Self-regulation of the anterior insula: Reinforcement learning using real-time fMRI neurofeedback

Emma J. Lawrence a,b,⁎, Li Su a,c,1, Gareth J. Barker b, Nick Medford d, Jeffrey Dalton b, Steve C.R. Williams b, Niels Birbaumer e, Ralf Veit c, Sitaram Ranganatha e, Jerzy Bodurka f,g, Michael Brammer b, Vincent Giampietro b,2, Anthony S. David a,b

a Department of Psychiatry Studies, Institute of Psychiatry, King’s College London, London SE5 8AF, UK
b Department of Neuroimaging, Institute of Psychiatry, King’s College London, London SE5 8AF, UK
c Department of Psychology, University of Cambridge, Cambridge, UK
d Brighton and Sussex Medical School, Brighton, UK
e Laureate Institute for Brain Research, Tulsa, USA
f Institute of Medical Psychology and Behavioral Neurobiology, Eberhard-Karls-University, Tübingen, Germany
g College of Engineering, University of Oklahoma, Tulsa, USA

A R T I C L E   I N F O

Article history:
Accepted 29 October 2013
Available online 11 November 2013

Keywords:
Real-time fMRI
Insula
Neurofeedback
Emotion
Self-regulation
Reinforcement learning

A B S T R A C T

The anterior insula (AI) plays a key role in affective processing, and insular dysfunction has been noted in several clinical conditions. Real-time functional MRI neurofeedback (rtfMRI-NF) provides a means of helping people learn to self-regulate activation in this brain region. Using the Blood Oxygenated Level Dependant (BOLD) signal from the right AI (RAI) as neurofeedback, we trained participants to increase RAI activation. In contrast, another group of participants was shown ‘control’ feedback from another brain area. Pre- and post-training affective probes were shown, with subjective ratings and skin conductance response (SCR) measured. We also investigated a reward-related reinforcement learning model of rtfMRI-NF. In contrast to the controls, we hypothesised a positive linear increase in RAI activation in participants shown feedback from this region, alongside increases in valence ratings and SCR to affective probes. Hypothesis-driven analyses showed a significant interaction between the RAI/control neurofeedback groups and the effect of self-regulation. Whole-brain analyses revealed a significant linear increase in RAI activation across four training runs in the group who received feedback from RAI. Increased activation was also observed in the caudate body and thalamus, likely representing feedback-related learning. No positive linear trend was observed in the RAI in the group receiving control feedback, suggesting that these data are not a general effect of cognitive strategy or control feedback. The control group did, however, show diffuse activation across the putamen, caudate and posterior insula which may indicate the representation of false feedback. No significant training-related behavioural differences were observed for valence ratings, or SCR. In addition, correlational analyses based on a reinforcement learning model showed that the dorsal anterior cingulate cortex underpinned learning in both groups. In summary, these data demonstrate that it is possible to regulate the RAI using rtfMRI-NF within one scanning session, and that such reward-related learning is mediated by the dorsal anterior cingulate.

© 2013 Elsevier Inc. All rights reserved.

Introduction

Developments in functional magnetic resonance imaging (fMRI) now allow very rapid data transfer and analysis within a few seconds of data collection. Such "real-time" fMRI (rtfMRI) provides a novel brain–computer interface (BCI) with extensive applications (DeCharms, 2008). RtfMRI may have important therapeutic implications because it is a non-invasive technique that allows brain activation levels to be relayed to participants using neurofeedback (NF) paradigms, with the aim of aiding the self-regulation of brain states. Using fMRI-based feedback overcomes many of the limitations of other feedback methods, such as skin conductance and electroencephalography (EEG), by directly targeting specific brain regions, with superior spatial resolution. This method has already been shown to have clinical utility; for example, chronic pain patients report reduced pain perception after feedback from the anterior cingulate cortex (DeCharms et al., 2005).

RtfMRI-NF is based on the premise that participants can learn to regulate brain activation in a target brain region when provided with feedback as to their success, and has been demonstrated in several studies (Caria et al., 2007; McCreight et al., 2011; Ruiz et al., 2011; Veit et al., 2011; Weiskopf et al., 2003; Zotev et al., 2011). However, it is not yet...
clear whether robust NF-related behavioural effects occur after brief training (Caria et al., 2010; DeCharms et al., 2005; Johnston et al., 2011; Rota et al., 2009; Ruiz et al., 2011), and this needs to be determined if the technique is to be of clinical use. In addition, there is still a lack of understanding of the underlying neurocognitive processes and neuronal specificity involved in NF-based self-regulation.

The anterior insula (AI) has been implicated in affective processing, including emotional and interoceptive awareness (Craig, 2009; Simmons et al., 2012; Zaki et al., 2012). This region has been shown to be hypoactive in clinical conditions involving attenuated affect, such as depersonalisation (Phillips et al., 2001). Grey matter volume reductions in the AI have also been found in adolescents with a diagnosis of conduct disorder (Fairchild et al., 2011) and men diagnosed with schizophrenia (Shapleske et al., 2002), which may underpin deficits in empathy (Singer et al., 2009). In addition, insular dysfunction is associated with the increased threat perception that can be observed in psychosis, with changes in insula activation observed after successful cognitive behaviour therapy (Kumari et al., 2011).

Previous studies have shown that it is possible to learn to regulate the AI using rtfMRI-NF (Caria et al., 2007; Veit et al., 2011) and that this is accompanied by increased valence ratings to aversive stimuli (Caria et al., 2010). In addition, it has been shown that rtfMRI-led regulation of the AI is accompanied by increased accuracy in recognising facial expressions of disgust in participants with a diagnosis of schizophrenia (Ruiz et al., 2011). Building on this previous work, we investigated whether participants who learn to regulate the AI within one scanning session would also show changes in subjective ratings and skin conductance response (SCR) to affective probes of normatively positive and negative valence. We also examined a new mental strategy to aid NF-based self-regulation, adapted from the mindfulness literature (Farb et al., 2007), which required participants to focus on their bodily sensations as they arise. Consistent with the mental strategies we aimed to investigate, we choose the right anterior insula (RAI) as our NF target due to its role in both interoception and emotional experience (Zaki et al., 2012).

It has been argued that the underlying neural mechanism of rtfMRI-NF may be based on the principles of operant learning i.e. participants use the feedback of their Blood Oxygenated Level Dependent (BOLD) signal as intrinsic reward to adjust their mental strategies and learn to control their own brain activity (Caria et al., 2011). However, little is known about what underlying neural systems are involved in reward-related learning in the context of NF. Previous literature suggests the ventral striatum to be associated with learning from reward and prediction error (Schonberg et al., 2007), and dorsal anterior cingulate (dACC) to be involved with reward based decision making (Bush et al., 2001; Holroyd and Yeung, 2012), so these are possible candidate regions. To examine this further, we applied a reinforcement learning model (Sutton and Barto, 1998) to predict the learning related fMRI activation pattern during NF.

