Learning Control Over Emotion Networks Through Connectivity-Based Neurofeedback

Yury Koush[1,2], Djalel-E. Meskaldji[1,2], Swann Pichon[3,4,5], Gwladys Rey[3], Sebastian W. Rieger[3,4], David E.J. Linden[6], Dimitri Van De Ville[1,2], Patrik Vuilleumier[3,4] and Frank Scharnowski[1,2,7,8]

[1]Institute of Bioengineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1202 Geneva, Switzerland, [2]Department of Radiology and Medical Informatics, [3]Geneva Neuroscience Center, Department of Neuroscience, [4]Swiss Center for Affective Sciences, [5]Faculty of Psychologyand Educational Science,University of Geneva,CH-1211 Geneva, Switzerland, [6]School of Medicine, Institute of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff CF24 4HQ, UK, [7]Department of Psychiatry, Psychotherapy and Psychosomatics, Psychiatric Hospital, University of Zürich, 8032 Zürich, Switzerland and [8]Neuroscience Center Zürich, University of Zürich and Swiss Federal Institute of Technology, 8032 Zürich, Switzerland

Abstract

Most mental functions are associated with dynamic interactions within functional brain networks. Thus, training individuals to alter functional brain networks might provide novel and powerful means to improve cognitive performance and emotions. Using a novel connectivity-neurofeedback approach based on functional magnetic resonance imaging (fMRI), we show for the first time that participants can learn to change functional brain networks. Specifically, we taught participants control over a key component of the emotion regulation network, in that they learned to increase top-down connectivity from the dorsomedial prefrontal cortex, which is involved in cognitive control, onto the amygdala, which is involved in emotion processing. After training, participants successfully self-regulated the top-down connectivity between these brain areas even without neurofeedback, and this was associated with concomitant increases in subjective valence ratings of emotional stimuli of the participants. Connectivity-based neurofeedback goes beyond previous neurofeedback approaches, which were limited to training localized activity within a brain region. It allows to noninvasively and non pharmacologically change interconnected functional brain networks directly, thereby resulting in specific behavioral changes. Our results demonstrate that connectivity based neurofeedback training of emotion regulation networks enhances emotion regulation capabilities. This approach can potentially lead to powerful therapeutic emotion regulation protocols for neuropsychiatric disorders.

Key words: connectivity-based neurofeedback, dynamic causal modeling (DCM), emotion networks regulation, functional magnetic resonance imaging (fMRI), positive emotions

Introduction

When Bertrand Russell wrote: “Control your emotion, or it will control you”, he expressed emphatically a necessity that we all experience in everyday life: the capacity to control our emotions. Emotion regulation allows us to adaptively cope with negative and positive events; failure to do so can result in burdening affective disorders. The psychological and the neural processes underlying emotion regulation have been intensely studied over the past decade, leading to the formulation of well-established Emotion regulation allows us to adaptively cope with negative and positive events; failure to do so can result in burdening affective disorders. The psychological and the neural processes underlying emotion regulation have been intensely studied over the past decade, leading to the formulation of well-established feasibility of simple correlation measures between brain areas as an index of brain connectivity has been explored offline (Zilverstand et al. 2014) or has been integrated as an add-on to standard activity-based neurofeedback (Kim et al. 2015). Rather than computing simple statistical dependencies for the feedback, we have recently shown, using dynamic causal modeling (DCM), that even causal interactions within brain networks can be used for neurofeedback (Koush et al. 2013). Such a DCM-based approach allows to determine the directionality of connectivity, describes how neural dynamics propagates through a network, and allows for modeling effective connectivity at the neuronal level (Friston et al. 2003; Stephan et al. 2010). Here, for the first time, we used a connectivity-based neurofeedback signal for brain training and applied it to emotion regulation. We hypothesized that such training would allow healthy participants to learn control over specific aspects of the emotion regulation network, and that such training would affect the subjective response to emotional stimuli. To test these hypotheses, we trained 15 healthy volunteers (9 participants in the experimental group; 6 in a matched control group) to voluntarily increase top-down effective connectivity from the dmPFC onto the bilateral amygdala, which is one of the key connections of emotion regulation networks (Banks et al. 2007; Barbas and Zikopoulos 2007; Pessoa 2008; Ochsner 2010; Ochsner et al. 2012) (Fig. 1). Neurofeedback training was accomplished by 1) processing fMRI signals in near real-time, 2) comparing a model representing top-down modulation from the dm PFC onto the amygdala with a model representing bottom-up flow of information from the amygdala onto the dmPFC using dynamic causal modeling (DCM) (Fig. 1C), and 3) providing the participants in the fMRI scanner with online feedback about the dominance of the top-down model (Koush et al. 2013). If the top-down model fitted the ongoing brain activity during a training trial better than the bottom-up model, the feedback signal was positive; if the bottom-up model dominated, the feedback signal was negative (Fig. 1B). Using this connectivity-based neurofeedback information across several successive training sessions, participants attempted to learn, by trial and error and using a freely chosen strategy, to increase the top-down connectivity from the dmPFC onto the amygdala. Traditional imaging research has generally focused on the downregulation of negative emotions, but here we focus the neurofeedback training on upregulating positive emotions. This approach is more relevant for targeting the anhedonia aspect of emotion regulation disorders such as depression and anxiety (Disner et al. 2011; Treadway and Zald 2011), and avoids exposure to disturbing visual scenes that might be problematic in these patients. Hence, participants were presented with images depicting moderately positive social situations during training and asked to appraise the positive content with a personal perspective. The goal of the neurofeedback training was to strengthen the top-down connectivity from the dmPFC onto the amygdala, thus potentially mimicking the effect produced by reappraisal (Ochsner and Gross 2005; Banks et al. 2007; Kim et al. 2011; Zotev et al. 2013).

