prone to implicit reinforcement- and reward-based learning [65]. Detailed reports of the cognitive strategies that participants used during the NF experiment would be useful to further elaborate this speculation, but no such reports are available for this study. Those coactivations maps could potentially be used for additional purposes, for example, monitor self-regulation aptness. Recent method proposed to detect online arousal level using only fMRI data [87] can be adapted for this purpose, coactivations maps could be projected onto online rt-fMRI volumes and generating rt-fMRI self-training index.

C. Limitations

The first limitation is that we used a linear model as a first-order approximation of the NF learning effect. Although the cross-validation results confirm that the linear assumption holds, it might have not been the optimal model.

The second limitation is that in this multivariate analysis we included all brain regions, including the right auditory region based on the Greicius atlas. Since the atlas-based definition of the auditory region differs from that of the auditory NF target ROI (which was based on a functional localizer for each participant), these regions are not identical and thus there is no "double-dipping." Including the atlas auditory ROI in the model allows discovering whether it plays a role (or not) in the learning effects of NF.

Finally, our sample size is low for clustering the participants into two groups. However, our unsupervised clustering showed a clear separation into two groups, and despite the lower sample size per group, the cross-validation results are superior in the subgroups compared to all subjects combined. Since intersubject variability in training strategies is an important topic in NF training studies, future rt-fMRI studies in larger samples might use similar clustering approaches to identify different learning strategies.

D. Conclusion

The proposed multivariate approach revealed interactions between distributed brain regions that contributed to learning control over a ROI through NF training. Using a cross-validation scheme, we examined the generalization and consistency of the model, as well as similarities and differences between NF learn-

ing strategies across participants. Our results suggest that future NF research could exploit distributed information in the brain to improve the efficiency of the NF signal [18], [83], [84], or to monitor and even guide the control strategy used by the participants. Finally, this approach is not limited to analyzing data from NF experiments, but can in principle be useful for gaining new insights in other types of longitudinal data from learning experiments.

References

[1] G. Rota *et al.*, "Self-regulation of regional cortical activity using real-time fMRI: The right inferior frontal gyrus and linguistic processing," *Human Brain Mapping*, vol. 30, no. 5, pp. 1605–1614, 2009.

[2] K. C. Berridge *et al.*, "What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience?," *Brain. Res. Brain. Res. Rev.*, vol. 28, no. 3, pp. 309–369, 1998.

[3] B. Abler *et al.*, "Investigating directed influences between activated brain areas in a motor-response task using fMRI," *Magn. Reson. Imag.*, vol. 24, no. 2, pp. 181–185, 2006.

[4] R. A. Barkley, "Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD," *Psychol. Bull.*, vol. 121, no. 1, pp. 65–94, 1997.

[5] A. Caria *et al.*, "Volitional control of anterior insula activity modulates the response to aversive stimuli. A real-time functional magnetic resonance imaging study," *Biol. Psychiatry*, vol. 68, no. 5, pp. 425–432, 2010.

[6] V. Zotev *et al.*, "Self-regulation of human brain activity using simultaneous real-time fMRI and EEG neurofeedback," *Neuroimage*, vol. 85, no. Pt. 3, pp. 985–995, 2014.

[7] V. Zotev *et al.*, "Self-regulation of amygdala activation using real-time FMRI neurofeedback," *PLoS One*, vol. 6, no. 9, 2011, Art. no. e24522.

[8] N. Weiskopf *et al.*, "Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging fMRI: Methodology and exemplary data," *Neuroimage*, vol. 19, no. 3, pp. 577–586, 2003.

[9] D. Van De Ville *et al.*, "Recovery of the default mode network after demanding neurofeedback training occurs in spatio-temporally segregated subnetworks," *Neuroimage*, vol. 63, no. 4, pp. 1775–1781, 2012.

[10] L. Subramanian *et al.*, "Real-time functional magnetic resonance imaging neurofeedback for treatment of Parkinsons disease," *J. Neurosci.*, vol. 31, no. 45, pp. 16309–16317, 2011.