We measured behavioural change using pre- and post-training affective probes i.e. images from the International Affective Picture System (IAPS: Lang et al., 2005), and recorded subjective valence and arousal ratings, as well as SCR. In addition, bearing in mind the potential clinical applications of this technique, our affective probe included normatively positive stimuli, enabling us to assess the impact of rtfMRI-NF on positive affect (Johnston et al., 2011).

We predicted that by using BOLD response from the RAI as neurofeedback signal, participants would display a positive linear increase in RAI activation indicative of a learned modulation of this region; conversely, we predicted that this would not occur in participants shown feedback from the middle parahippocampal region, and so not specific to the mental strategies used. We also hypothesised the involvement of brain regions associated with reinforcement learning during training runs i.e. ventral striatum and dorsal anterior cingulate. In terms of behavioural effects, we predicted post-training increases in autonomic arousal, i.e. SCR, and subjective valence ratings in response to both positive and negative stimuli.

Methods

Participants

Twenty-four healthy right-handed participants (12 females) were recruited. Participants were proficient in written and spoken English and had no current neurological or psychiatric diagnosis. All participants provided written informed consent and received compensation for their participation. This study was approved by the local ethics committee at King’s College London, UK (Ethics No: CREC/07/08-131). Sixteen participants were assigned to the RAI feedback group and received rtfMRI-NF from the RAI (mean age 27 ± 5 years, 8 females). Eight participants were assigned to the control ‘sham’ feedback group (mean age 28 ± 4 years, 4 females); they were given identical training instructions to the RAI NF group but received feedback from another brain region. Prior to this, 24 participants (different from those in the NF study) were recruited for an offline pilot to examine the impact of cognitive strategy without NF (mean age 31 ± 8 years, 12 females).

Experimental procedure

Prior to the experiment, participants were told that they would be shown information from their brain, with a few seconds delay, and that their task was to increase/decrease the level of the thermometer bar (used to visually display the feedback on the projector screen) using suggested strategies (see below). Participants were asked to keep their head still, eyes open and focussed on the thermometer bar, and were familiarised with the graphical user interface (GUI) beforehand in a mock scanner using SCR derived feedback. The Beck Anxiety Inventory (BAI: Beck et al., 1988a), Beck Depression Inventory (BDI: Beck et al., 1988b), Cambridge Depersonalisation Scale (CDS: Sierra and Berrios, 2000), Friburg Mindfulness Inventory (FMI: Walach et al., 2006) and Vividness of Visual Imagery Questionnaire (VVIQ: Marks, 1973) were administered before entering the mock scanner.

The fMRI session consisted of 6 runs: the first and the final runs were affective probe runs and the middle four runs were training runs. Each rtfMRI-NF training run consisted of 5 alternate baseline/increase blocks lasting 30 s each. Participants were presented with the feedback by means of a fluctuating thermometer bar (Caria et al., 2007; McCraig et al., 2011 – see Fig. 1A). A red background colour indicated when participants were required to try to increase activation (Increase blocks), and a blue background colour indicated when participants were required to return their activation to baseline levels (Baseline blocks). For the first 38 s of each run, participants were presented with a blank screen to allow the baseline NF signal to be calibrated, see Real-time fMRI data processing section for more details.

Both affective probe runs consisted of 9 alternate Increase/Baseline feedback blocks. Each block was followed by an affective probe consisting of IAPS images of normatively positive, negative and neutral valence (see Fig. 1B). IAPS images were presented for 8 s each. Four sets of pseudorandomised IAPS lists were generated to be used in each of the 4 conditions: (1) Baseline – 1st affective probe run, (2) Increase – 1st affective probe run, (3) Baseline – 2nd affective probe run and (4) Increase – 2nd affective probe run. Each set had 3 negative, 3 positive and 3 neutral stimuli. In affective probe runs, Baseline blocks were always presented first, followed by Increase blocks i.e. ABAB. To control for order effects, the IAPS sets were presented in a counterbalanced order, using a 4 × 4 Latin Square. Within each set, no significant correlations were observed between order of presentation and valence, suggesting no particular valence was over-represented at the beginning or end of the trials. After each IAPS image was presented, a modified Self-Assessment Manikin (SAM: Bradley and Lang, 1994)

According to the Edinburgh Handedness Inventory (Oldfield, 1971).
scale was shown to allow valence/arousal ratings by joystick, with initial cursor position randomised (see Fig. 1C and D).

Across groups, a 2 (baseline/increase) × 3 (neutral/positive/negative IAPS) mixed design was used. For the affective probe runs, the dependent variables (DV) were SCR, valence and arousal ratings. For the training runs the DV was BOLD signal change in the RAI.

Cognitive strategies

Participants were given standardised instructions suggesting different cognitive strategies that might help them increase the RAI activation. In line with Caria et al. (2007), one strategy involved emotional memory recall:

“Recall and relive personal memories of situations and events in which you were very physically aroused. These could be pleasant or unpleasant.”

In addition a new strategy was adapted from the mindfulness literature (Farb et al., 2007) and from data suggesting that the RAI underpins interoception (Critchley et al., 2004):

“Become aware of the tension of the muscles in your face, or limbs, the sensations created by the noise and vibration of the scanner, heartbeat or breathing rate. Do not spend too long focussing on any one sensation, but just continue to move your attention around your body.”

Participants were free to vary body/memory strategy across runs, but were asked not to switch strategies or switch between the valences (pleasant and unpleasant) of the memories recalled within any one block. Hence, participants had a choice of three different mental strategies for all increase blocks. To return RAI activation to “baseline” levels, participants were asked to silently count back from 100 in 3’s (Caria et al., 2007). At the end of the scanning session, all participants were debriefed using standardised pre-piloted questions.

Behavioural pilot study

In order to test the impact of emotion regulation strategies without NF on subjective ratings and SCR, we conducted an offline pilot study (N = 24). These participants did not take part in the subsequent NF experiment. When shown a red screen, participants in each group were required to adopt a pre-specified emotion regulation strategy (i.e. pleasant memory recall, unpleasant memory recall or body sensation focus). As in the fMRI study, participants were required to silently count back from 100 in 3’s when shown a blue screen, and were also shown a series of affective probes after each trial.

