NeuroImage 56 (2011) 753–765

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Real-time support vector classification and feedback of multiple emotional brain states

Ranganatha Sitaram [a,*,1], Sangkyun Lee [a,b,*,1], Sergio Ruiz [a,b,d], Mohit Rana [a,b], Ralf Veit [a], Niels Birbaumer [a,c]

[a] Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Gartenstr. 29, 72074 Tübingen, Germany
[b] Graduate School of Neural & Behavioural Sciences, International Max Planck Research School, 72074 Tübingen, Germany
[c] Ospedale San Camillo, Istituto di Ricovero e Cura a Carattere Scientifico, Venezia, Lido, Italy
[d] Department of Psychiatry, Faculty of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile

ARTICLE INFO

Article history:
Received 25 February 2010
Revised 31 July 2010
Accepted 3 August 2010
Available online 6 August 2010

ABSTRACT

An important question that confronts current research in affective neuroscience as well as in the treatment of emotional disorders is whether it is possible to determine the emotional state of a person based on the measurement of brain activity alone. Here, we first show that an online support vector machine (SVM) can be built to recognize two discrete emotional states, such as happiness and disgust from fMRI signals, in healthy individuals instructed to recall emotionally salient episodes from their lives. We report the first application of real-time head motion correction, spatial smoothing and feature selection based on a new method called Effect mapping. The classifier also showed robust prediction rates in decoding three discrete emotional states (happiness, disgust and sadness) in an extended group of participants. Subjective reports ascertained that participants performed emotion imagery and that the online classifier decoded emotions and not arbitrary states of the brain. Offline whole brain classification as well as region-of-interest classification in 24 brain areas previously implicated in emotion processing revealed that the frontal cortex was critically involved in emotion induction by imagery. We also demonstrate an fMRI-BCI based on real-time classification of BOLD signals from multiple brain regions, for each repetition time (TR) of scanning, providing visual feedback of emotional states to the participant for potential applications in the clinical treatment of dysfunctional affect.

Introduction

Prediction of emotional states from brain activity constitutes a major scope of affective neuroscience and would solve several pressing clinical problems such as the assessment of affect in verbally incompetent people with dementia, minimally conscious state and locked-in-syndrome, and the detection of deception. Recent advances in multivariate pattern classification of functional magnetic resonance imaging (fMRI) signals are especially important due to the high spatial resolution, whole brain coverage and non-invasiveness of fMRI. It has been pointed out (Haynes and Rees, 2006) that multivariate approaches are more sensitive in decoding brain states as they integrate spatial and temporal information from different regions of the brain, in contrast to univariate statistical parametric mapping (SPM) which analyses each brain location in isolation.

Studies on pattern classification of fMRI signals can be grouped into three major themes. Firstly, a number of methodological studies aimed to incorporate and adapt the existing methods in the field of machine learning to the classification of fMRI signals (LaConte et al., 2003, 2005; Martinez-Ramon et al., 2006; Shaw et al., 2003; Strother et al., 2004). LaConte and colleagues examined the classification of block-design fMRI data using linear discriminant analysis (LDA) (LaConte et al., 2003) and support vector machines (SVM) (LaConte et al., 2005) in contrast to canonical variates analysis (CVA). Mourao-Miranda and colleagues (2005) compared SVM and the Fisher linear discriminant (FLD) classifier and demonstrated that SVM outperforms FLD in prediction accuracy as well as in robustness of the spatial maps obtained. Shaw and colleagues (2003) showed that preprocessing strategies such as spatial and temporal smoothing and classification parameters could be derived in a subject-specific manner to result in optimal prediction. Martinez-Ramon and colleagues (2006) later developed an approach for multi-class classification by segmenting partially normalized activation maps into functional areas using a neuroanatomical atlas, classified each area separately with local classifiers, and finally performed a weighted aggregation of the multiclass outputs.

A second topic of research pertains to the application of pattern classification for obtaining greater insight into spatial and temporal

* Corresponding authors. Institute of Medical Psychology and Behavioural Neurobiology, Eberhard-Karls-University of Tübingen, Gartenstr. 29, D-72074 Tuebingen, Germany. Fax: +49 7071 295956.
E-mail addresses: sitaram.ranganatha@uni-tuebingen.de (R. Sitaram), lee.sangkyun@gmail.com (S. Lee).
[1] Authors contributed equally.

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.08.007

patterns of brain activity during cognitive, affective and perceptual states, represented by the following recent studies: neural antecedents of voluntary movement (Soon et al., 2008), visual processing (Kamitani and Tong, 2005), memory recall (Polyn et al., 2005), detection of deception (Davatzikos et al., 2005) and emotion perception (Pessoa and Padmala, 2005).

A third class of studies in pattern classification of brain signals, to which the present work belongs, is related to the rapidly advancing field of brain–computer interfaces (BCIs) and neurorehabilitation. A series of studies (Caria et al., 2006; deCharms et al., 2004, 2005; Posse et al., 2003; Rota et al., 2006; Sitaram et al., 2005; Veit et al., 2006; Weiskopf et al., 2003, 2004; Yoo et al., 2004; Yoo and Jolesz, 2002) with fMRI-BCIs demonstrated that healthy individuals as well as patients can learn voluntary regulation of localized brain regions and presented evidence for behavioural modifications that accompany self-regulation training. If the neurobiological basis of a disorder is known in terms of abnormal activity in certain regions of the brain, fMRI-BCI can be potentially targeted to modify activity in those regions with high specificity for treatment. However, a disadvantage of the existing fMRI-BCIs is in the restriction to one single region of interest (ROI) for extracting signals and providing feedback of activation to the participant. A major argument for moving away from deriving feedback signals from single ROIs is that perceptual, cognitive or emotional states generally recruit a distributed network of brain regions rather than single locations. Training subjects to merely increase or decrease BOLD signals in single ROIs may not completely model the network dynamics of the target brain state. Applying multivariate pattern classification presents itself as a potential solution to this problem as it does not make prior assumptions about functional localization and cognitive strategy. However, due to the limited computational time and resources available during online decoding, much of the offline classification methods developed so far might not be directly applicable to real-time fMRI. LaConte and colleagues (2007) reported the first real-time fMRI system with multivariate classification. Using SVM, the authors showed the feasibility of online decoding and feedback from single (repetition time, TR) fMRI data during block-design left- and right-hand motor imagery and further demonstrated the classifier's ability to decode other forms of cognitive and emotional states. The present study is in line with the above work and extends by demonstrating robust online classification and feedback of multiple emotional states.