[11] W. R. Shirer *et al.*, "Decoding subject-driven cognitive states with whole-brain connectivity patterns," *Cerebral Cortex*, vol. 22, no. 1, pp. 158–165, 2012.

[12] A. K. Seth, "A MATLAB toolbox for Granger causal connectivity analysis," *J Neurosci. Methods*, vol. 186, no. 2, pp. 262–273, 2010.

[13] A. K. Seth, "Causal connectivity of evolved neural networks during behavior," *Network*, vol. 16, no. 1, pp. 35–54, 2005.

[14] E. Seifritz, "Spatiotemporal pattern of neural processing in the human auditory cortex," *Science*, vol. 297, no. 5587, pp. 1706–1708, 2002.

[15] S. Ruiz *et al.*, "Acquired self-control of insula cortex modulates emotion recognition and brain network connectivity in schizophrenia," *Human Brain. Mapping*, vol. 34, no. 1, pp. 200–212, 2013.

[16] G. Rota *et al.*, "Reorganization of functional and effective connectivity during real-time fMRI-BCI modulation of prosody processing," *Brain Lang.*, vol. 117, no. 3, pp. 123–132, 2011.

[17] M. E. Raichle *et al.*, "A default mode of brain function," *Proc. Natl. Acad. Sci. USA*, vol. 98, no. 2, pp. 676–682, 2001.

[18] Y. Koush *et al.*, "Connectivity-based neurofeedback: Dynamic causal modeling for real-time fMRI," *Neuroimage*, vol. 81, 422–430, 2013.

[19] S. Haller *et al.*, "Dynamic reconfiguration of human brain functional networks through neurofeedback," *Neuroimage*, vol. 81, pp. 243–252, 2013.

[20] S. Haller *et al.*, "Real-time fMRI feedback training may improve chronic tinnitus," *Eur. Radiol.*, vol. 20, no. 3, pp. 696–703, 2010.

[21] S. Haller *et al.*, "Mapping continuous neuronal activation without an ON-OFF paradigm: Initial results of BOLD ceiling fMRI," *Eur. J. Neurosci.*, vol. 24, no. 9, pp. 2672–2678, 2006.

[22] K. Emmert *et al.*, "Comparison of anterior cingulate vs. insular cortex as targets for real-time fMRI regulation during pain stimulation," *Frontier Behav. Neurosci.*, vol. 8, p. 350, 2014.

[23] R. C. de Charms *et al.*, "Control over brain activation and pain learned by using real-time functional MRI," *Proc. Natl. Acad. Sci. USA*, vol. 102, no. 51, pp. 18626–18631, 2005.

[24] N. Weiskopf, "Real-time fMRI and its application to neurofeedback," *Neuroimage*, vol. 62, no. 2, pp. 682–692, 2012.

[25] J. Sulzer *et al.*, "Real-time fMRI neurofeedback: Progress and challenges," *Neuroimage*, vol. 76, pp. 386–399, 2013.

[26] N. Weiskopf, "Real-time fMRI and its application to neurofeedback," *Neuroimage*, vol. 62, no. 2, pp. 682–692, 2012.

[27] T. Harmelech *et al.*, "The day-after effect: Long term, Hebbian-like restructuring of resting-state fMRI patterns induced by a single epoch of cortical activation," *J. Neurosci.*, vol. 33, no. 22, pp. 9488–9497, 2013.

[28] D. E. Linden *et al.*, "Real-time self-regulation of emotion networks in patients with depression," *PLoS One*, vol. 7, no. 6, 2012, Art. no. e38115.

[29] K. Shibata *et al.*, "Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation," *Science*, vol. 334, no. 6061, pp. 1413–1415, 2011.

[30] A. Anderson *et al.*, "Real-time functional MRI classification of brain states using Markov-SVM hybrid models: Peering inside the rt-fMRI black box," in *Machine Learning and Interpretation in Neuroimaging*. Berlin, Germany: Springer, 2012, pp. 242–255.