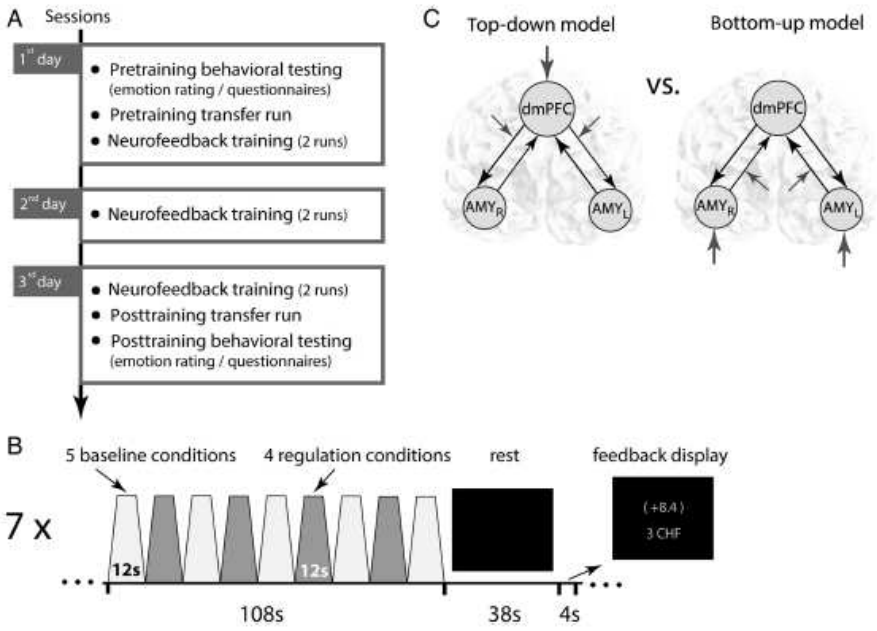

Figure 1. Experimental design. (A) Prior to neurofeedback training, participants rated their subjective responses to target pictures, filled out psychological questionnaires, and performed self-regulation without neurofeedback (pretraining transfer run). They then started neurofeedback training, which took place in six 17.5 min runs spread over 3 days. After training, participants again performed self-regulation in the absence of neurofeedback (posttraining transfer run), rated their subjective responses to target pictures, and filled out psychological questionnaires. (B) Example of a neurofeedback trial. Per neurofeedback training day, participants performed 2 runs, which were each composed of 7 trials. Each neurofeedback trial was composed of 5 baseline and 4 regulation blocks of 12 s each, a rest period of 38 s, and a feedback display lasting 4 s. During the baseline blocks, images of neutral objects were presented, and participants were asked to passively look at them. During regulation blocks, moderately positive social images were presented, and participants were asked to imagine experiencing the depicted positive social situation. During the rest period, a black screen was presented. The feedback display consisted of the logarithmic Bayes factor value (which was red if the trial was successful, i.e., positive, and blue otherwise), and the cumulative reward that had been earned until then. (C) During neurofeedback training, participants learned to voluntarily increase top-down effective connectivity from the dmPFC onto the bilateral amygdala. This was accomplished by providing a feedback signal that indicated the degree of dominance of a top-down model (=target model, left panel) compared with a bottom-up model (right panel).

Materials and Methods

Participants

Fifteen healthy human volunteers (7 male, 8 female, age 26.2 ± 1.4 years) gave written informed consent to participate in the experiment, Which was approved by the local ethics committee. All participants had normal or corrected-to-normal vision and had no prior history of neurological or psychiatric diseases. Before the experiment, participants received written instructions describing that they would perform a 3-day neurofeedback training experiment during which they would be asked to attempt to control the emotion networks. The instructions included explanations of the experimental procedure and of the neurofeedback display. Participants were informed that they should regulate their brain activity to maximize positive feedback and recommended as potential regulation strategies to imagine one-self engaged in the depicted positive social situation. In common with established practice in the neurofeedback field, we were not prescriptive about the strategy that participants should adopt during learning (Sulzer et al. 2013). It was emphasized that participants should find an individual strategy that worked best for them. In addition to a fixed amount of 20 CHF/h for their participation in the experiment, a bonus of 1 CHF was rewarded for each successful neurofeedback trial. Six of the 15 participants were allocated to an age- and gender-matched control group (control group: 3 males, 3 females, 25.7 ± 2.9 years; experimental group: 4 males, 5 females, 26.4 ± 4.7), to determine whether the hypothesized learning effects can also be achieved with unrelated feedback. Participants in this control group were provided with the same instructions and underwent identical training procedures but unknown to them received sham feedback. Sham feedback was derived from the feedback values of one of the 6 best performing participants in the learning group, rather than their own brain activity. Debriefing interviews conducted after the experiment confirmed that the control participants were unaware that they had received sham feedback. After each scanning session, participants were asked to fill in a written questionnaire and, among other questions, describe how they tried to manipulate the feedback signal, how effective the strategy was, and how they rated the attentional demands (i.e., rate on a scale from 1 to 5 if they were focused or absent-minded during the training runs).

Stimuli

The stimuli consisted of 2 sets of photographs that were taken from the International Affective Picture Set (IAPS) (Lang et al. 1993), the Nencki Affective Picture System (NAPS) (Marchewka et al. 2013), and the Geneva Affective Picture Database (GAPED) (Dan-Glauser and Scherer 2011). Images in the first set (684 photographs) depicted social situations with a moderately positive content, among which 504 images were used for neurofeedback training (mean and standard deviation for normative valence 6.73 ± 0.92, arousal 4.48 ± 1.00), and 180 images were used for the pre- and pos ttraining tests (mean and standard deviation for normative valence 6.40 ± 0.93, arousal 4.22 ± 0.92). From the test sets, we randomly selected 60 images for the preand pos ttraining transfer runs (mean and standard deviation for normative valence 6.43 ± 0.95, arousal 4.20 ± 0.92), of which 30 images were randomly selected for the pretraining and 30 for the posttraining transfer runs. The remaining 120 images from the test set were used for the pre- and posttraining behavioural ratings (mean and standard deviation for normative valence 6.38 ± 0.92, arousal 4.23 ± 0.92), of which 60 images were randomly selected for the pretraining and 60 images for the posttraining behavioral ratings. There were no valence and arousal differences between pre- and posttraining image sets ($P > 0.20$). Note that the images in the test setswere of significantly lower valence and arousal levels than the images used during the neurofeedback training (2-tailed 2-sample t-test comparing 180 images used for tests and 504 images used for training; valence: $t_{(682)} = 4.16$, $P < 0.01$, arousal: $t_{(682)} = 2.97$, $P < 0.01$). We used slightly less emotional images during testing to avoid ceiling effects and to thereby allow detection of training-related changes. We also wanted to test whether learned self-regulation transfers not only to situations without neurofeedback, but also to situations with different valence levels. Due to these differences, the logarithmic Bayes factors for the training and