fMRI-BCI is a promising tool for affective neuroscience and has shown potential for neurorehabilition to alleviate emotional disorders such as anxiety, sociopathy, chronic pain and schizophrenia and brain communication in the locked-in syndrome (Caria et al., 2007; deCharms et al., 2005; Rota et al., 2008; Ruiz et al., 2008). However, Phan et al. (2002) in a meta-analysis indicated that emotions such as happiness, sadness, fear, anger and disgust activate networks of several brain regions such as insula, amygdala, hippocampus, thalamus, medial and lateral prefrontal cortex, orbitofrontal cortex and anterior and posterior cingulate. In consideration of the above evidence and the limitations of single ROI feedback, we intended to implement a high performance real-time SVM classifier that could decode as well as provide feedback of emotional states of the brain from fMRI signals. It would be beneficial if the classifier could learn to generalize across intra-subject variations of emotion regulation.

The purpose of the present study is to demonstrate that multiple discrete emotional states (e.g., happy, disgust and sad) could be classified and the information could be fed back to the participant in real time. We first demonstrate two-class classification between happy and disgust emotions as these two emotions are quite distinct from each other in emotional valence. Further, these emotions are interesting from the point of view of clinical applications—namely, rehabilitation in patients with dysfunctional affect. A novelty of the present work is the demonstration of the generalization of classification by training the classifier on one type of imagery or scenario and testing it on another type of imagery. In other words, we show that the classifier is able to distinguish between two different imageries, e.g., winning a lottery and playing with a pet, as the same emotion, i.e., happy.

However, performing two-class (happy vs. disgust) classification alone may raise the question whether the classification is between positive and negative emotions in general, or whether specific discrete emotions could indeed be classified. To answer this question, we further aimed to demonstrate multiclass classification of 3 discrete emotional states, namely, happy, disgust and sad. We chose sad as the third emotion to show that the classifier can distinguish between two emotions that are similar in valence (e.g., negative), namely, disgust and sad.

We first demonstrate in 12 healthy volunteers, a high performance, real-time, two-class prediction of positive (happy) and negative (disgust) emotions and visual feedback of the predicted brain states for each TR. We show that robust classification performance could be obtained even with limited computational time by the application of preprocessing (realignment, detrending and spatial smoothing) and a method of feature selection called Effect mapping. Our previous work (Lee et al., 2010) has already reported results of rigorous comparisons between the Effect mapping method and some of the existing methods, showing clearly that the Effect mapping method fares better in classification performance.

Further to the two-class classification, we illustrate for the first time a multiclass (happiness, disgust and sadness) prediction and neurofeedback on 4 more healthy volunteers. To test whether classifier-based feedback training can help participants to improve emotion regulation, we trained two additional healthy participants for three sessions of feedback training. Finally, to obtain greater insight into the brain regions activated in emotion imagery and regulation, we carried out the following offline analyses: (1) SVM classification and Effect mapping of whole brain fMRI signals and (2) ROI classification of fMRI signals extracted from 24 brain regions previously implicated in emotion imagery and regulation.

Methods

Our aim was to develop a system that could train a classification model based on an initial set of fMRI data and brain state labels determined by the experimental protocol (training data set), thereafter to use the classification model to predict brain states from every volume of brain images (at the end of each repetition time, TR) that are acquired from the scanner, and to subsequently update the visual feedback based on the classifier's prediction.

Real-time Implementation

The real-time brain state classification system (Fig. 1) comprises of the following subsystems: (1) image acquisition subsystem written in the C programming language, (2) fMRI-BCI that performs image preprocessing, brain state classification and visual feedback, coded in a combination of C programming language and Matlab scripting language (Mathworks, Natwick, MA). The image acquisition system is centered around a modified echo-planar imaging (EPI) sequence developed in-house on the Siemens 3T TIM whole body scanner (Siemens Magnetom Trio Tim, Siemens, Erlangen, Germany), based on the image acquisition software, *Syngo* version VB13. We used the same real-time EPI sequence that was employed in our earlier studies on real-time fMRI (Rota et al., 2006, 2008; Weiskopf et al., 2003). The original implementation was described in Weiskopf et al. (2003). In brief, a standard echo-planar imaging sequence provided by the firm Siemens (Erlangen, Germany) was modified in collaboration with the manufacturer. The modifications enabled storage of functional image files in the ANALYZE format (*.hdr, *.img) to a specified directory of the host computer of the MRI scanner. The EPI sequence, otherwise, had facilities for changing the MR parameters, such as repetition time (TR), echo time (TE), number of slices and so on, similar to a standard