[31] S. Bray, S. Shimojo, and J. P. O. Doherty, "Direct instrumental conditioning of neural activity using functional magnetic resonance imaging-derived reward feedback," *J. Neurosci.*, vol. 27, no. 28, pp. 7498–7507, 2007.

[32] N. Weiskopf *et al.*, "Self-regulation of local brain activity using real-time functional magnetic resonance imaging fMRI," *J. Physiology*, vol. 98, no. 4, pp. 357–73, 2004.

[33] F. Scharnowski *et al.*, "Improving visual perception through neurofeedback," *J. Neurosci.*, vol. 32, no. 49, pp. 17830–17841, 2012.

[34] X. Li *et al.*, "Volitional reduction of anterior cingulate cortex activity produces decreased cue craving in smoking cessation: A preliminary real-time fMRI study," *Addiction Biol.*, vol. 18, no. 4, pp. 739–748, 2013.

[35] C. A. Seger and C. M. Cincotta, "The roles of the caudate nucleus in human classification learning," *J Neurosci*, vol. 25, no. 11, pp. 2941–2951, 2005.

[36] E. Gunter and R. Romand, *The Central Auditory System*, New York, NY, USA: Oxford Univ. Press, 1997, p. 404.

[37] C. J. Price *et al.*, "Hearing and saying. The functional neuro-anatomy of auditory word processing," *Brain*, vol. 119, no. Pt. 3, pp. 919–931, 1996.

[38] P. M. Grasby *et al.*, "Functional mapping of brain areas implicated in auditory–verbal memory function," *Brain*, vol. 116, no. Pt 1, pp. 1–20, 1993.

[39] H. O. Karnath, B. Baier, and T. Nagele, "Awareness of the functioning of ones own limbs mediated by the insular cortex?," *J. Neurosci.*, vol. 25, no. 31, pp. 7134–7138, 2005.

[40] A. D. Craig, "How do you feel-now? The anterior insula and human awareness," *Nature Rev. Neurosci.*, vol. 10, no. 1, pp. 59–70, 2009.

[41] J. G. McHaffie *et al.*, "Subcortical loops through the basal ganglia," *Trends Neurosci.*, vol. 28, no. 8, pp. 401–407, 2005.

[42] B. Draganski *et al.*, "Evidence for segregated and integrative connectivity patterns in the human Basal Ganglia," *J. Neurosci.*, vol. 28, no. 28, pp. 7143–7152, 2008.

[43] G. Isabelle, S. Gunn, and M. Nikravesh, *Feature Extraction: Foundations and Applications [With CDROM] Studies in Fuzziness and Soft Computing*. Berlin, Germany: Springer, 2006, p. 804.

[44] K. J. Friston *et al.*, "Functional topography: Multidimensional scaling and functional connectivity in the brain," *Cerebral Cortex*, vol. 6, no. 2, pp. 156–164, 1996.

[45] A. Petacchi *et al.*, "Cerebellum and auditory function: An ALE meta-analysis of functional neuroimaging studies," *Human Brain Mapping*, vol. 25, no. 1, pp. 118–128, 2005.

[46] A. Brent, *Cingulate Neurobiology and Disease*. London, U.K.: Oxford Univ. Press, 2009, p. 829.

[47] A. J. Shackman *et al.*, "The integration of negative affect, pain and cognitive control in the cingulate cortex," *Nat. Rev. Neurosci.*, vol. 12, no. 3, pp. 154–167, 2011.

[48] J. D. Wallis and S. W. Kennerley, "Heterogeneous reward signals in prefrontal cortex," *Curr. Opin. Neurobiol.*, vol. 20, no. 2, pp. 191–198, 2010.

[49] N. Weiskopf *et al.*, "Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging fMRI: Methodology and exemplary data," *Neuroimage*, vol. 19, no. 3, pp. 577–586, 2003.

[50] R. C. F. deCharms *et al.*, "Control over brain activation and pain learned by using real-time functional MRI," *Proc. Natl. Acad. Sci. USA*, vol. 102, no. 51, pp. 18626–18631, 2005.