transfer runs (for details see “Connectivity-based neurofeedback signal” section) are not directly comparable (Fig. 2A,C). The order of presentationwas pseudorandomized, and no imagewas presented more than once to any participant. Images in the second set (696 photographs) depicted non-social neutral objects that were used for neurofeedback training (630 images, mean and standard deviation for normative valence 5.34 ± 0.72, arousal 3.67 ± 0.97) and for the pre- and posttraining transfer runs (66 images, mean and standard deviation for normative valence 5.41 ± 0.52, arousal 3.65 ± 0.84). From the 66 images for the transfer runs, we randomly selected 33 for the pretraining and 33 for the posttraining transfer run. There were no valence and arousal differences between training and transfer run sets, and between the images used for pre- and posttraining transfer runs (P> 0.40). Neurofeedback Training Participants took part in 3 neurofeedback training sessions spread over 3 days. On average, neurofeedback training was provided 4.7 ± 0.8 days apart (the duration between sessions did not correlate with learning success [i.e., the slope of the learning curve]: Pearson’s $\rho = -0.05$, $P = 0.85$). Every training session started with a structural scan to coregister the current head position with the anatomical template containing the ROIs. This ensured that the same ROIs were targeted across the 3 different training sessions (for details about the selection of the ROIs, see Supplementary Materials and Methods). In each training session, participants performed 2 training runs of 17.5 min each. Each of the 2 training runs consisted of 7 neurofeedback trials. A neurofeedback trial was composed of four 12 s regulation blocks that were interleaved with 5 baseline blocks of the same duration (Fig. 1B). During the baseline blocks, images of neutral objects were presented and participants were asked to passively look at them. During the regulation blocks, moderately positive social images were presented, and we suggested potential regulation strategies such as, for example, imagining experiencing the depicted positive social situation. However, it was emphasized that participants should find individual strategies that work best for them. Note that the feedback signal did not depend on a comparison between baseline and regulation conditions (as in most previous neurofeedback studies), but that it was based on a comparison between how well our 2 network model alternatives described the fMRI data acquired during each trial, which included both baseline and regulation conditions (for details, see “Connectivity-based neurofeedback signal” section). Per block, 3 images with a diameter of 12° visual angle were presented centrally for 4 s each using the Psychtoolbox-3 (Brainard 1997). After each repetition of the 5 baseline and the 4 regulation blocks, participants were given the chance to rest for 38 s, while a black screen was presented. After this rest period, participants were presented with feedback about their success for 4 s. The feedback display consisted of the logarithmic Bayes factor value (for details, see “Connectivity based neurofeedback signal” section), which was red if the trial was successful (i.e., a positive Bayes factor) and blue otherwise (i.e., a negative Bayes factor), and the total reward that had been earned until the present trial. We used intermittent rather than continuous feedback because 1) it allows for improved feedback signal quality due to more scans being available for averaging, 2) the intrinsic hemodynamic delay does not have to be taken into account by the participant, and 3) it is easier for the participants to focus on self-regulation (there is no dual-task interference with the simultaneous evaluation of the feedback signal). At least for fMRI, these advantages seem to be more important for promoting efficient neurofeedback learning than is tight temporal contiguity (Johnson et al. 2012; Sulzer et al. 2013). To promote learning, we applied a shaping procedure according to which the threshold for reward was gradually increased across training days: On the 1st day, all positive log Bayes factors were rewarded, on the 2nd day, Bayes factors needed to be larger than 2 to qualify for reward, and on the 3rd day, reward was given for Bayes factors larger than 3 (Skinner 1953).

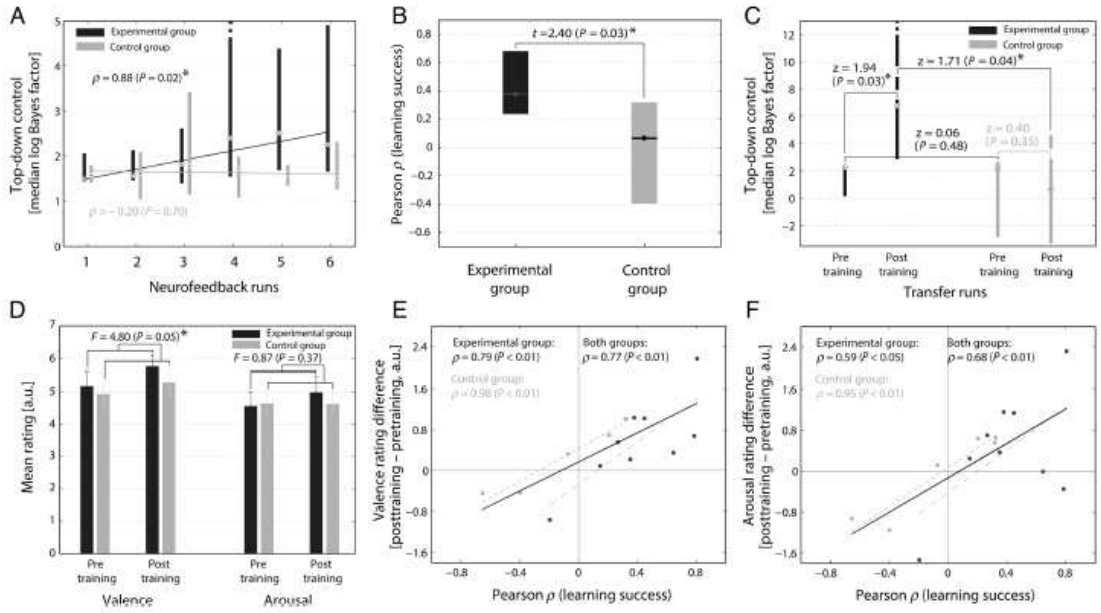

Figure 2. Neurofeedback learning and behavioral effects. (*A*) Top-down control of the dmPFC over the amygdala was measured as median log Bayes factor, which indicated the degree of dominance of the top-down model over the bottom-up model. Participants in the experimental group ($n = 9$) showed an increase in top-down control with training. Participants in the control group ($n = 6$) did not learn to control the emotion network. (*B*) Participants in the experimental group showed significantly larger training success than those in the control group, that is, the slopes of the learning curves were significantly steeper than those of the control group. (*C*) Learned control over the emotion networks when feedback was no longer available. Note that the median values of the transfer and the training runs are not directly comparable and are thus scaled differently. (*D*) The experimental group, but not the control group, showed significantly more positive responses to the target stimuli after than before training. (*E* and *F*) Significant positive correlation between the degree to which an individual learned control over the emotion network (i.e., the slope of the learning curve), and post- versus pretraining differences in valence and arousal ratings across all participants. In (*A–C*), symbols represent the median and error bars represent the lower and upper quartile; in (*D*) error bars represent the standard error of the mean. Asterisks denote statistical significance.

MRI Data Acquisition

MRI data were acquired on a 3T MRI scanner (Trio Tim, Siemens Medical Solutions, Erlangen, Germany) equipped with a 32-channel head receive coil at the Brain and Behavior Laboratory (University of Geneva). At the beginning of each scanning session, for each participant we acquired a T_1-weighted structural image (3D MPRAGE, voxel size = 1mm^3 isotropic, flip angle $\alpha = 9°$, TR = 1900 ms, TI = 900 ms, TE = 2.27 ms), and a double-echo FLASH fieldmap (TE1 = 5.19 ms, TE2 = 7.65 ms, voxel size = 3 × 3 × 2.2 mm^3). Functional images were acquired with a single-shot gradient-echo T_2*-weighted EPI sequence, with 1050 and 252 repetitions for training and transfer sessions, respectively (TR = 1100 ms, TE = 30 ms, 18 slices, matrix size = 120 × 120, voxel size = 1.8 × 1.8 × 1.8 mm^3, flip angle $\alpha = 70°$, bandwidth = 1.54 kHz/pixel, TE = 30 ms, GRAPPA, iPAT = 3). The EPI protocol had a high resolution to allow for a precise subdivision of the preselected frontal and limbic brain areas, and had a short TRto limit the effects of slice timing differences on the DCM (Kiebel et al. 2007; Koush et al. 2013). Positive phase-encoding polarity and slice tilt

approximately −42° in combination with relatively high spatial resolution at 3T was applied to optimize sensitivity for frontal and limbic brain areas (Weiskopf et al. 2006, 2007).
Visual stimuli were displayed using a rectangular projection screen at the rear of the scanner bore and were viewed with a mirror positioned on top of the head coil. All participants were instructed to breathe steadily and to remain as still as possible.
Heart rate and respiration were continuously monitored throughout the experiment and did not show any differences between the experimental conditions (for details about the cardio-respiratory variables, see Supplementary Materials and Methods).