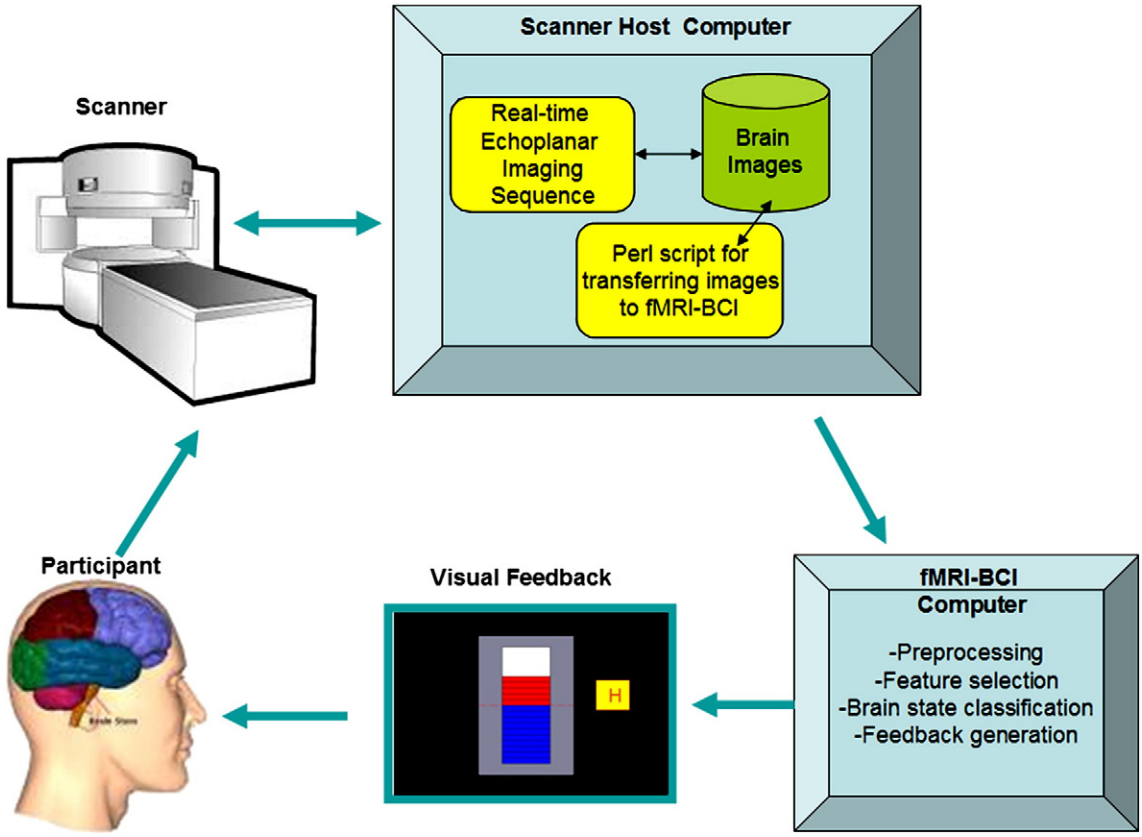

Fig. 1. The real-time fMRI brain state classification system comprises of the following subsystems: (1) image acquisition subsystem, which is a modified version of the standard echo-planar imaging (EPI) sequence written in C and executed on the scanner host computer, and (2) fMRI-BCI subsystem, which performs image preprocessing, brain state classification and visual feedback, implemented in C and Matlab scripts (Mathworks, Natwick, MA) and executed on a 64-bit Windows desktop. A Perl-script on the scanner host transfers the acquired images after every scan (at an interval of 1.5 s) to the fMRI-BCI computer.

EPI sequences available from the manufacturer. In other words, no parameters needed be hard-coded into the sequence.

Functional images were acquired with a standard 12-channels head coil by the real-time EPI sequence and were then stored in a user-specified directory on the scanner host computer. Sixteen images pertaining to 16 slices (voxel size = $3.3 \times 3.3 \times 5.0$ mm^3, slice gap = 1 mm) were accessible for preprocessing and further analysis at the end of each repetition time. The real-time sequence incorporated the following image acquisition parameters: TR = 1.5 s, matrix size = 64×64, echo time TE = 30 ms, flip angle $\alpha = 70°$, bandwidth = 1.954 kHz/pixel. For superposition of functional maps upon brain anatomy, during offline analysis, a high-resolution T1-weighted structural scan of the whole brain was collected from each subject (MPRAGE, matrix size = 256×256, 160 partitions, 1 mm^3 isotropic voxels, TR = 2300 ms, TE = 3.93 ms, TI = 1100 ms, $\alpha = 8°$). The Perl-script was a simple program that copied the functional image files in the ANALYZE format from a folder in the host computer to a Windows workstation that hosted the fMRI-BCI system.

The fMRI-BCI reads the images as they arrive, and performs online preprocessing, brain state classification and generation of visual feedback to the participant. Unlike LaConte and colleagues (2007), we did not alter the image reconstruction (IR) system of the scanner in implementing the fMRI-BCI. As the present study incorporates additional preprocessing steps such as online realignment, spatial smoothing and feature extraction, we anticipated that these procedures may introduce unpredictable delays that may affect the online reconstruction of the functional images in the image reconstruction system (IRS). To avoid this potential problem, we implemented online preprocessing, feature extraction and classification as a separate software module and executed it on the fMRI-BCI workstation. We implemented this subsystem in a 64-bit version of Matlab (version 7.1, Mathworks, Natwick, MA) by building around the C-language implementation of the core engine from SVMlight (Joachims, 1999). This approach offered us flexibility in repeated modify-and-test software life cycle of several intermediate implementations of preprocessing, brain masking, feature/voxel selection and feedback algorithms.

Real-time brain state classification on each participant is performed in several steps, as depicted in the Fig. 2. First, signal preprocessing is performed on each volume of brain images as they arrive from the scanner. In this step, real-time realignment for head motion correction and spatial smoothing are performed. In our implementation, the preprocessing was performed online during SVM training as well as testing (see Fig. 2). In the next step, a first-pass feature selection is done, whereby brain voxels are selected by choosing voxels above a user-specified intensity threshold. Subsequently, a second-pass thresholding is done, by which informative voxels are chosen from the first-pass brain mask by training an SVM classifier on the training data set. In the penultimate step, the classifier is retrained on the same data set but now with the features selected in step 3. Finally, the classifier is tested on new data collected from the same participant.