[51] J. H. Lee, J. Kim, and S. S. Yoo, "Real-time fMRI-based neurofeedback reinforces causality of attention networks," *Neurosci. Res.*, vol. 72, no. 4, pp. 347–354, 2012.
[52] R. N. Spreng, R. A. Mar, and A. S. Kim, "The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis," *J. Cogn. Neurosci.*, vol. 21, no. 3, pp. 489–510, 2009.
[53] R. A. Mar, "The neural bases of social cognition and story comprehension," *Annu. Rev. Psychol.*, vol. 62, pp. 103–134, 2011.
[54] E. T. Rolls, "The functions of the orbitofrontal cortex," *Brain Cogn.*, vol. 55, no. 1, pp. 11–29, 2004.
[55] T. Hinterberger *et al.*, "Neuronal mechanisms underlying control of a brain-computer interface," *Eur. J. Neurosci.*, vol. 21, no. 11, pp. 3169–3181, 2005.
[56] A. E. Cavanna and M. R. Trimble, "The precuneus: A review of its functional anatomy and behavioural correlates," *Brain*, vol. 129, no. Pt. 3, pp. 564–583, 2006.
[57] G. Zhang, L. Yao, and Z. Long, "Alteration of functional connectivity during real-time fMRI regulation of PCC," *Proc. SPIE*, vol. 8317, pp. 83170C–83170C, Mar. 2012.
[58] S. S. Yoo *et al.*, "Neurofeedback fMRI-mediated learning and consolidation of regional brain activation during motor imagery," *Int. J. Imag. Syst. Technol.*, vol. 18, no. 1, pp. 69–78, 2008.
[59] M. Chiew, S. M. LaConte, and S. J. Graham, "Investigation of fMRI neurofeedback of differential primary motor cortex activity using kinesthetic motor imagery," *Neuroimage*, vol. 61, no. 1, pp. 21–31, 2012.
[60] I. I. Goldberg, M. Harel, and R. Malach, "When the brain loses its self: Prefrontal inactivation during sensorimotor processing," *Neuron*, vol. 50, no. 2, pp. 329–339, 2006.
[61] X. Liu *et al.*, "Common and distinct networks underlying reward valence and processing stages: A meta-analysis of functional neuroimaging studies," *Neurosci. Biobehav. Rev.*, vol. 35, no. 5, pp. 1219–1236, 2011.
[62] R. C. de Charms *et al.*, "Control over brain activation and pain learned by using real-time functional MRI," *Proc. Natl. Acad. Sci. USA*, vol. 102, no. 51, pp. 18626–18631, 2005.
[63] R. C. McCaig *et al.*, "Improved modulation of rostrolateral prefrontal cortex using real-time fMRI training and meta-cognitive awareness," *Neuroimage*, vol. 55, no. 3, pp. 1298–1305, 2011.
[64] S. E. Kober *et al.*, "Learning to modulate ones own brain activity: The effect of spontaneous mental strategies," *Frontier Human Neurosci.*, vol. 7, p. 695, 2013.
[65] J. M. Lacroix, "Mechanisms of biofeedback control: On the importance of verbal conscious processing," in *Consciousness and Self-Regulation: Vol. 4*, R. J. Davidson, G. E. Schwartz and D. Shapiro, Eds. New York, NY, USA: Plenum, vol. 137162, 1986.
[66] C. R. Gallistel and L. D. Matzel, "The neuroscience of learning: Beyond the Hebbian synapse," *Annu. Rev. Psychol.*, vol. 64, pp. 169–200, 2013.
[67] R. L. Buckner, J. R. Andrews-Hanna, and D. L. Schacter, "The brains default network: Anatomy, function, and relevance to disease," *Ann. N. Y. Acad. Sci.*, vol. 1124, pp. 1–38, 2008.
[68] F. Scharnowski *et al.*, "Manipulating motor performance and memory through real-time fMRI neurofeedback," *Biol. Psychol.*, vol. 108, pp. 85–97, 2015.