Connectivity-Based Neurofeedback Signal

The connectivity-based neurofeedback signal was calculated using our recently developed real-time DCM approach (Koushet al. 2013). DCM is a Bayesian framework for modeling a functional brain network as a set of differential equations that describe the architecture of the network (i.e., the ROIs and the connections) and dynamic changes within the network due to external inputs (e.g., the presentation of images) and due to contextual modulations (e.g., imagining experiencing the depicted scenes) (Friston et al. 2003). Using Bayesian model comparison, DCM allows to test which model architecture explains the fMRI data best (Penny et al. 2004). The specific models that we compared using real-time DCM were a target model that represented top-down modulation from the dmPFC onto the bilateral amygdala, and a model that represented bottom-up flow of information from the bilateral amygdala onto the dmPFC (Fig. 1C; for details about the selection of the ROIs, see Supplementary Materials and Methods). For the top-down model, external inputs entered the network via the dmPFC, and modulatory inputs affected the top-down connections from the dmPFC onto the amygdala. For the bottom-up model, external inputs entered the network via the dmPFC, and modulatory inputs affected the top-down connections from the dmPFC onto the amygdala. For the bottom-up model, external inputs entered the network via the amygdala, and modulatory inputs affected the bottom up connections from the amygdala onto the dmPFC. To compare these model alternatives in real time, the fMRI images were exported to a high-end PC immediately after the acquisition (CPU Intel Core i7-3930 K 3.2 GHz, 32 GB RAM). On this computer, custom-made, real-time fMRI software running on Matlab (Mathworks Inc., Natick,MA, USA)was used to performonline motion correction; extraction of the time courses from the ROIs; removal of signal drift, spikes, and high frequency noise; and calculation of the feedback signal (Koush et al. 2012, 2013). The feedback signal was the result of the Bayesian model comparison between the 2 model alternatives using our real-time adaptation of DCM10 (as implemented in SPM8; Wellcome Trust Centre for Neuroimaging, Queen Square, London, UK, http://www.fil.ion.ucl. ac.uk). If the top-down model fitted the data of a neurofeedback trial better than the bottom-up model, the logarithmic Bayes factorwas positive, and the participantwas rewarded for a successful trial. If the bottom-up model dominated, the logarithmic Bayes factor was negative, and the participant was not rewarded (Fig. 1B). Note that the feedback signal calculation for a neurofeedback trial was based on the entire ROI time series of this trial, including baseline and regulation conditions. The feedback signal is thus not a comparison between baseline and regulation conditions (as in most previous neurofeedback studies), but is a comparison between how 2 model alternatives describe the data acquired during a trial, which includes both baseline and regulation conditions.

Pre- and Posttraining Tests

To test whether learned self-regulation transfers to situations where neurofeedback is no longer available, participants performed self-regulation in the absence of neurofeedback before and after neurofeedback training. The transfer trials were identical to the training trials except that they were composed of 11 baseline blocks interleaved with 10 regulation blocks (4.2 min run duration), that the presented images were of lower valence and arousal levels (for details, see “Stimuli” section), and that no neurofeedback was presented. During the pretraining transfer run, participants were asked to passively look at the images depicting neutral objects (baseline condition) or to imagine one-self engaged in the depicted positive social situation (regulation condition). Using the same design in 7 pilot participants (different from those recruited in the main experiment), we found increased top-down modulations from the dmPFC onto the amygdala when they imagined themselves engaged in the depicted positive social situation compared with when they passively viewed the depicted images (for details, see Supplementary Materials and Methods). During the posttraining transfer run, participants were asked to apply their newly learned self-regulation skills. To test whether neurofeedback training modulated the participants’ affective responses to visual stimuli, we asked them to rate their subjective response to visual stimuli similar to those used during the pre- and posttraining transfer runs in terms of valence and arousal using the standard self-assessment manikin ratings (SAM) (Lang et al. 1993). The pictures used for these ratings were less positive than the pictures used during training (for details, see “Stimuli” section). These ratings were performed once before and once after the neurofeedback training, using the SAM 9-point rating scale (Lang et al. 1993).

Psychometric Questionnaires

Before the experiment, we asked participants to complete the Emotion Regulation Questionnaire (ERQ) (Gross and John 2003), the Thought Control Ability Questionnaire (TCAQ) (Luciano et al. 2005), the White Bear Suppression Inventory (WBSI) (Wegner and Zanakos 1994), the State-Trait Anxiety Inventory (STAI) (Spielberger et al. 1983), the Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPRSQ) (Torrubia et al. 2001; Lardi et al. 2008), and the Beck Depression Inventory (BDI) (Beck et al. 1961). None of the participants suffered from depression (BDI scores were ~1.5 ± 1.1). Average questionnaire scores did not differ between participants in the experimental group and those of the control group (2-tailed 2-sample t-tests, $P > 0.20$ uncorrected).

Statistical Analyses

As dependent variables, we included valence and arousal rating differences, and learning success. The differences between the pre- and post training valence and arousal ratings were calculated for each participant and converted to z-scores. The learning success was indicated by the slope of the learning curve, i.e., the linear regression of the logarithmic Bayes factor across training sessions (as indicated by Pearson rho). To study the covariance patterns between dependent variables and to further examine whether valence and arousal rating differences were similarly modulated by the learning success ,we performed a multivariate analysis of variance (MANOVA) and a principal component analysis (PCA, for details, see Supplementary Materials and Methods). To further elucidate the effect of neurofeedback training on subjectively experienced levels of valence and arousal, we performed 2 separate analyses of variance (ANOVAs) with the factor group (experimental vs. control) and the covariate learning success We estimated post hoc the difference between the learning success of the experimental and control groups using a 2-tailed 2-sample t-test. Next, we analyzed the difference in the resulting logarithmic Bayes factors between the pre- and posttraining transfer runs, as well as between the participants in the experimental group and those in the control group using 1-tailed Wilcoxon rank sum tests and z-statistics. This approach was used, because a Jarque–Bera test established that the logarithmic Bayes factors of these runs were not normally distributed. To ensure that pre- and posttraining transfer run performance could be compared, we constructed participant-specific ROIs that were applied to both of these runs. These ROIs were based on the disjunction of the ROIs that were used in the best training run of each participant. The transfer run time courses of the so-defined ROIs were extracted, corrected for possible signal drifts, high frequency noise, and spikes, and were fed into the DCM analysis.
Subsequently, pre- and posttraining valence and arousal ratings were converted to z-scores and compared post hoc using 1-tailed paired t-tests (separately for the experimental and for the control group). We also calculated post hoc Pearson correlations between each participant’s pre- and posttraining valence and arousal rating differences (i.e., posttraining–pretraining) and learning success (i.e., the slope of their learning curve).
To investigate the DCM model parameter differences between the pre- and posttraining transfer runs, we applied 2-tailed paired t-tests to the corresponding sets of DCM model parameters from the pre- and posttraining transfer run models (separately for the experimental and the control group). To assess how the dynamic ROIs (based on the incremental general linear model [iGLM]; for details, see Supplementary Materials and Methods) changed across neurofeedback training runs, we analyzed 1) the mean signal change, 2) the mean number of voxels within each ROI, and 3) the cytoarchitectonic composition of the amygdala for each training run (for details, see Supplementary Materials and Methods).