SCR data acquisition

Participants washed their hands before electrodes were attached to the distal phalanges of the 1st and 2nd digits of the non-dominant hand for the offline study and mock scanner training session, and the middle phalanges for the rtfMRI-NF study. The SC data were collected using a PsychLab Skin Conductance module with standard skin conductance gel as an electrolyte, and SC collars to standardise the skin area. The SCR time series were smoothed using a single pass of a 3-point equally weighted moving average filter. A triple pass was applied to 5 datasets that were particularly noisy, and yielded further event-related fluctuations in two of them.
fMRI data acquisition

Functional images were acquired on a 3T GE HDx MRI system (General Electric, Waukesha, USA) using GE’s standard receive-only 8-channel head coil for signal reception, and the body coil for transmission. Single shot gradient recalled EPI sequence with a parallel imaging (‘ASSET’, reduction factor 2) was used for BOLD fMRI. For the functional runs, 38 2.3 mm slices with 1 mm gap were collected, with a 22 cm FOV and a 64 x 64 matrix (giving 3.4 mm in plane voxel size), TR = 2 s, TE = 30 ms, and an Ernst flip angle of 75° was used which is optimal for a grey matter T1 of approximately 1600 ms at the TR used (Wright et al., 2008).

Real-time fMRI data processing

A custom rtfMRI system (developed by Dr Jerzy Bodurka) was run on the MR scanner computer system to provide immediate access to the fMRI images as they were reconstructed (Bodurka and Bandettini, 2008). These images were then transferred to an external Linux workstation where they were pre-processed using the AFNI software (Cox, 1996), which has built-in real-time capacities, e.g. real-time co-registration for motion correction, to increase analysis sensitivity and diminish artefacts. An image mask of pre-selected regions of interest (ROI) based on neuroanatomical criteria was applied to the pre-processed images, to extract the mean neural activation, in real time, from each ROI. As each new brain volume was acquired, AFNI calculated a new set of values for each ROI. These values were used to generate running time series for the feedback signal, which were smoothed using a 3-point moving average with weightings set at 0.125, 0.25 and 0.625. Global brain effects were controlled by subtracting activation from a large background region (c.f. below) to cancel any non-specific global BOLD changes e.g. from respiration, etc. The feedback signal was calculated as:

\[
ROI_{\text{EXP}} \left( \text{BOLD}_{\text{increase}} - \text{BOLD}_{\text{baseline}} \right) - ROI_{\text{BACK}} \left( \text{BOLD}_{\text{increase}} - \text{BOLD}_{\text{baseline}} \right)
\]

The first half of the above equation calculates the averaged BOLD response within the neurofeedback ROI_{\text{EXP}}, and the second half calculates the averaged BOLD response from a bigger background ROI_{\text{BACK}} (see the next section for the selection of these ROIs). BOLD_{\text{increase}} denotes the instantaneous activation at each volume during the increase block, and BOLD_{\text{baseline}} denotes the mean activation of the proceeding baseline block. Feedback was close to ‘real time’, although unavoidably lagged due to the haemodynamic response and the time necessary for data processing (5–7 s). However, careful debriefing suggested that none of our pilot participants found this to be problematic.

Regions of interest selection

For each experiment, participants underwent a preliminary localiser scan to allow the ROI to be defined for each participant. A structural rather than functional localiser was used, as this was thought more likely to be amenable to clinical applications. The structural localiser used was a high resolution T2 weighted Fast Spin Echo (FSE) scan, which does not suffer from the distortions inherent in an Echo Planar Imaging (EPI) scan. In addition, an EPI scan matched to the fMRI sequence for the geometric distortion was also collected from the same slice locations and used to guide ROI placement. As shown in Fig. 2A, the ROI_{\text{EXP}} was
a $3 \times 3 \times 3$ voxel cube, equivalent to $10.2 \times 10.2 \times 10.2$ mm, in the RAI; it was selected in native space according to the following anatomical landmarks: the Sylvian fissure, anterior/superior peri-insular sulcus, falciform, and central insular sulci (Craig, 2009; Crespo-Facorro et al., 2000; Ture et al., 1999). Control feedback was taken from an ROICTRL in the middle parahippocampal region (Weiskopf, 2004) of the same size as the ROILExp (see Fig. 2B). The larger background ROIBACk was selected on the basis of areas not activated in our experimental pilots and sufficiently distal to the ROILExp for cross contamination to be unlikely. It was again selected in native space in the left inferior parietal region (see Fig. 2C), and size was $6 \times 6 \times 6$ voxels cube, equivalent to 8 times the size of ROILExp. Piloting confirmed negligible signal variation in this region in either condition.

**Off-line fMRI data analysis**

In addition to the “real-time” processing described above, the fMRI data were also analysed offline, in a more conventional manner. These analyses employed the XBAM software (version 4) developed at the King’s College London’s Institute of Psychiatry (Brammer et al., 1997), which uses a permutation based nonparametric approach to minimise normality assumptions (Thirion et al., 2007). We firstly modelled the fMRI time series and contrasted between regulation and baseline blocks. Then, the statistical maps were normalised onto a Talairach template (Talairach and Tournoux, 1988) for group analysis.

To analyse the specific effect of neurofeedback on modulating RAI activity, we first performed a hypothesis driven ROI analysis for the interaction between the groups (RAI and control NF) and the linear trend of the correlation coefficient at each voxel inside a bilateral mask of insular cortex. The insular ROI was defined anatomically in the standard space using the Talairach Daemon database (Lancaster et al., 1997, 2000). After the confirmatory ROI analysis, we performed an exploratory whole brain linear trend analyses for both the RAI NF and control NF groups separately. This was to investigate which other brain regions were involved when attempting to self-regulation the RAI. False positives due to multiple comparisons were controlled for when the results were extended from the voxel to the 3D cluster level (Bullmore et al., 1999).

**Individual analysis**

The data were first realigned (Bullmore et al., 1999) to minimise motion-related artefacts and smoothed using a Gaussian filter (FWHM 7.2 mm). Time series analysis was then carried out by first convolving the experimental condition with two Gamma Variate functions with delays of 4 and 8 s, respectively, to allow for variability within this range. The weighted sum of these two convolutions that gave the best fit (least-squares) to the time series at each voxel was then determined, and a goodness of fit statistic consisting of the ratio of the sum of squares of deviations from the mean intensity value due to the model (fitted time series) divided by the sum of squares due to the residuals (original time series minus model time series) was computed. This statistic is known as the SSQ-ratio.

The appropriate null distribution for assessing significance of any given SSQ-ratio was then computed using a resampling method (Bullmore et al., 2001) and applying the model-fitting process to the resampled data. This process was repeated 20 times at each voxel and the data combined over all voxels to give the overall null distribution of SSQ-ratio. The same permutation strategy was applied at each voxel to the data combined over all voxels to give the overall null distribution of resampled data (Brammer et al., 1997). Median values were used to minimise outlier effects. To increase sensitivity and reduce the multiple comparisons problem encountered in fMRI, hypothesis testing was carried out at the 3D cluster level using the method initially developed by Bullmore et al. (1999) for structural image analysis, and subsequently shown to give excellent cluster-wise type I error control in both structural and functional fMRI analyses. The voxel and cluster statistical thresholds were adjusted in such a way as to produce less than one false positive 3D cluster per map.