Real-time preprocessing

To correct for head motion artifacts, all brain scans were realigned to the first brain scan of the first run. The real-time implementation was adapted from the realignment algorithm (Ashburner et al., 2003; see Supplementary information and http://www.fil.ion.ucl.ac.uk/spm/doc/books/hbf2/) of SPM2. After realignment, the images were spatially smoothed in real-time with an 8-mm fixed width half maximum (FWHM) window.

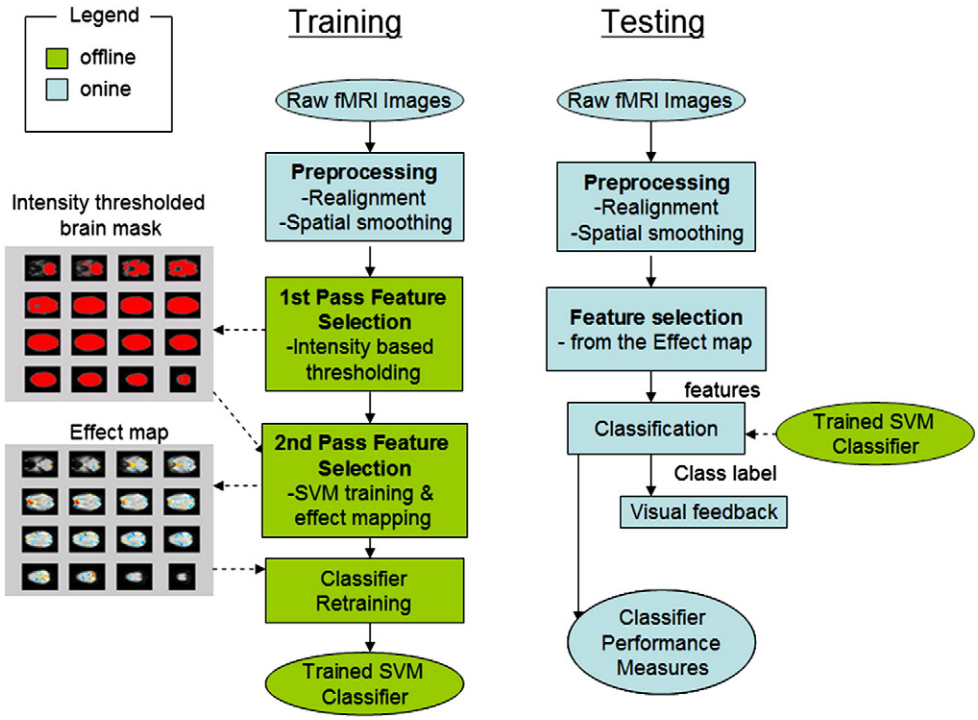

Fig. 2. Flow chart for fMRI signal preprocessing and classification. Brain state classification is performed in the following steps: (1) signal preprocessing for online realignment and spatial smoothing, (2) first-pass feature selection for selecting brain voxels by applying an intensity threshold resulting in a brain mask, (3) second-pass thresholding for selecting informative voxels from the first-pass brain mask by the method of Effect mapping resulting in the final brain mask, (3) classifier retraining based on the brain mask obtained in step 3, and (4) real-time classifier testing on new data using the second-pass brain mask.

Support vector classification

To classify an individual scan of fMRI data, all brain voxels or selected brain voxels from each repetition time (TR) can be composed of an input vector $\mathbf{x}^i$. SVM determines a scalar class label Y^i ($Y^i = \mathrm{sgn}(y^i = \mathbf{w}^T\mathbf{x}^i + b) = \pm 1$, $i = 1,\ldots, N$, where N is the number of input vectors, T is the transpose of a vector, b is a constant value, sgn(.) is a signum function, $\mathrm{sgn}(x) = +1, 0, -1$ if $x>0, x=0, x<0$, respectively) from the input vector.

When the input vectors $\mathbf{x}^i$ and the designed labels Y^i_L are taken from the training data set, the weight vector $\mathbf{w}$ of SVM is obtained by minimizing objective function L of Eq. (5) with constraints Eqs. (6) and (7),

$$L = \frac{1}{2}\mathrm{w}^T\mathrm{w} + C\sum_{i=1}^{N}\xi^i, \tag{5}$$

$$\text{with } Y^i_L(\mathrm{w}^T\mathrm{x}^i + b) \geq 1 - \xi^i, \tag{6}$$

$$\text{and } \xi^i \geq 0 \tag{7}$$

where slack variable ξ^i is introduced to describe a non-separable case, i.e., data that cannot be separated without classification error, C denotes the weighting on the slack variable, i.e., the extent to which misclassification is allowed, and a value of $C=1$ was used in our implementation because of the following reasons. Firstly, model selection to determine the C value is hard to perform in the context of real-time classification due to limitations of time available for SVM training. What is important is that real-time, online classification should work robustly in the majority of participants and sessions. In many previous fMRI classification studies (Haynes et al., 2007; LaConte et al., 2005, 2007; Mourao-Miranda et al., 2005, 2007, 2006; Soon et al., 2008), $C=1$ was successfully used. Furthermore, LaConte et al. (2005) showed that prediction accuracy does not vary much with the selection of C.