[69] F. Scharnowski *et al.*, "Connectivity changes underlying neurofeedback training of visual cortex activity," *PLoS One*, vol. 9, no. 3, 2014, Art. no. 91090.
[70] K. J. Friston *et al.*, "Psychophysiological and modulatory interactions in neuroimaging," *Neuroimage*, vol. 6, no. 3, pp. 218–229, 1997.
[71] T. Calinski and J. Harabasz, "A dendrite method for cluster analysis," *Commun. Statist.*, vol. 3, pp. 1–27, 1974.
[72] K. J. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *Neuroimage*, vol. 19, no. 4, pp. 1273–1302, 2003.
[73] M. Tsakiris, M. R. Longo, and P. Haggard, "Having a body versus moving your body: Neural signatures of agency and body-ownership," *Neuropsychologia*, vol. 48, no. 9, pp. 2740–2749, 2010.
[74] C. Farrer *et al.*, "The angular gyrus computes action awareness representations," *Cerebral Cortex*, vol. 18, no. 2, pp. 254–261, 2008.
[75] A. E. Cavanna and M. R. Trimble, "The precuneus: A review of its functional anatomy and behavioural correlates," *Brain*, vol. 129, no. Pt. 3, pp. 564–583, 2006.
[76] T. Higo *et al.*, "Distributed and causal influence of frontal operculum in task control," *Proc. Natl. Acad. Sci. USA*, vol. 108, no. 10, pp. 4230–4235, 2011.
[77] F. Scharnowski *et al.*, "Manipulating motor performance and memory through real-time fMRI neurofeedback," *Biol. Psychol.*, vol. 108, pp. 85–97, 2015.
[78] N. Birbaumer, S. Ruiz, and R. Sitaram, "Learned regulation of brain metabolism," *Trends Cogn. Sci.*, vol. 17, no. 6, pp. 295–302, 2013.
[79] S. R. Arnott *et al.*, "Assessing the auditory dual-pathway model in humans," *Neuroimage*, vol. 22, no. 1, pp. 401–408, 2004.
[80] J. R. Binder *et al.*, "Human brain language areas identified by functional magnetic resonance imaging," *J. Neurosci.*, vol. 17, no. 1, pp. 353–362, 1997.
[81] J. Downar *et al.*, "A multimodal cortical network for the detection of changes in the sensory environment," *Nature Neurosci.*, vol. 3, no. 3, pp. 277–283, 2000.
[82] C. Alain, Y. He, and C. Grady, "The contribution of the inferior parietal lobe to auditory spatial working memory," *J. Cogn. Neurosci.*, vol. 20, no. 2, pp. 285–295, 2008.
[83] A. Zilverstand *et al.*, "Windowed correlation: A suitable tool for providing dynamic fMRI-based functional connectivity neurofeedback on task difficulty," *PLoS One*, vol. 9, no. 1, 2014, Art. no. e85929.
[84] S. Y. Kim *et al.*, "The inclusion of functional connectivity information into fMRI-based neurofeedback improves its efficacy in the reduction of cigarette cravings," *J. Cogn. Neurosci.*, vol. 27, no. 8, pp. 1552–1572, 2015.
[85] R. A. Poldrack and J. D. Gabrieli, "Characterizing the neural mechanisms of skill learning and repetition priming: evidence from mirror reading," *Brain.*, vol. 124, no. Pt. 1, pp. 67–82, 2001.
[86] K. Emmert *et al.*, "Meta-analysis of real-time fMRI neurofeedback studies using individual participant data: How is brain regulation mediated?" *Neuroimage*, vol. 124, no. Pt A, pp. 806–812, 2015.
[87] C. Chang *et al.*, "Tracking brain arousal fluctuations with fMRI," *Proc. Natl. Acad. Sci. USA*, vol. 113, no. 16, pp. 4518–4523, 2016.