**Linear trend analysis**

Following transformation of the statistical (SSQ-ratio) maps for each individual into standard space, it is possible to perform a randomisation-based test for interaction. First, the linear trend of the median SSQ-ratio values across each run was calculated at each voxel. At each voxel, participants’ linear trend was then randomly and independently reallocated to the four runs. We used the standard XBAM resampling strategy, which permutes between individuals but not between groups because we make the assumption that observations on an individual are correlated but observations across individuals are not correlated. Indeed, it is the existence of a correlation on an individual (i.e. that there is a time effect so that individuals behave differently at different runs during the real time NF training) that we wish to test. In this case, the null hypothesis is that there is no effect of runs. The latter operation yields the distribution of linear correlation coefficient under the null hypothesis of no effect of linear trend across runs with a sufficiently large number of permutations. Voxel-wise maps of significant linear trend at any desired level of type I error can then be obtained using the appropriate threshold from the null distribution. Provided that identical permutations are carried out at each voxel (to preserve spatial correlations) this method can then be extended to yield cluster-level effects (Bullmore et al., 1999). The voxel and cluster statistical thresholds were adjusted in such a way as to produce less than one false positive cluster per map.

In fMRI analysis, most commonly used assessments of the significance of the fit of the resulting model use normal theory, and the validity of the normality assumption is rarely tested. The method of analysis described earlier (XBAM) makes no such assumptions. Instead, it uses median statistics to control outlier effects and permutation rather than normal theory-based inference. Furthermore, the test statistic (SSQ-ratio) is computed by standardising for individual difference in residual noise before embarking on second level, multi-subject, testing using robust permutation-based methods. This allows a mixed effects approach to analysis – an approach that has recently been recommended following a detailed analysis of the validity and impact of normal theory based inference in fMRI in large number of participants.

**Analysis based on reinforcement learning model**

During NF, participants were asked to control the feedback signal shown on the screen, and their success or failure at this task could be formulated as an implicit reward or punishment. For this reason, it has been suggested that rtfMRI-NF follows the principles of operant learning (Caria et al., 2011). However, it is important to know whether NF depends on operant conditioning or conscious cognitive control, or both, and whether the learning mechanism changes depending on feedback conditions (Scharnowski et al., 2012). In addition, a detailed learning model is thought to be critical for the development of rtfMRI (Weiskopf, 2012). We argue that the same learning process is likely to be implemented for both the RAI and the control NF group, because they were given identical behavioural instructions, although the control NF group had less likelihood of success based on the suggested strategies.
In our study, participants were asked to use a single mental strategy within each run and only make a decision at the end of the run as to whether a different strategy should be used in the next run. This means that participants needed to estimate the overall reward value (an index of how successfully they have performed the task) for the strategy they used in the previous run, and to choose which strategy to use in the next run. In this study, we modelled this decision-making process as a “N-arms bandit problem”, which can be learnt by reinforcement-learning algorithm (Sutton and Barto, 1998). In a typical N-arms bandit problem, participants face several possible choices, which produce a variable reward/punishment according to different probability distributions. This is similar to the current study whereby participants needed to try different mental strategies and decide on the one that resulted in optimal control of the feedback signal.

Temporally discounted reward

During NF, the success of a mental strategy in controlling the NF signal changes from time to time, and provides an implicit reward associated with the mental strategy. When a reward is non-stationary, the recent reward tends to be more informative than the past reward in predicting the effectiveness of the mental strategies in the future. In reinforcement learning, this is reflected by down-weighting the past reward and up-weighting the recent reward. We use the formula below to estimate the temporally discounted reward values for all six runs in the experiment since a linear model is unlikely to capture the general dynamic of biological learning.

\[ R_t = \sum_{i=1}^{t} r_i (1 - \gamma)^{t-i} \]

We refer to the above as the exponentially recency-weighted average (Sutton and Barto, 1998), where the total amount of reward \( R_t \) at the end of the training run is a function of the reward values at each fMRI volume, at which the feedback signal for the participants was updated. The reward at each volume is denoted by \( r_i \), where \( i \) is an index of volume, and \( r_i \) was updated every 2 s. Hence, the total reward \( R_t \) reflects the total amount of BOLD signal change for each run. \( \gamma \) is a discounting parameter that ranges between 0 and 1, and we have chosen to use \( \gamma = 0.1 \), which is a standard value based on the literature (Sutton and Barto, 1998). Note that when calculating the reward value for baseline blocks, we flipped the sign of the feedback signal since participants were instructed to return their brain activation to baseline levels. We then estimated reward values across all six runs and correlated this with the BOLD signal. This analysis is based on the model-based fMRI approach (O’Doherty et al., 2007).

Exploration or exploitation?

One of the simplest types of reinforcement learning used to solve the N-arms bandit problem is called the ε-greedy algorithm (Sutton and Barto, 1998). Based on this model, a participant applies the best mental strategy at the majority of the time to maximise the reward i.e. exploitation, and only occasionally selects another strategy i.e. exploration (with a small probability 1–ε randomly sampled from a uniformed distribution) in order to see if it might give a better payoff. \( \varepsilon \) is a parameter that controls the exploration rate that is set between 0 and 1 inclusive. The ε-greedy algorithm often assumes that, at the beginning of learning, \( \varepsilon \) is small, so that participants explore the different mental strategies more extensively. Then, \( \varepsilon \)-gradually increases to encourage the exploitation of the best strategy. In our study, it was difficult to directly estimate the \( \varepsilon \) parameter because the learning was too short; instead, we have calculated an alternative metric called the exploration width, which is defined as the number of different mental strategies a participant has applied in the current run, the previous run and the following run. This means the exploration width for training run 1 is the number of different strategies used in affective probe run 1, training runs 1 and 2. This value indicates how much the participants have explored the usefulness of different mental strategies in controlling the feedback signal. We have computed this metric and predicted that, if the participants were employing the ε-greedy algorithm, the exploration width should be high at the start of the training as participants would not yet know which strategy leads to control of the target brain region, and it will progressively decay across runs as participants learn to control the feedback.

Behavioural data analysis

Where the data met parametric assumptions, the general linear model was used to examine hypotheses with paired t tests used to examine simple effects. Otherwise, alternative non-parametric tests were used with Wilcoxon Signed Ranks tests. Two-tailed p values are given and family-wise error rate was controlled using the False Discovery rate (Benjamini and Hochberg, 1995). Where assumptions for sphericity were not met, corrected Greenhouse-Geisser values are reported.