Feature selection and SVM training

The performance of a pattern classifier could be improved by a procedure called feature selection, where informative features from the data are extracted for input to the classifier. We performed feature selection in 2 steps: (1) a first-pass intensity thresholding and (2) a second-pass voxel selection using Effect mapping (Lee et al., 2010). For the first-pass intensity thresholding, an image of the brain, called intensity thresholded mask, was created by removing voxels below a certain BOLD intensity value in a mean image of all the brain scans. For the second-pass feature selection, training data from 2 runs were divided into 5 sequential subsets. That is, each subset consisted of 4.5 blocks (for example, in the case of binary classification: 2.5 blocks of condition happy, 2 blocks of condition disgust or 2.5 blocks of condition disgust, 2 blocks of condition happy) from a total number of 24 blocks. *Z*-normalization (*z*-value: (*x*-mean(*x*))/standard deviation(*x*), *x*: samples) was applied across all the time-course signals at each voxel to account for the variability of BOLD signals across runs. *Z*-normalization was performed to correct for the variability of the BOLD signal and to convert it to a signal of zero mean and unit variance. BOLD values from brain voxels of the subsets were used to create input vectors for training SVMs and subsequently to generate five Effect maps (please see Section 3.5 for a description of the Effect maps). Informative voxels were selected by applying a threshold on the mean of five Effect maps (Fig. 2). The online SVM classification software written in Matlab allows for specifying user-specified threshold. The threshold that was incorporated in this study was 0.5 for all participants and sessions. In our experience, this threshold provided a good balance between reduction of the size of the data input while maintaining good prediction accuracy. Finally, an SVM was trained from the BOLD values of the feature-selected voxels after *Z*-normalization across all the training samples.

Ideally, SVM output y^i should be centered about zero, so that when the output is greater than zero the brain state is assigned to one emotion (e.g., happy), and when it is less than zero it is assigned to the other emotion (e.g., disgust). However, due to participants' head movements and other systemic changes, a gradual drift in the classifier output can be expected (LaConte et al., 2007). To remove this bias during online classification, we subtracted the mean of the SVM outputs during the rest condition from each SVM output during the active condition (i.e., happy or disgust).

Effect mapping

Based on the parameters of the trained SVM model, we analyzed the fMRI data with a new measure called *effect value* (Lee et al., 2010) which estimates the effect of the BOLD activation at each voxel on the classifier output by first computing the *mutual information* between the voxel and the output, and then multiplying it with the weight value of the voxel as estimated by SVM. Mutual information (MI) is derived from relative entropy or the Kullback–Leibler divergence, which is defined as the amount of information that one random variable contains about another random variable (Cover and Thomas, 1991). Hence, the effect value (EV) E_k of a voxel k is defined as:

$$E_k = w_k I(x_k; y), \quad k = 1, \cdots, M \tag{8}$$

where $I(x_k; y)$ is the mutual information between the voxel and the output, y is the SVM output after excluding the sign function, w_k and x_k are weight value and activation at voxel k, respectively. To reduce the variability of EVs from Eq. (8) dependent on data set, normalization of E_k from Eq. (9) was applied as follows:

$$lE_k = \mathrm{sgn}(E_k)\log(1 + \mathrm{std}(|E|)), k = 1, \cdots, M \tag{9}$$

where sgn(.) is a signum function, and std(|E|) is standard deviation across all the brain voxels. For ease of voxel selection based on the EVs, we have used Eq. (9) for generating the Effect maps at the contrast: happy vs. disgust in binary classification, and happy vs. disgusted, happy vs. sad, and disgusted vs. sad, in our multiclass classification, respectively. Such maps can be thresholded at a suitable level by visual inspection by the experimenter and the resulting map can be used as a brain mask. This brain mask is subsequently applied to functional images arriving at each time point (TR), to reduce the data dimension and to choose the most important voxels for classification.

Generation of visual feedback

Visual feedback was provided in the form of a graphical thermometer whose bars increase or decrease in number from the baseline value (represented by the middle red dashed line of the thermometer) proportional to the classifier output after each TR. The thermometer could be operated in two modes by selection in our Matlab toolbox: (1) feedback mode and (2) non-feedback mode. In the non-feedback mode, the thermometer was displayed without any bars for the whole duration of the experiment. In the feedback mode, the graphical thermometer is shown in an animated fashion by increasing or decreasing the number of bars. The classifier output after correction of the classifier drift was used to generate the visual feedback. Depending on the classification result, i.e., +1 or −1, the thermometer bars were updated with respect to the previous TR in the positive or negative direction, respectively. The thermometer was designed to indicate not only the correctness of classification (in terms of the bar changing in the positive or negative direction) at the end of each TR, but also the cumulative performance of the past trials in terms of the height of the bars above or below the baseline. This way the participant receives feedback about his instantaneous (defined by the length of the TR) brain state as well as his past performance.

Multiclass classification

The objective of the multiclass classification experiment was to test whether the online classifier is able to classify three discrete emotions, namely, happy, sad and disgust. The multiclass problem was formulated as a problem of finding the best classification among three trained binary classifiers (i.e., classifier 1: happy vs. disgust, classifier 2: happy vs. sad, and classifier 3: disgust vs. sad) based on the selected voxels from an Effect map of each combination as described above.

Multiclass classification was performed based on the framework of error correcting output code (ECOC) from 3 binary classifications (Mourao-Miranda et al., 2006). The ECOC method assigns a unique binary string of length n, called the "code word". Subsequently, n binary classifiers are trained, one for each bit position of the binary string. During the classifier testing, new data are classified by evaluating each of the binary classes to arrive at an n-bit string, which is then compared to all the different code words. The test data is finally assigned to the code word that is the closest based on a distance measure. In our approach to the ECOC, each class has its own code $\mathbf{m}_i$ as follows: $\mathbf{m}_1 = [1\ 1\ 0]$ for classifier 1 (happy vs. disgust), $\mathbf{m}_2 = [-1\ 1\ 0]$ for classifier 2 (happy vs. sad), $\mathbf{m}_3 = [0\ -1\ -1]$ for classifier 3 (disgust vs. sad). Let $\mathbf{p}$ be a vector of predictions from the 3 classifiers. Then, the final decision was made by selecting the closest code from the prediction vector $r = \min_{\arg} i = 1^3 \mathrm{dist}(m_i, p)$. We used the Euclidian distance measure in our implementation.