We calculated Pearson ρ correlations between the participant's normalized z-scores on the questionnaires and 1) the slope of the learning curve, and 2) the slope of the ROI activity changes across training runs. We also calculated Pearson ρ correlations between the participants' normalized z-scores on the questionnaires and 1) the logarithmic Bayes factor differences, 2) the DCM parameter differences, and 3) the ROI activity differences between the pre- and the posttraining transfer runs. The correlation with the ROIs was calculated separately for each of the 3 ROIs. The correlation with the DCM parameter differences was only reported for significant differences (i.e., between the top-down and the bottom-up connection between the dmPFC and the right amygdala). These correlation analyses were done separately for the participants in the experimental group and for those in the control group. The statistical significance was corrected for multiple comparisons using Bonferroni familywise error rate correction (FWE).
Finally, we performed a random-effect whole-brain group-level analysis using a paired t-test comparing posttraining > pretraining transfer runs (for details, see Supplementary Materials and Methods).

Results

The MANOVA, which assesses the effect of neurofeedback training conjointly on learning success (i.e., the slope of the learning curve) and the differences in valence and arousal ratings, revealed a significant effect of training (factor group: $F_{3,11} = 6.93$, $P < 0.01$; for details, see Supplementary Materials and Methods). The ANOVA on valence rating differences (global fit: $F_{2,12} = 14.47$, $P < 0.01$) confirmed a main effect of group (factor group: $F_{1,12} = 4.80$, $P = 0.05$), with higher positive valence rating differences in the experimental than in the control group. The ANOVA on arousal rating differences (global fit: $F_{2,12} = 6.03$, $P = 0.02$) indicated no effect of training (factor group: $F_{1,12} = 0.86$, $P = 0.37$). Interestingly, both valence and arousal rating differences were positively influenced by learning success (learning success covariates: $F_{1,12} = 28.12$, $P < 0.01$ and $F_{1,12} = 10.89$, $P < 0.01$, respectively). The PCA further illustrates significant differences between the participants in the experimental group and in the control group in terms of learning success and concomitant changes in behavioural ratings (see Supplementary Fig. 1; for details, see Supplementary Materials and Methods).

Learning to Increase Top-Down Control

Follow-up tests revealed that over the course of training, participants in the experimental group successfully learned to increase the dominance of the top-down model compared with the bottom-up model (Fig. 2A; experimental group: Pearson's $\rho = 0.88$, $P = 0.02$). Participants in the control group did not learn to increase top-down control (Fig. 2A; control participants: Pearson's $\rho = -0.20$, $P = 0.70$). The participants in the experimental group showed significantly higher learning success than those in the control group, i.e., the slopes of the learning curves were significantly steeper (Fig. 2B; 2-tailed 2-sample t-test, $t_{(13)} = 2.40$, $P = 0.03$). The learned ability to control top-down connectivity was subsequently maintained in the absence of neurofeedback. In posttraining transfer runs without feedback, we found that participants in the experimental group could control top-down connectivity significantly better after than before neurofeedback training (Fig. 2C; 1-tailedWilcoxon rank sumtests and z-statistics; experimental group, $z = 1.94$, $P = 0.03$). No such improvement was found in the control participants who received sham neurofeedback (1-tailed Wilcoxon rank sum tests and z-statistics; control group, $z = 0.40$, $P = 0.35$). In the posttraining transfer runs, dominance of the top-down model was significantly higher for participants in the experimental group compared with the control participants (Fig. 2C; 1-tailed Wilcoxon rank sum tests and z-statistics; posttraining transfer run comparison: $z = 1.71$, $P = 0.04$), whereas pretraining transfer runs did not reveal such a difference (1-tailed Wilcoxon rank sum tests and z-statistics; pretraining transfer run comparison: $z = 0.06$, $P = 0.48$). These data show that once participants in the experimental group had learned control over top-down connectivity, this learned skill could be employed even in the absence of neurofeedback.

Behavioral Effects of Self-Regulation

Follow-up tests revealed that the difference between the experimental and the control group was driven by a significant increase in valence ratings after compared with before neurofeedback training that was related to learned control over the top-down connectivity.
It was evident only for the participants in the experimental group (Fig. 2D; experimental group; 1-tailed paired t-test, $t_{(8)} = 2.04$, $P = 0.04$), but not in the control participants (Fig. 2D; control participants; 1-tailed paired t-test, $t_{(5)} = 1.33$, $P = 0.12$). The effect sizes for valence ratings were large and medium, respectively (experimental group Cohen's $d = 0.53$; control group Cohen's $d = 0.42$). When assuming equal group sizes between the experimental and the control group, the increase in valence ratings that reached almost trend level in the control group does not become significant, and is thus not an effect of power (t_9 ¼ ð ffiffiffiffi 9 p = ffiffiffiffi 6 p Þ × t_6; $t_9 = 1.63$, $P_9 = 0.07$). Follow-up tests also indicated that arousal ratings did not change with training (Fig. 2D; posttraining–pretraining; experimental group: 1-tailed paired t-test, $t_{(8)}=1.12$, $P = 0.15$; control group: 1-tailed paired t-test, $t_{(5)}=0.05$, $P = 0.52$), nor did we find any other significant intra- and inter-group valence and arousal rating ($P > 0.15$).
Over and above group-specific effects, collapsing across both groups, there was a significant positive correlation between the degree to which an individual improved top-down control across training runs (i.e., the slope of the learning curve) and the increase in valence and arousal ratings that they exhibited (Fig. 2E,F; Pearson correlation; $\rho = 0.77$, $P < 0.01$ and $\rho = 0.68$, $P < 0.01$, respectively). This finding accords with the notion that regulation abilities mediated by the dmPFC–amygdala network were directly related to the subjective emotion appraisals of the participants, but that they were differentially boosted in those participants undergoing neurofeedback training.
Unlike for emotion ratings, there were no significant differences in psychological questionnaire scores assessing mood and anxiety after compared with before neurofeedback training (2-tailed paired t-tests, all $P > 0.20$ uncorrected; Emotion Regulation Questionnaire [ERQ], Thought Control Ability Questionnaire [TCAQ], White Bear Suppression Inventory [WBSI], State-Trait Anxiety Inventory [STAI], Sensitivity to Punishment and Sensitivity to Reward Questionnaire [SPRSQ], Beck Depression Inventory [BDI]).