SCD data analysis

Amplitude was defined as the highest phasic increase in conductance initiated 1–8 s after stimulus onset and exceeding 0.01 μA. Mean SCD amplitude was calculated as the mean value for trials whereby a measurable response occurred. Mean SCD magnitude data were calculated as the mean value including non-responses, which were coded as zero (Dawson et al., 2007). Probability of a response regardless of amplitude (number of responses above 0.01 over total number of presentations) was also calculated. Where sample sizes are small and/or skewed due to the number of “no measurable responses”, non-parametric statistics were used with exact two-tailed significance reported. One participant had missing SCD data due to technical difficulties with the software.

Results

Pilot data

3 × 2 × 3 mixed ANOVAs were conducted to examine subjective ratings of arousal and valence. For valence ratings, no main effect for condition was observed \((F(1,42) = .01, p = .92)\) or group x condition interaction \((F(2,42) = 0.5, p = .95)\), suggesting no difference in valence ratings as a result of using emotion regulation strategies without NF. Similarly, no main effect for condition \((F(1,42) = .014, p = .91)\) or group x condition interaction \((F(2,42) = 1.26, p = .31)\) were observed for arousal ratings.

A hypothesis-led Friedman’s test revealed no significant between-condition differences in SCD magnitude. A marginal between-condition difference in SCR amplitude in participants who were asked...
to recall pleasant memories was observed ($X^2 = 9.7, p = .084$) driven by higher SCR to negative stimuli although this did not survive correction for multiple testing ($\text{Inc.} > \text{Baseline}$, Wilcoxon Signed Ranks: $z = -1.76, p = .24, n = 5$).

Overall, these data suggest that, in the absence of neurofeedback, emotion regulation strategies do not impact on arousal, valence or SCR ratings in contrast to a baseline control condition. Debrieval also confirmed the feasibility of the body sensation focus strategy.

**Demographic, behavioural and psychophysiological data**

There were no group differences in age, verbal IQ, RAI, BDI, state CDS, trait CDS, FMI or VVIQ. Eyes Open (See Table 1). One participant in the RAI NF group completed the CDS trait questionnaire but had missing data for the other variables due to time constraints on the day of testing.

**Valence ratings**

A $2 \times 2 \times 3$ (baseline/increase $\times$ group $\times$ stimuli valence) mixed ANOVA for subjective valence ratings and revealed a main effect for normative stimuli valence ($F_{1,44} = 31.26, p < .001$) in the expected direction. No main effect was observed for experimental group i.e. control vs. real feedback ($F_{1,22} = .1, p = .74$). A Valence $\times$ Group interaction was observed ($F_{1,22} = 6.63, p = .003$), driven by more pleasant ratings for positive stimuli in the control group ($t = 3, df 22, p = .02$). However, no significant Time $\times$ Group interactions were observed, suggesting an absence of pre-training or training related differences (Table 2).

**Arousal ratings**

A $2 \times 2 \times 3$ mixed ANOVA for arousal ratings revealed a main effect for normative stimuli valence ($F_{1,44} = 32.64, p < .001$) again in the expected direction. No main effect was observed for group ($F_{1,22} = .54, p = .47$). A marginal Valence $\times$ Group interaction was observed ($F_{2,22} = 2.6 p = .08$), driven by higher arousal ratings for negative stimuli in the control group, but this did not survive correction for multiple testing ($t = -2.1, df 22, p = .17$). However, no significant Time $\times$ Group interactions were observed, suggesting an absence of pre-training or training related differences (Table 2).

**Skin conductance data**

All participants showed measurable skin conductance (SC) in the 30 s prior to the start of the paradigm and there were no differences

---

**Table 2**

Mean and SD for valence and arousal ratings, and median and interquartile range (IQR) for skin conductance response (SCR) magnitude following affective probes.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Neutral stimuli</th>
<th>Positive stimuli</th>
<th>Negative stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Valence</td>
<td>Arousal</td>
<td>SCR</td>
</tr>
<tr>
<td>RAI NF group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-train baseline</td>
<td>$4.88 \pm .58$</td>
<td>$3.03 \pm 2.47$</td>
<td>.00 (.12)</td>
</tr>
<tr>
<td>Pre-train increase</td>
<td>$4.84 \pm 6.7$</td>
<td>$3.03 \pm 2.36$</td>
<td>.00 (.00)</td>
</tr>
<tr>
<td>Total pre-train</td>
<td>$4.86 \pm .54$</td>
<td>$3.03 \pm 2.39$</td>
<td>.00 (.07)</td>
</tr>
<tr>
<td>Post-train baseline</td>
<td>$4.85 \pm .5$</td>
<td>$3.19 \pm 2.34$</td>
<td>.046 (.18)</td>
</tr>
<tr>
<td>Post-train increase</td>
<td>$4.81 \pm .57$</td>
<td>$3.4 \pm 2.4$</td>
<td>.00 (.22)</td>
</tr>
<tr>
<td>Control group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total post-train</td>
<td>$4.83 \pm .4$</td>
<td>$3.3 \pm 2.3$</td>
<td>.023 (.2)</td>
</tr>
<tr>
<td>Overall total</td>
<td>$4.85 \pm .28$</td>
<td>$3.16 \pm 2.32$</td>
<td>.012 (.17)</td>
</tr>
<tr>
<td>Pre-train baseline</td>
<td>$4.83 \pm .78$</td>
<td>$2.49 \pm .97$</td>
<td>.01 (.17)</td>
</tr>
<tr>
<td>Pre-train increase</td>
<td>$4.96 \pm .49$</td>
<td>$2.66 \pm 1.22$</td>
<td>.01 (.04)</td>
</tr>
<tr>
<td>Total pre-train</td>
<td>$4.9 \pm .59$</td>
<td>$2.56 \pm 1.03$</td>
<td>.025 (.08)</td>
</tr>
<tr>
<td>Post-train baseline</td>
<td>$5.04 \pm .63$</td>
<td>$2.79 \pm 1.34$</td>
<td>.07 (.14)</td>
</tr>
<tr>
<td>Post-train increase</td>
<td>$4.75 \pm .3$</td>
<td>$2.58 \pm .83$</td>
<td>.007 (.07)</td>
</tr>
<tr>
<td>Overall total</td>
<td>$4.9 \pm .37$</td>
<td>$2.69 \pm 1$</td>
<td>.07 (.09)</td>
</tr>
</tbody>
</table>

**Fig. 3.** (A) Brain activation that shows a significant Group $\times$ Linear Trend interaction in the ROI analysis. (B) Scatter plot depicting the SSQ-ratio (a goodness of fit statistic from the fMRI time series modelling) of the peak voxel within the significant RAI cluster for the RAI NF group in red, and within the same brain region for the control NF group in blue. The solid lines in the scatter plot represent linear regression trend lines for each group.