Experimental paradigm

Our experiment (Fig. 3) was divided into three parts: experiment 1 for investigating real-time binary classification, experiment 2 for further feasibility testing of multiclass prediction, and experiment 3 for assessing the effect of extended feedback training with a real-time classifier. Twelve healthy students from the Department of Medicine, University of Tuebingen, Germany, participated in experiment 1 and four more students participated in experiment 2. The participants were aged in the range of 22–26 years, with a mean age of 25 years. All participants signed a written informed consent, and the study was approved by the local institutional review board.

Experiment 1 (see Fig. 3) was envisioned in 4 succeeding stages, each stage prepared as a block-design protocol with alternating blocks of happy and disgust imagery of 24-s duration, interleaved with rest blocks of 4.5-s duration. In each run, there were 6 blocks of happy imagery, 6 blocks of disgust imagery and 12 blocks of rest. In total, 233 scans were collected in a run including the 5 initial scans that were not used in the analysis, due to magnetic equilibration effects that could potentially distort the data. The rest blocks were incorporated for the sole purpose of avoiding cognitive and emotional overload due to sudden changes of imagery. During classifier training (stage 1) containing 2 runs of the above protocol, there were 12 blocks of happy imagery, 12 blocks of disgust imagery and 24 blocks of rest. During each type of the emotional imagery block, an empty thermometer (without feedback) was shown at the centre of the screen with a letter beside it indicating the participant to perform happy ('H') or disgust ('D') mental imagery. The purpose of the empty thermometer during the SVM training stage was to maintain consistency in visual stimulation with the testing stage during which the thermometer could be updated for visual feedback of the brain state. Two runs of the stage 1 protocol were performed to collect sufficient amount of data for training the SVM classifier (stage 2).

Participants were instructed well in advance of the experiment to identify one or more emotional episodes from their personal lives for each type of emotion (e.g., happy, sad or disgust) that they were

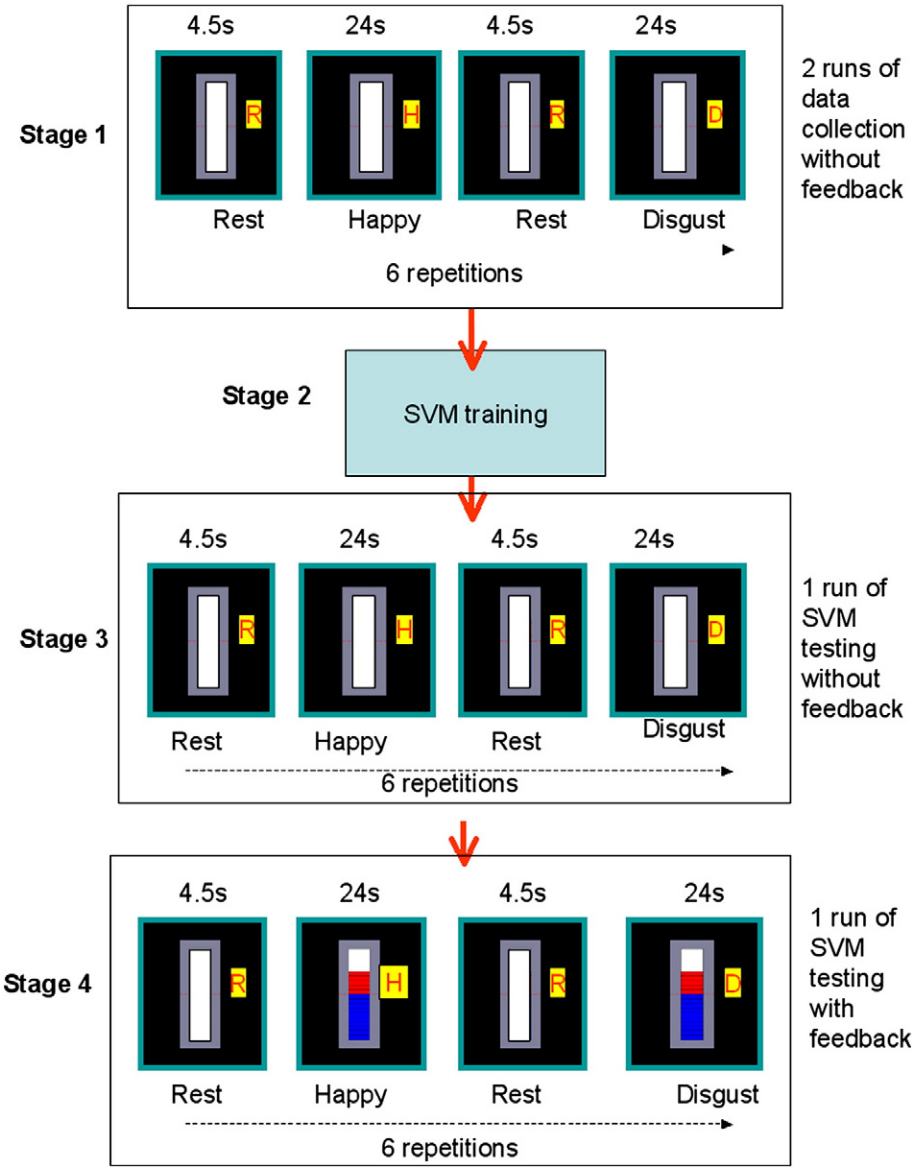

Fig. 3. Experimental protocol. Both binary (happy and disgust) and multiclass (happy, sad and disgust) classification experiments were performed in 4 succeeding stages, each stage prepared as a block-design protocol with alternating blocks of emotion imagery each of 24-s duration, interleaved with rest blocks of 4.5-s duration. During SVM training, an empty thermometer was shown at the centre of the screen with a letter beside it indicating the participant to perform happy ('H'), disgust ('D') or sad ('S') mental imagery. Two runs of the stage 1 protocol were performed to collect data for training the SVM classifier (stage 2). Stages 3 and 4 were performed to test the classifier without and with feedback, in that order.