Mental Processes Underlying Self-Regulation

How did our participants learn to increase top-down connectivity of the dmPFC over the amygdala? In common with established practice in the neurofeedback field, we were not prescriptive about the strategy that participants should adopt during learning (Sulzer et al. 2013). In debriefing, most reported attempting a strategy related to imagining themselves as personally involved and experiencing the depicted positive social situations (see Supplementary Table 1). However, participants in the control group used similar strategies (although with somewhat greater variability, see Supplementary Table 1), but they nevertheless failed to learn self-regulation. This indicates that in addition to feedback- guided search for an explicit control strategy, other, more implicit learning mechanisms attributable to operant conditioning based on reinforcement by the feedback might be a factor (Thorndike 1898; Skinner 1953; Bray et al. 2007; Birbaumer et al. 2013). Furthermore, attentional factors alone cannot explain our findings, because the participants in the experimental group and the control participants showed statistically indistinguishable attentional effort (2-tailed 2-sample t-test, $t_{(13)} = 0.32$, $P = 0.76$; rating experimental group: 4.33 ± 0.87, rating control group: 4.17 ± 1.17), yet the training success diverged substantially (Fig. 2B). Likewise, motivational factors alone cannot explain our findings, because feedback reward levels in both experimental groups were identical, and participants in the control groupwere unaware that they had received sham feedback.

Neural Substrates of the Learning Effect

To further explore the neural substrates of the learning effect that we observed, we compared the individual model parameters of the pretraining and the posttraining transfer runs. This comparison revealed a significant increase in top-down modulation and a significant decrease in bottom-up processing between the dmPFC and the right amygdala that were found only in the participants in the experimental group, and only in the trained model (i.e., the top-down model; Fig. 3; Pearson correlation; dmPFC→ AMY_R connectivity: 2-tailed paired t-test, $t_{(8)} = 2.38$, $P = 0.05$ uncorrected; AMY_R→dmPFC connectivity: 2-tailed paired t-test, $t_{(8)}= 2.49$, $P= 0.04$ uncorrected).

An analysis of brain activations in the 3 ROIs that comprised our emotion network models revealed that dmPFC activity significantly increased, and that right amygdala activity significantly decreased in the participants in the experimental group, with training (see Supplementary Fig. 2A; Pearson correlation; dmPFC: $\rho = 0.81$, $P = 0.05$ uncorrected; AMY_R: $\rho = -0.81$, $P = 0.05$ uncorrected).

Participants in the experimental group with higher levels of state anxiety (STAI-S score) showed a smaller increase in dmPFC activity, which is in line with previous findings that showed decreased levels of prefrontal activity in highly anxious individuals during negative emotion processing (Fig. 4, see Supplementary Table 2; Pearson correlation; $\rho = -0.94$, $P < 0.01$ corrected for 8 STAI correlations; for further results related to the psychometric questionnaire scores, see Supplementary Fig. 3 and Material and Methods) (Bishop et al. 2004). In contrast, participants in the control group showed significantly increased left amygdala activity with training (see Supplementary Fig. 2B; Pearson correlation; $\rho = 0.86$, $P = 0.03$ uncorrected). Interestingly, for the experimental group, the increased percent signal change in the dmPFC was associated with a decrease of its number of active voxels (see Supplementary Fig. 4A), whereas for all other areas in both groups, the similar trend for percent signal change and the number of active voxels was observed (see Supplementary Figs 2 and 4). The increased left amygdala activity in the control participants is also reflected in an increasing number of voxels of particularly the laterobasal nuclei that the control participants recruited with training (see Supplementary Fig. 5; Pearson correlation; $\rho = 0.90$, $P = 0.02$ uncorrected; our ROIs were defined dynamically to allow for shaping of brain activity (Skinner 1953); for details, see Supplementary Fig. 6 and Materials and Methods). These differences between the experimental group and the control group were also evident when directly contrasting whole brain activation of the pre- and posttraining transfer runs, which revealed a significant increase indmPFC activity in the experimental group and a significant increase in left amygdala activity in the control group (see Supplementary Fig. 7).

Discussion

Our results establish for the first time that participants can employ connectivity-based neurofeedback to learn to increase topdown connectivity from the dmPFC onto the amygdala in a self-organized, endogenous fashion (Fig. 2A). Control participants who received sham feedback did not learn such control, indicating that the learning effects observed in the experimental group cannot be achieved with unrelated feedback (Fig. 2A,B). The sample size in our study was rather small and, as a consequence, the behavioral effects were somewhat statistically weak. Nonetheless, the fact that increasing the level of top-down connectivity was associated with increased valence ratings (Fig. 2D,E) might indicate the potential for enhancing emotion regulation through connectivity-based neurofeedback training.

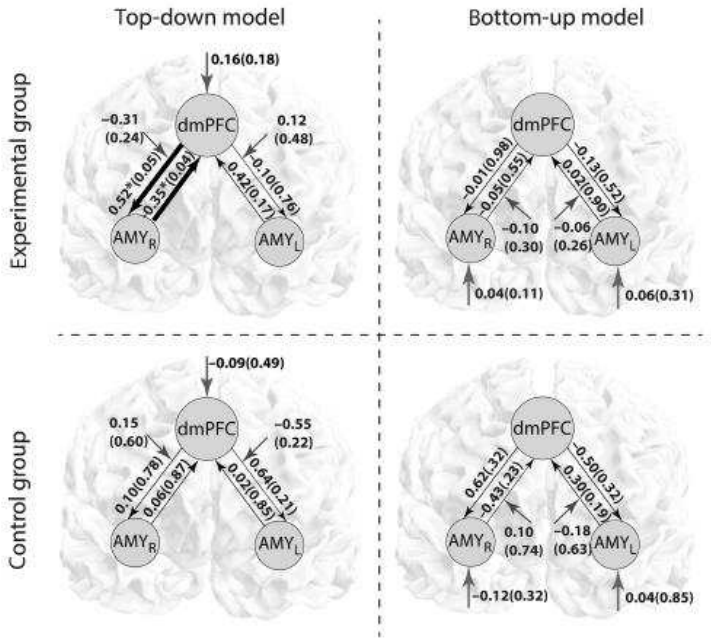

Figure 3. Model parameter differences. (A) Comparing pre- and posttraining transfer runs, the connectivity between the dmPFC and the right amygdala changed significantly in the experimental group. Specifically, top-down connections from the dmPFC onto the amygdala were increased, and bottom-up connections from the right amygdala onto the dmPFC were decreased. These changes were specific to the trained top-down model (i.e., they were not found in the bottom-up model), and (B) specific to the participants in the experimental group (i.e., they were not found in the control participants). Asterisks and bold lines denote statistical significance; *P*-values are indicated in brackets.