---

**Notes:**

- **$p < .05$.**
- Pre-training refers to the first affective probe run and Post-training refers to the last affective probe run; Baseline refers to trials where participants were required to return brain activation to baseline by counting back from 100 in 3’s; Increase refers to trials where participants were required to increase brain activation; Valence and arousal ratings range from 1 to 9 with 9 representing most unpleasant, and most arousing; SCR Magnitude i.e. individual participants mean value for trials including non-responses which were coded as zero.

---

**References:**


---

**Table 1:**

Mean and SD for valence and arousal ratings, and median and interquartile range (IQR) for skin conductance response (SCR) magnitude following affective probes.
between groups in mean SC level ($t = .56$, df 21, $p = .59$). Five participants did not show a measurable SCR to any affective probe i.e. 21% of the sample, but 3 of these participants showed non-event related fluctuations. These participants were all in the RAI feedback group. There were no significant between-group differences in probability of response for neutral ($t = −.05$, df 21, $p = .97$), negative ($t = .16$, df 21, $p = .87$) or positively valenced stimuli ($t = −.46$, df 21, $p = .65$).

Friedman’s test revealed no significant between-condition differences in mean SCR magnitude in the control group ($\chi^2 = 11.5$, $p = .4$). However, a main effect was observed in the RAI NF group ($\chi^2 = 20.4$, $p = .04$). This effect was driven by increased SCR to positive stimuli shown after Increase blocks during the post-training affective probe run, however, this comparison did not survive correction for multiple testing ($Increase > Baseline$, Wilcoxon signed ranks: $z = −2.2$, $p = .18$). In addition, in the RAI group, only 4 participants showed a measurable SCR to positive stimuli in both conditions. Two participants showed a measurable SCR in only one condition i.e. Increase or Baseline, so this score was contrasted statistically with a score of zero representing a non-response, and 9 participants did not show a measurable SCR response. When examining the amplitude data, which do not include non-responses, no statistically significant effect was observed ($Baseline > Increase$ Wilcoxon Signed ranks: $z = −1.8$, $p = .41$).

**Functional MRI data**

**Region-of-interest analyses**

Within the bilateral insular ROI, we found a significant group x linear trend interaction across the 4 training runs (see Fig. 3A). The single significant cluster overlaps with the predefined ROI$_{INS}$ (cluster maxima: Talairach coordinates $x = 43$, $y = 15$, $z = 3$, $p = .03$). Fig. 3B shows
the SSQ-ratio of the peak voxel within the significant RAI cluster for the RAI NF group in red, and within the same brain region for the control group in blue. The solid lines in the scatter plot represent linear regression trend lines for each group showing a greater training effect for the RAI NF group than the control NF group. Please see Supplementary Information for an additional figure showing the average BOLD response (SSQ values) in both groups, over the 4 training runs, extracted from the RAI ROI contrasting Increase vs. Baseline blocks.

Whole brain analyses

In the exploratory whole brain analysis, there was a group \times linear trend interaction but it did not survive whole brain correction, so we tested the linear trend for each group separately. We found a significant positive linear trend in RAI activation across the 4 training runs (linear increase cluster maxima: Talairach coordinates \(x = 32, y = 18, z = 16\)) in the group receiving RAI feedback (see Fig. 4A), but not in participants receiving control feedback (see Fig. 4B). The group receiving RAI feedback also showed a positive linear trend in activation in the caudate body and posterior thalamus across training (see Table 3 and Fig. 4A). A significant negative linear trend was observed in the dorsomedial prefrontal cortex (DMPFC, linear decrease cluster maxima: Talairach coordinates \(x = 25, y = 22, z = 16\)) in the RAI feedback group (see Table 3, Figs. 4A and S1).

No positive linear increase in RAI activation was observed in the group receiving control feedback from the middle parahippocampus region. However, a positive linear trend in activation across training runs was observed in the putamen, extending into the mid and posterior insula, caudate and thalamus (see Table 4, Figs. 4B and S1). In addition, neither group displayed a linear increase in the left anterior insula. In the control NF group, a negative linear trend across training was also found in the primary visual cortex and cerebellum (see Table 4 and Fig. S1).

Neural correlates of reinforcement learning parameters

For the RAI NF group (except for 1 participant whose reward value data were missing), activations in the dorsal anterior cingulate cortex (dACC: positive correlation cluster maxima: Talairach coordinates \(x = 18, y = 44, z = 29\)) and left supramarginal gyrus (positive correlation cluster maxima: Talairach coordinates \(x = -36, y = -37, z = 29\)) were positively correlated with the temporally discounted reward values (see Table 5, Figs. 4C and S2). A negative correlation with the reward values was seen in the anterior cingulate (ACC: negative correlation cluster maxima: Talairach coordinates \(x = 25, y = -22, z = 26\)) and left inferior and rostral to the positive cluster (see Fig. 4C). The primary visual cortex (negative correlation cluster maxima: Talairach coordinates \(x = 25, y = -29, z = -3\)) was also negatively correlated with the reward in the RAI NF group (see Table 5, Figs. 4C and S2).

For the control feedback group, activation in the same region of the dACC (positive correlation cluster maxima: Talairach coordinates \(x = 28, y = 22, z = 26\)) was positively correlated, but to a lesser extent, with temporally discounted reward values (see Table 5, Figs. 4D and S2). A negative correlation was found between activation in the posterior cingulate cortex (cluster maxima: Talairach coordinates

### Table 3

<table>
<thead>
<tr>
<th>Cluster size</th>
<th>Tal(x)</th>
<th>Tal(y)</th>
<th>Tal(z)</th>
<th>Cluster P value</th>
<th>Peak of cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive linear trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D cluster</td>
<td>180</td>
<td>32</td>
<td>18</td>
<td>16</td>
<td>.002316</td>
</tr>
<tr>
<td>2D clusters¹</td>
<td>39</td>
<td>14</td>
<td>11</td>
<td>23</td>
<td>Caudate body</td>
</tr>
<tr>
<td>29</td>
<td>18</td>
<td>30</td>
<td>0</td>
<td>Caudate body</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>18</td>
<td>19</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>18</td>
<td>33</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>16</td>
<td>19</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative linear trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D cluster</td>
<td>211</td>
<td>7</td>
<td>33</td>
<td>50</td>
<td>.001228</td>
</tr>
<tr>
<td>2D clusters¹</td>
<td>45</td>
<td>11</td>
<td>33</td>
<td>50</td>
<td>Medial frontal gyrus (BA6)</td>
</tr>
<tr>
<td>27</td>
<td>14</td>
<td>7</td>
<td>53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>11</td>
<td>12</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>29</td>
<td>48</td>
<td>26</td>
<td>Superficial frontal gyrus (BA10)</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>18</td>
<td>41</td>
<td>43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>29</td>
<td>48</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>18</td>
<td>48</td>
<td>33</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹ 2D clusters are derived from the 3D cluster rather than from additional analyses, and are given to provide further information as to the spatial extent of the 3D cluster.