required to recall during the experimental blocks. To maintain consistency between classifier training and testing conditions we showed participants pre-selected pictures as a reminder of the specific type of emotion imagery that they need to use in each session. Towards this end, participants were asked to identify one picture from the International Affective Picture System (IAPS) (Lang et al., 1997) that best epitomised each type of emotional episode. At the beginning of a training or testing run, participants were shown each selected picture for a duration of 30 s to remind and strengthen their emotional recall strategies. We could have alternately displayed text to describe and remind them of the specific imagery to use, but we found that pictures captured more succinctly the complexity of an emotional recall scenario. However, to avoid a potential problem that subjects might just recall the images rather than the emotion, we specifically instructed subjects to use the IAPS images as mere references to their imagery and to elaborate on the images to invoke the required discrete emotion (happy, disgust or sad emotion). We later confirmed the type and level of emotion imagery they performed with an interview and subjective rating (see Table 1 and Supplementary Table 2).

Stages 3 and 4 were meant to test the classifier without and with feedback, respectively. These two tests were performed to compare real-time classification performance in the absence and presence of feedback to evaluate the eventual application of real-time brain state decoding for BCI and clinical rehabilitation. At the end of each run, participants were asked, over the scanner audio system, to complete a self-assessment scale (1–9, 1 indicating poor performance and 9 indicating best performance) of their level of emotion regulation for each type of emotion. The purpose of this self-report was to perform an offline analysis of the correlation between subjective report and online classifier performance.

Four out of the 12 participants in experiment 1 underwent an additional run of the stage 4 protocol to test whether the classifier could robustly predict emotion states even when participants do not use the same emotion inducing strategies that were used during SVM training runs. To assess whether feedback training helps participants to improve emotion recall, we recruited two more participants and trained them on two-class (happy vs. disgust) classification for 3 sessions (in addition to 2 sessions for collection of training samples).

Table 1
PANAS scores, subjective ratings and strategies for binary and multiclass emotion imagery.

Subject	Mean positive affect sub-score of PANAS	Mean negative affect sub-score of PANAS	Subjective ratings of emotion imagery (scales 1–9, 1 = worst, 9 = best)*				Strategies for happy (H), disgust (D) and sad (S) emotions
			Training run 1	Training run 2	Test run (no feedback)	Test run (with feedback)	
Binary classification with same strategies for training and testing							
S1	30	24	7.5	7.5	6	6.5	H: romantic love D. toilet/feces
S2	41	21	6	6	6	5	H: family reunion D: eating insects
S3	26	12	8	7.5	7.5	7.5	H: good marks in exams D: blood in an open wound
S4	26	13	6	6.5	6	7	H: graduation day D: a dirty toilet
S5	27	26	6.5	5	4.5	4.5	H: riding a bike D: infected skin wound
S6	20	23	2.5	3	4	3.5	H: romantic love D: feces
S7	28	30	6	6	5.5	6.5	H: romantic love D: a dirty toilet
S8	28	18	9	6	6	7	H: baby D: a dirty toilet
S9	29	22	6.5	6	5	6	H: friends gathering D: a dirty toilet
S10	34	10	7	7	7	9	H: cats D: an used toilet
S11	23	16	7	8	9		H: friends gathering D: dirty and sick people
S12			4	8	9	9	H: falling in love D: spiders
Binary classification with different strategies for training and testing							
S9	28	13	7	6.5	6.5	5.5	Training runs: H: romantic love D: baby with feces Testing runs: H: family reunion D: dirty bathroom
S10	14	17	6	7	6	7.5	Training runs: H: romantic love D: spiders Testing runs: H: party with friends D: dirty bathroom
S11	39	10	8.5	8.5	8	9	Training runs: H: romantic love D: toilet Testing runs: H: playing the guitar D: rotten food
S12	43	11	8	8.5	8	6.5	H: getting the visa D: a dirty laboratory Testing runs: H: romantic love D: toilet
Multiclass classification with same strategies for training and testing							
S13	31	10	7.7	8.4	8	7.7	H: pets D: toilet S: braking up with boyfriend
S14	24	14	7.7	7.8	7.7	7.8	H: graduation day D: toilet S: death of a friend
S15	27	11	6.2	8	8	9	H: a boat trip D: bloody car accident S: braking up with boyfriend
S16	14	16	6.8	7.3	6.3	7.8	H: romantic love D: spiders S: view of blind beggars

*Values correspond to the mean between happy and disgust score in binary classification, or between happy, disgust and sad scores in the multiclass classification.

To adapt the classifier to the changes in the participant's brain activation with thermometer feedback and learned regulation, we retrained the classifier after each feedback session on a combined data set of the training data set (collected in the absence of feedback) and the feedback data set (collected in the presence of feedback).

Experiment 2 was designed to test real-time multiclass classification, both with and without feedback. Four more student participants (mean age = 25) were recruited for this experiment. Conceptually, experiment 2 was designed similar to experiment 1, except that additional blocks pertaining to a third emotion (sadness) was added

to train and test a three-class classifier. That is, each run consists of 4 blocks of each emotion with 12 blocks of rest condition. The block length is exactly same as the binary case. The total duration of each run of experiment 2 was also 233 scans (5.825 min).

Experiment 3 was conducted to answer the question: How can a pattern classifier provide contingent feedback in a self-regulation application where participants are normally expected to improve their ability with training to induce themselves into desired brain states? One way to approach this would be by adapting the classifier to participant's learning by incrementally retraining the classifier with new samples at the end of each session of feedback training. We have adapted this method in an extended experiment in two participants to incrementally enhance the classifier's ability to recognize changes in brain activity with self-regulation learning. We trained participants on 2-class (happy vs. disgust) classification, with 3 sessions of feedback training (in addition to 2 sessions for collection of training samples) in a day. We retrained the classifier after each feedback session by combining the previous training set with the new test data. Thermometer feedback was provided as described in section 3.6. Participants were instructed to perform emotion recall in order to increase the bars of the thermometer. They were told that thermometer bars going up meant that they were doing the task right and bars going down meant that they were probably not doing it right.