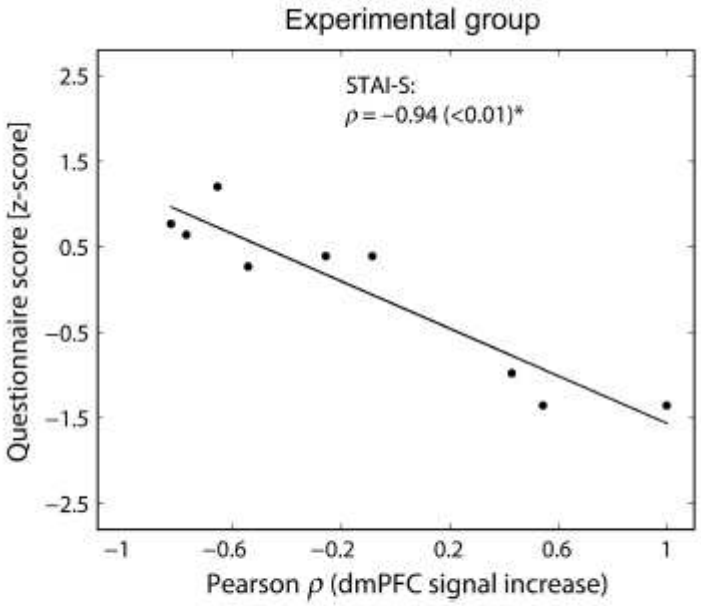

Figure 4. State anxiety predicts neurofeedback learning success. STAI-S score was correlated negatively with the degree to which dmPFC activity increased across neurofeedback training runs. Asterisk indicates that the result survived Bonferroni FWE correction for multiple comparisons; P-value is indicated in brackets.

Our new connectivity-based neurofeedback approach goes beyond previous neurofeedback approaches that were limited to training localized brain activity within a ROI (Sulzer et al. 2013). Although connectivity-based neurofeedback training can result in accompanying changes in the level of activity in ROIs (as was found in the present study, see Supplementary Figs 2 and 4, for details, see Supplementary Materials and Methods), we have previously shown that it is a new and distinct feedback measure that reflects connectivity between areas, and that is qualitatively different from activity-based feedback (i.e., activity- based feedback signals and connectivity-based feedback signals can be uncorrelated) (Koush et al. 2013). Connectivity-based neurofeedback thus provides a novel way to noninvasively and nonpharmacologically change interactions within interconnected functional brain networks, to provide brain-based training for causing specific behavioral changes.

Learning to increase top-down control of the dmPFC over the amygdala was associated with decreased bottom-up processing between the dmPFC and the right amygdala (Fig. 3A). It was also associated with increased activity in the dmPFC and decreased activity in the right amygdala (see Supplementary Fig. 2A), which is particularly involved in inducing negative emotions (Lanteaume et al. 2007). Such training might thus be used to directly target emotion regulation disorders such as depression and anxiety, which are often characterized by insufficient topdown inhibition, excess bottom-up processing, hyperactivity in the amygdala, and hypoactivity in the dmPFC (Disner et al. 2011). It is particularly important for clinical applications that after training, self-regulation of effective connectivity between the dmPFC and the amygdala can be applied even without neurofeedback, as demonstrated by our transfer runs (Fig. 2C) (Sulzer et al. 2013).

Our connectivity-based neurofeedback approach is not limited to enhancing positive emotions through training of topdown emotion regulation connectivity, which may help counteract the frequent anhedonia component of emotion regulation disorders. It might also be applied to improve emotion regulation capabilities by decreasing negative emotions associated with these disorders. In principle, any functional brain network might be targeted with this new method to study causal relationships with noninvasive neuroimaging methods, or to normalize dysfunctional brain networks in psychiatric and neurological disorders (Stoeckel et al. 2014). In the future, neurofeedback training of functional brain networks may provide a powerful and highly specific mean to promote plasticity and learning in various conditions, by modulating patterns of interactions between brain areas—rather than the level of activity within one brain area only, as traditionally implemented by most neuromodulation techniques such as single-ROI neurofeedback, transcranial magnetic stimulation (TMS), or deep brain stimulation (DBS). This in turn opens new perspectives for treatment of neuropsychiatric disorders associated with emotion regulation failure, including depression and anxiety, possibly in combination with other procedures based on psychotherapy and medication.

Authors' Contributions

Y.K., S.P., P.V., F.S. conceived and designed the experiments. Y.K., G.R. performed the experiments. Y.K., D.E.M., S.W.R., D.E.J.L., D.V. D.V., F.S. analyzed the data/contributed analysis tools. Y.K., F.S. wrote the paper and all authors critically reviewed and approved the final manuscript.

Funding

This studywas supported by the Center for Neuroscience and the National Center of Competence in Research (NCCR) for Affective Sciences financed by the Swiss National Science Foundation (no 51NF40-104897) and hosted by the University of Geneva, the Geneva Academic Society (Foremane Fund), the Ernest Boninchi Foundation, the SwissNational Science Foundation (PZ00P3_131932, BSSGI0_155915), and the European Union.

Notes

We thank Narly Golestani for helpful discussions. Conflict of Interest: None declared.

References

Banks SJ, Eddy KT, Angstadt M, Nathan PJ, Phan KL. 2007. Amygdala- frontal connectivity during emotion regulation. Soc Cogn Affect Neurosci. 2:303–312.

Barbas H, Zikopoulos B. 2007. The prefrontal cortex and flexible behavior. Neuroscientist. 13:532–545.

Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. 1961. An inventory for measuring depression. Arch Gen Psychiatry. 4:561–571.

Birbaumer N, Ruiz S, Sitaram R. 2013. Learned regulation of brain metabolism. Trends Cogn Sci. 17:295–302.

Bishop S, Duncan J, Brett M, Lawrence AD. 2004. Prefrontal cortical function and anxiety: controlling attention to threat-related stimuli. Nat Neurosci. 7:184–188. Brainard DH. 1997. The Psychophysics Toolbox. Spat Vis. 10:433–436.

Bray S, Shimojo S, O'Doherty JP. 2007. Direct instrumental conditioning of neural activity using functional magnetic resonance imaging-derived reward feedback. J Neurosci. 27:7498–7507.

Dan-Glauser ES, Scherer KR. 2011. The Geneva affective picture database (GAPED): a new730-picture database focusing on valence and normative significance. Behav Res Methods. 43:468–477.

Disner SG, Beevers CG, Haigh EAP, Beck AT. 2011. Neural mechanisms of the cognitive model of depression. Nat Rev Neurosci. 12:467–477.

Friston KJ, Harrison L, PennyW. 2003. Dynamic causal modelling. NeuroImage. 19:1273–1302.

Gross JJ, John OP. 2003. Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. J Pers Soc Psychol. 85:348–362.

Haller S, Kopel R, Jhooti P, Haas T, Scharnowski F, Lovblad KO, Scheffler K, Van De Ville D. 2013. Dynamic reconfiguration of human brain functional networks through neurofeedback. Neuroimage. 81:243–252.

Johnson KA, Hartwell K, LeMatty T, Borckardt J, Morgan PS, Govindarajan K, Brady K, George MS. 2012. Intermittent "realtime" fMRI feedback is superior to continuous presentation for a motor imagery task: a pilot study. J Neuroimaging. 22:58–66.

Kiebel SJ, Kloppel S,Weiskopf N, Friston KJ. 2007. Dynamic causal modeling: a generative model of slice timing in fMRI. NeuroImage. 34:1487 1496.