Rotem Kopel received her bachelor and M.S. degrees in Electrical Engineering from Tel-Aviv University Israel in 2001 and 2006, respectively. From 2000 to 2009, she worked in Texas Instruments Inc. in various positions, including research and program management. Currently, she is a PhD student in Ecole Polytechnique Fédérale de Lausanne (EPFL), Institute of Bioengineering, School of Engineering. Under the supervision of Prof. Dimitri van-de-Ville and Sven Haller. Her research interest includes rt-fMRI methodology, different functional connectivity measures, and their application on neurofeedback.

Frank Scharnowski received the B.Sc. degree with honors in Cognitive Science, the M.Sc. degree in Neural and Behavioral Sciences with honors from the Max Planck Research School in Tübingen, Germany, and the Ph.D. degree from the Swiss Institute of Technology Lausanne, Switzerland, in 2001, 2004, and 2007, respectively. He was Senior Fellow at the UCL Wellcome Trust Centre for Neuroimaging and the Institute of Cognitive Neuroscience, London, UK. In 2011 he became Senior Lecturer at the University of Geneva and the Swiss Institute of Technology Lausanne, Switzerland. Currently he is Assistant Professor at the University of Zürich, Switzerland. His research interests cover brain imaging techniques, (real-time) fMRI, brain-computer interfaces, neurofeedback, cognitive neuroscience, and neurological as well as psychiatric disorders. He was awarded a Marie-Curie Fellowship of the European Union, an Ambizione Fellowship, and a Swiss National Science Foundation Professorship.

Dimitri Van De Ville (M'02–SM'12) received the M.S. degree in engineering and computer sciences and the Ph.D. degree from Ghent University, Belgium, in 1998, and 2002, respectively. After a post-doctoral stay (2002-2005) at the lab of Prof. Michael Unser at the Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland, he became responsible for the Signal Processing Unit at the University Hospital of Geneva, Switzerland, as part of the Centre d'Imagerie Biomédicale (CIBM). In 2009, he received a Swiss National Science Foundation professorship and since 2015 became Professor of Bioengineering at the EPFL and the University of Geneva, Switzerland. His research interests include wavelets, sparsity, pattern recognition, and their applications in computational neuroimaging. He was a recipient of the Pfizer Research Award 2012, the NARSAD Independent Investigator Award 2014, and the Leenaards Foundation Award 2016.

Dr. Van De Ville served as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING from 2006 to 2009 and the IEEE SIGNAL PROCESSING LETTERS from 2004 to 2006, as well as Guest Editor for several special issues. He was the Chair of the Bio Imaging and Signal Processing (BISP) TC of the IEEE Signal Processing Society (2012-2013) and is the Founding Chair of the EURASIP Biomedical Image & Signal Analytics SAT. He is Co-Chair of the biennial Wavelets & Sparsity series conferences, together with V. Goyal and M. Papadakis.

Kirsten Emmert obtained her Bachelor of Science in Molecular Medicine from the University of Göttingen (Germany) in 2011. In 2013, she finished her Master of Science in Neuroscience at the International Max Planck Research Schools and the University of Göttingen. She is currently a Ph.D. student at the Lemanic Neuroscience Doctoral School conducting fMRI neurofeedback research under the guidance of Prof. Sven Haller and Prof. Dimitri Van De Ville at the University of Geneva.

Sven Haller had obtained his MD from Medical school at the University of Tübingen, Germany, and a Master of Sciences at the Max Planck research school of Neural and Behavioral in Tübingen, Germany. He had specialized in Radiology and diagnostic Neuroradiology at the University Hospital of Basel, Switzerland. He had completed his Privatdozent (venia docendi, senior lecturer) and obtained the title of Assistant Professor in Neuroradiology at the University Hospital Geneva, Switzerland.

Currently, he serves as a Neuroradiologist at Affidea CDRC Carouge, Switzerland and associate visiting professor at Uppsala University. His interest is mainly in advanced neuroimaging techniques in particular in the domains of neurodegenerative and neurovascular diseases.

Dr. Haller had recived two awards by the European Society of Neuroradiology, two awards by the Swiss Society of Neuroradiology, one award by the Swiss Society of Radiology and one award by the German Society of Neuroradiology.