### Table 4

<table>
<thead>
<tr>
<th>Cluster Size</th>
<th>Tal(x)</th>
<th>Tal(y)</th>
<th>Tal(z)</th>
<th>Cluster P Value</th>
<th>Peak of Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive linear trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D cluster</td>
<td>770</td>
<td>25</td>
<td>11</td>
<td>0</td>
<td>.000165</td>
</tr>
<tr>
<td>2D clusters</td>
<td>72</td>
<td>40</td>
<td>0</td>
<td>3</td>
<td>Posterior insula</td>
</tr>
<tr>
<td>70</td>
<td>11</td>
<td>7</td>
<td>7</td>
<td>Caudate body</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>0</td>
<td>-22</td>
<td>0</td>
<td>Red nucleus</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>25</td>
<td>11</td>
<td>0</td>
<td>Putamen</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>0</td>
<td>-52</td>
<td>-16</td>
<td>Colmnus (cerebellum)</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>Hypothalamus</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>14</td>
<td>-48</td>
<td>-23</td>
<td>Dentate (cerebellum)</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>4</td>
<td>-44</td>
<td>-10</td>
<td>Cerebellar lingual</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>7</td>
<td>11</td>
<td>10</td>
<td>Caudate body</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>11</td>
<td>-30</td>
<td>-3</td>
<td>Pulvinar (Thalamus)</td>
<td></td>
</tr>
<tr>
<td>Negative linear trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D cluster</td>
<td>205</td>
<td>-4</td>
<td>-67</td>
<td>13</td>
<td>.002621</td>
</tr>
<tr>
<td>158</td>
<td>0</td>
<td>-48</td>
<td>40</td>
<td>.001979</td>
<td>Cerebellum</td>
</tr>
</tbody>
</table>

### Table 5

<table>
<thead>
<tr>
<th>Cluster size</th>
<th>Tal(x)</th>
<th>Tal(y)</th>
<th>Tal(z)</th>
<th>Cluster P value</th>
<th>Peak of cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAI NF group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive correlation</td>
<td>3D cluster</td>
<td>132</td>
<td>-36</td>
<td>-37</td>
<td>29</td>
</tr>
<tr>
<td>117</td>
<td>18</td>
<td>44</td>
<td>29</td>
<td></td>
<td>Anterior cingulate</td>
</tr>
<tr>
<td>Negative correlation</td>
<td>3D cluster</td>
<td>365</td>
<td>25</td>
<td>-89</td>
<td>-3</td>
</tr>
<tr>
<td>126</td>
<td>18</td>
<td>37</td>
<td>3</td>
<td></td>
<td>Anterior cingulate</td>
</tr>
<tr>
<td>Control NF group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive correlation</td>
<td>3D cluster</td>
<td>225</td>
<td>-3</td>
<td>19</td>
<td>26</td>
</tr>
<tr>
<td>113</td>
<td>-25</td>
<td>-48</td>
<td>-7</td>
<td></td>
<td>Fusiform gyrus</td>
</tr>
<tr>
<td>Negative correlation</td>
<td>3D cluster</td>
<td>134</td>
<td>25</td>
<td>11</td>
<td>56</td>
</tr>
<tr>
<td>103</td>
<td>4</td>
<td>-37</td>
<td>33</td>
<td></td>
<td>Cingulate gyrus</td>
</tr>
<tr>
<td>77</td>
<td>51</td>
<td>-48</td>
<td>33</td>
<td></td>
<td>Supramarginal gyrus</td>
</tr>
<tr>
<td>72</td>
<td>0</td>
<td>-51</td>
<td>46</td>
<td></td>
<td>Precuneus</td>
</tr>
<tr>
<td>58</td>
<td>-40</td>
<td>-44</td>
<td>-20</td>
<td></td>
<td>Fusiform gyrus</td>
</tr>
<tr>
<td>53</td>
<td>61</td>
<td>-25</td>
<td>7</td>
<td></td>
<td>Superior temporal gyrus</td>
</tr>
</tbody>
</table>
always (Delgado et al., 2004). Consistent with this, activation in the thalamus incentives and so they likely re of the thalamus (see Fig. S1). These regions have been previously of the thalamus (see Fig. S1). These regions have been previously

Use of mental strategy

The debriefing data suggested that 22/24 participants tried at least 2 strategies from the 3 available. Overall, 9 participants used the body strategy most frequently (6 in RAI group), 5 participants used the unpleasant recall strategy (3 in the RAI feedback group) and 8 chose to use pleasant recall more frequently (5 in the RAI group). The remaining 2 participants used each strategy twice. 12/16 participants who received NF from the RAI group and 4/8 participants who received control NF, had used the same strategy on the last training run as they used for the subsequent final affective probe run, and 10/24 used the same strategy on the last training run as the strategy they used most overall.

Participants were also asked “Did you manage to stay focussed on the level of the thermometer whilst performing the suggested mental strategies?” The response options were “not at all, sometimes and always”. 50% of the control group and 87.5% of the RAI feedback group endorsed “sometimes” and with the remainder endorsing “always” (χ = 4, df 1, p = .046). This suggests that the control feedback group stayed more focussed on the thermometer whilst performing the mental strategies, which may mean they exerted more effort.

Participants were also asked “at which run, approximately, did they think they managed to consistently increase the level of the thermometer?”. 50% of the control feedback group and 37.5% of the RAI feedback group responded that they did not manage to consistently increase the thermometer, with others endorsing either run 1, 2, 3 or 4 with no difference between groups (χ = 3.3, df 1, p = .05). In addition, 87.5% of the control feedback group, and 75% in the RAI feedback group, reported feeling calm as opposed to anxious about their performance or the scanner environment.

As shown in Fig. 5, there is a decrease in exploration width across the runs i.e. participants explored different strategies more in the beginning than in the end of the training for both groups (R2 = 0.623 for the RAI NF group and R2 = 0.361 for the control NF group). However, in the control NF group, the exploration width is more variable than that in the RAI NF group.

Discussion

These data provide further evidence that participants can learn to self-regulate BOLD signal in the right anterior insula (Caria et al., 2007) within one scanning session (Johnston et al., 2010), and confirmed the utility of a body focus mental strategy as an aid to NF based self-regulation of the RAI. However, we did not find evidence that RAI neurofeedback within one scanning session led to changes in SCR or subjective ratings to affective probes.