Kim DY, Yoo SS, Tegethoff M, Meinlschmidt G, Lee JH. 2015. The inclusion of functional connectivity information into fMRIbased neurofeedback improves its efficacy in the reduction of cigarette cravings. J Cogn Neurosci. 27:1552–1572.

Kim MJ, Loucks RA, Palmer AL, Brown AC, Solomon KM, Marchante AN, Whalen PJ. 2011. The structural and functional connectivity of the amygdala: from normal emotion to pathological anxiety. Behav Brain Res. 223:403–410.

Koush Y, Rosa MJ, Robineau F, Heinen K, Rieger SW, Weiskopf N, Vuilleumier P, Van De Ville D, Scharnowski F. 2013. Connectivity- based neurofeedback: dynamic causal modeling for realtime fMRI. Neuroimage. 81:422–430.

Koush Y, Zvyagintsev M, Dyck M, Mathiak KA, Mathiak K. 2012. Signal quality and Bayesian signal processing in neurofeedback based on real-time fMRI. NeuroImage. 59:478–489.

Lang PJ, Greenwald MK, Bradley MM, Hamm AO. 1993. Looking at pictures - affective, facial, visceral, and behavioral reactions. Psychophysiology. 30:261–273.

Lanteaume L, Khalfa S, Regis J, Marquis P, Chauvel P, Bartolomei F. 2007. Emotion induction after direct intracerebral stimulations of human amygdala. Cereb Cortex. 17:1307–1313.

Lardi C, Billieux J, d'Acremont M, Van der Linden M. 2008. A French adaptation of a short version of the Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ). Pers Individ Dif. 45:722–725.

Lee JH, Kim J, Yoo SS. 2012. Real-time fMRI-based neurofeedback reinforces causality of attention networks. Neurosci Res. 72:347–354.

Lee S, Ruiz S, Caria A, Veit R, Birbaumer N, Sitaram R. 2011. Detection of cerebral reorganization induced by real-time fMRI feedback training of insula activation: a multivariate investigation. Neurorehabil Neural Repair. 25:259–267. Linden DEJ, Habes I, Johnston SJ, Linden S, Tatineni R, Subramanian L, Sorger B, Healy D, Goebel R. 2012. Real-time self-regulation of emotion networks in patients with depression. PLoS One 7:e38115.

Luciano JV, Algarabel S, Tomas JM, Martinez JL. 2005. Development and validation of the thought control ability questionnaire. Pers Individ Dif. 38:997–1008.

Marchewka A, Zurawski L, Jednorog K, Grabowska A. 2013. The Nencki Affective Picture System (NAPS): introduction to a novel, standardized, wide-range, high-quality, realistic picture database. Behav Res Methods. 46:596–610. Ochsner K. 2010. The cognitive control of emotion: from basic mechanisms to mechanistic breakdowns. Biol Psychiatry. 67:1S. Ochsner KN, Gross JJ. 2005. The cognitive control of emotion. Trends Cogn Sci. 9:242–249.

Ochsner KN, Silvers JA, Buhle JT. 2012. Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. Ann N Y Acad Sci. 1251: E1–E24.

Penny WD, Stephan KE, Mechelli A, Friston KJ. 2004. Comparing dynamic causal models. Neuroimage. 22:1157–1172.

Pessoa L. 2008. On the relationship between emotion and cognition. Nat Rev Neurosci. 9:148–158.

Rota G, Handjaras G, Sitaram R, Birbaumer N, Dogil G. 2011. Reorganization of functional and effective connectivity during real-time fMRI-BCI modulation of prosody processing. Brain Lang. 117:123–132.

Sarkheil P, Zilverstand A, Kilian-Hutten N, Schneider F, Goebel R, Mathiak K. 2015. fMRI feedback enhances emotion regulation as evidenced by a reduced amygdala response. Behav Brain Res. 281:326–332.

Scharnowski F, Rosa MJ, Golestani N, Hutton C, Josephs O, WeiskopfN,Rees G. 2014.Connectivity changes underlying neurofeedback training of visual cortex activity. PLoS One. 9:e91090. Shen J, Zhang G, Yao L, Zhao X. 2015. Real-time fMRI training-induced changes in regional connectivity mediating verbal working memory behavioral performance. Neuroscience. 289:144–152.

Shibata K, Watanabe T, Sasaki Y, Kawato M. 2011. Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. Science. 334:1413–1415.

Skinner BF. 1953. Science and human behavior. Oxford, UK: Macmillan. Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. 1983 Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press. Stephan KE, Penny WD, Moran RJ, den Ouden HEM, Daunizeau J, Friston KJ. 2010. Ten simple rules for dynamic causal modeling. Neuroimage. 49:3099–3109.

Stoeckel LE, Garrison KA, Ghosh S, Wighton P, Hanlon CA, Gilman JM, Greer S, Turk-Browne NB, deBettencourt MT, Scheinost D, et al. 2014. Optimizing real time fMRI neurofeedback for therapeutic discovery and development. Neuroimage Clin. 5:245–255.

Sulzer J, Haller S, Scharnowski F, Weiskopf N, Birbaumer N, Blefari ML, Bruehl AB, Cohen LG, Decharms RC, Gassert R, et al. 2013. Real-time fMRI neurofeedback: progress and challenges. Neuroimage. 76C:386–399.

Thorndike EL. 1898. Animal intelligence: an experimental study of the associative processes in animals. NewYork: Macmillan. Torrubia R,Avila C,Molto J, Caseras X. 2001. The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. Pers Individ Dif. 31:837–862.

Treadway MT, Zald DH. 2011. Reconsidering anhedonia in depression: lessons from translational neuroscience. Neurosci Biobehav Rev. 35:537–555.

Wegner DM, Zanakos S. 1994. Chronic thought suppression. J Pers. 62:615–640.

Weiskopf N,Hutton C, Josephs O, Deichmann R. 2006. Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: a whole-brain analysis at 3T and 1.5T. NeuroImage. 33:493–504.

Weiskopf N, Hutton C, Josephs O, Turner R, Deichmann R. 2007. Optimized EPI for fMRI studies of the orbitofrontal cortex: compensation of susceptibility-induced gradients in the readout direction. Magn Reson Mater Phys Biol Med. 20:39–49.

Young KD, Zotev V, Phillips R, Misaki M, Yuan H, Drevets WC, Bodurka J. 2014. Real-time FMRI neurofeedback training of amygdala activity in patients with major depressive disorder. PLoS One. 9:e88785.

Yuan H, Young KD, Phillips R, Zotev V, Misaki M, Bodurka J. 2014. Resting-state functional connectivity modulation and sustained changes after real-time functional magnetic resonance imaging neurofeedback training in depression. Brain Connect. 4:690–701.
Zilverstand A, Sorger B, Zimmermann J, Kaas A, Goebel R. 2014. Windowed correlation: a suitable tool for providing dynamic fMRI-based functional connectivity neurofeedback on task difficulty. PLoS One. 9:e85929.
Zotev V, Phillips R, Young KD, Drevets WC, Bodurka J. 2013. Prefrontal control of the amygdala during real-time fMRI neurofeedback training of emotion regulation. PLoS one. 8: e79184.