Consistent with previous data, we observed a positive linear trend in RAI activation in participants who received feedback from this region in contrast to those receiving control feedback (Caria et al., 2007). In particular, the ROI analysis showed a significant group × linear trend interaction confirming that the RAI NF group regulated RAI activation significantly more than the control feedback group. In the whole brain analysis, the RAI NF group, but not the control NF group, also showed a significant effect for self-regulation in the RAI cluster. Although the peak of activation was slightly more superior than predicted, nonetheless the RAI was the nearest grey matter (within a 3 mm range). The cluster also extended to the caudate body and posterior lateral nucleus of the thalamus (see Fig. S1). These regions have been previously observed to be involved in paradigms that manipulate reward and incentives and so they likely reflect increasing feedback-related success (Delgado et al., 2004). Consistent with this, activation in the thalamus has previously been observed to increase during NF, albeit more inferiorly (Johnston et al., 2010; McCrae et al., 2011).

Fig. 5. Mean exploration widths for four training runs for (A) RAI NF group and (B) control NF group. Error bars represent the standard error of the mean over participants.
receiving RAI feedback showed a monotonic increase in RAI activation in a fairly circumscribed region, whereas the network activated in the control group was much more diffuse, likely due to increased error processing in this condition (Murphy and Garavan, 2004).

Behavioural data from the neural feedback study indicated no training-related group differences in arousal and valence ratings to affective probes. Previous studies have found NF training-related behavioural changes, including increased valence ratings for negative affective probes, after left AI upregulation (Caria et al., 2010; Ruiz et al., 2011) and increased perception of disgust after bilateral AI regulation (Ruiz et al., 2011). It is possible that this is due to laterality differences, as previous studies have found valence ratings to be associated with activation in the left anterior insula (Anders et al., 2004; Posner et al., 2009). However, NF training-related behavioural effects have not always been clearly observable after a single NF scanning session and may be confounded by many different variables e.g. short-training time and fatigue (Johnston et al., 2011). In our study, the control group also rated positive stimuli more pleasantly overall, but this was not related to either pre- or post-training differences.

In addition, our data did not show robust training-related changes in SCR. Post-training increases in SCR to positively-valenced stimuli did not survive corrections for multiple comparisons, were based on a small sample size and included non-responses as representing zero amplitude. Although the probability of SCR response during rtfMRI was low (RAI feedback group .23; control feedback group .24), the majority of participants showed some SCR fluctuations in response to known events. This suggests that the observed low probability of response may be due to the paradigm i.e. habituation, making it more difficult to detect a difference when an affective probe is shown. However, a low probability of response was not observed with the offline pilot data (.48), which used the same procedure. It may therefore be that this reflects technical differences such as the use of distal vs. medial phalanges, differences in the length of the paradigm and hence fatigue effects, or other artefacts related to the scanning environment.

How participants choose between different mental strategies or whether mental strategies should be provided at all in NF paradigms are important issues in applying rtfMRI-NF (Sulzer et al., 2013). Our experimental design is particularly rich in this respect since we provided participants with a number of different options for choosing their mental strategies. By monitoring the effectiveness of these mental strategies in real-time, with the aid of feedback, participants could optimise their training accordingly. In addition, by applying a reinforcement learning model to explain the behaviour of the participants during NF, we found that the dACC played a key role in monitoring the reward values over time and may potentially contribute to the decision-making process for optimising the strategies. This finding is consistent with well-established theories about the functions of ACC (Holroyd and Yeung, 2012; Holroyd et al., 2004). We also found reward values to be positively correlated with voxels in the left supramarginal gyrus, which has been shown to be active in learning from implicit feedback and reinforcement (Hester et al., 2009).

Moreover, dACC activation was observed in both RAI and control NF groups. This suggests that both groups may have employed the same learning strategy, based on the dynamically changing reward signals from rtfMRI-NF. This is consistent with the fact that both groups were given the identical instructions, and the exploration width decreased across runs for both groups, suggesting they might employ a common learning mechanism. However, as expected, these two groups did differ in the likely effectiveness of applying the learning algorithm i.e. potential rewards. In addition, the rostral ACC was negatively correlated with the rewards in the RAI NF group, suggesting a potential lateral inhibition between dorsal and rostral ACC (Margulies et al., 2007).

One possible limitation of our study is our choice of control task, which was to deliver feedback from a different brain region whilst providing participants with identical instructions to the RAI feedback group. It is possible that the lack of likely success may lead controls to abandon or become discouraged by the mental strategies suggested (Sulzer et al., 2013). However, our debriefing suggestions that participants in the control NF group continued to focus on the feedback whilst using the mental strategies suggested, did not differ in terms of their subjective experience of success and that the majority were not anxious about their performance. These data suggest that the main difference between groups was whether they received feedback from the RAI or from another brain region, and so that this condition is an adequate control for the RAI NF condition.

Regarding our target ROI, a recent meta-analysis of the utilization of insula function in affective processing, found bilateral anterior insula activation to be associated with the experience of emotion, with a peak in the left insula (Duerden et al., 2013). In addition, positive stimuli showed a left-hemisphere dominance, and negative stimuli activated the bilateral insula (Duerden et al., 2013). Given that interoceptive processing is likely to be underpinned by the right anterior insula (Zaki et al., 2012), future studies may benefit from either bilateral anterior insula as a target ROI, or tailoring their target ROI to specific mental strategies i.e. negative, positive memory recall, or interoceptive body focus.

In summary, our main aim was to replicate the findings for previous studies regarding the possibility of self-regulating the RAI with the aid of rtfMRI neurofeedback. We were also keen to determine whether this would be possible within just one scanning session, as this has a direct bearing on potential clinical applications. In addition, we further investigated the role of mental strategies in NF-related self-regulation of the RAI, and included a mindfulness-based strategy which may be more amenable to some clinical populations than generating imagery and/or recalling emotional autobiographical memories. Our data also suggest that participants receiving NF behaved like ‘greedy learning agents’, and that the dorsal anterior cingulate cortex may be a key neural structure that supports the reward related learning and the decision-making process. The next step is to examine the optimal training protocol for clinically relevant behavioural change.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2013.10.069.

Acknowledgments

We would like to thank Poppy Schoenberg, Mauricio Sierra-Siergert, Dave Lythgoe and Dave Gasston for their technical expertise. We would also like to acknowledge the support of the Psychiatry Research Trust, the Royal Society and the NIH Biomedical Research Centre at the South London & Maudsley NHS Trust and Institute of Psychiatry, King’s College London.

References