Offline whole brain and ROI analysis with classification

To assess the level of involvement of different brain regions, two additional types of SVM classifications were performed offline: (1) single-subject whole brain analysis with the E-map (Lee et al., 2010) (binary classification on 12 subjects and multiclass classification on 1 subject), and (2) ROI analysis (binary classification on 12 subjects). We expected that BOLD signals from brain regions known to be involved in emotion processing would show higher decoding accuracy for the discrete emotions under consideration (i.e., happy and disgust) than in other regions. Our ROI analysis is based on the expectation that classification accuracy will be higher in ROIs that have been previously implicated in emotion regulation and induction (Amodio and Frith, 2006; Ochsner et al., 2005, 2004; Phan et al., 2002). Towards this end, fMRI images for all participants were preprocessed with the following procedures: realignment, normalization to MNI space, and spatial smoothing with 8 mm Gaussian FWHM. First, single-subject whole brain classification was performed wherein BOLD signals from all brain voxels for each TR were collected into an input vector. For the ROI classification, 24 masks (see Supplementary Fig. 2) from brain regions previously reported to be involved in emotion imagery, recall and regulation (Amodio and Frith, 2006; Ochsner et al., 2005, 2004; Phan et al., 2002) were applied to the preprocessed data to extract BOLD signals as input to separate classifiers. *Z*-normalization was performed on the BOLD values at each voxel across the whole time course to correct for the variance of BOLD signals of different runs and different participants. The normalized BOLD values from each ROI of each TR were then collected into an input vector. Data from 2 training sessions and 1 testing session were used in both the single-subject whole brain and ROI classifications. Linear SVM, with the regularization parameter $C=1$, based on freely available SVM library SVMLight (Joachims, 1999) was used to perform both single-subject and ROI classifications. The single-subject whole brain analysis was performed based on the E-maps obtained after training a separate SVM model from each participant's data. The ROI classification performance was evaluated through 12-fold cross validation (CV). In each fold, the data of 11 participants were used to train the classifier, and then the data of 1 remaining participant were used to test the classifier.

Affect scores and subjective ratings of emotion recall

Before the beginning of fMRI data collection, we asked each participant to fill out the Positive Affect Negative Affect Schedule (PANAS), which is a psychometric scale developed to measure the independent constructs of positive and negative affect both as states and traits (Watson et al., 1988). PANAS contains a list of 10 descriptors for positive scale: attentive, interested, alert, excited, enthusiastic, inspired, proud, determined, strong and active; and 10 descriptors for negative scale: distressed, upset-distressed; hostile, irritable-angry; scared, afraid-fearful; ashamed, guilty; nervous, and jittery. The PANAS has been found to exhibit the following characteristics: (1) a significant level of stability in every time frame; (2) no consistent sex differences; (3) inter-correlations and internal consistency reliabilities are all acceptably high (ranging from 0.86 to 0.90 for PA and 0.84–0.87 for NA); (4) the reliability of the scales is unaffected by the time instructions used; and (5) has high scale validity and high item validity.

At the end of each SVM training or testing run, participants were instructed to rate their degree of success in being able to recall each type of emotion (i.e., happiness, disgust or sadness) for each block of the run on a scale of 1–9, where 1 represented the worst and 9 represented the best regulation (see Table 1).

Results

PANAS scores and subjective reports of emotion imagery

Participants used several emotion recall scenarios and imagery strategies for inducing themselves into happy, disgust and sad (see Table 1).

A negative correlation was found between the mean of the negative subscale of the PANAS and participants' ratings of their performance in emotion imagery tasks of session 3 ($r_s=-.74$, $p<.01$) and 4 ($r_s=-.72$, $p<.01$). A similar trend was also found for session 2 ($r_s=-.56$, $p=.07$, two-tailed). No significant correlations were found between positive subscale of the PANAS and the participants' ratings of their performance across the sessions of emotion imagery task. Also, no significant correlations were found between the prediction accuracy either with the PANAS scores, or with participants' ratings of their performance in emotion imagery.

In addition, we collected subjective reports of the intensity of the emotion attained, attention to the imagery task and the consistency of the mental strategy within each session, in a 9-point scale (1—worst, 9—best). Due to lack of scanning time, these data could be collected only in a subset of participants: 4 participants in the binary classification and 4 participants in the multiclass classification (see Supplementary Table 2).

Real-time binary and multiclass classification

Fig. 4 shows the prediction accuracies of the online classifier for all participants for SVM testing runs, with and without feedback, for binary classification (see Figs. 4a and b) and for multiclass classification (see Fig. 4c). Fig. 4a shows the performance of the binary classifier when participants used the same type of emotion imagery both during SVM training and testing runs (e.g., a participant used imageries of family reunion for happy emotion and eating insects for disgust emotion, respectively). Fig. 4b shows the performance of the binary classifier when participants used different types of emotion imagery during SVM training and testing runs (e.g. a participant used imageries of romantic love for happy emotion and baby with feces for disgust emotion for SVM training; and family reunion for happy emotion and dirty bathroom for disgust emotion for SVM testing). The online classifier showed the following classification accuracies across participants, indicated by mean and standard deviation of correct classification: (1) for binary classification (chance accuracy = 50%) testing with same emotion imagery, without feedback (92% ± 6%) and with feedback (80% ± 13%) (see Fig. 4a); (2) for binary classification testing with different emotion imagery, without feedback (80% ± 10%) and with feedback (65% ± 18%) (see Fig. 4b); and (3) for multiclass classification (chance accuracy = 33%) testing without

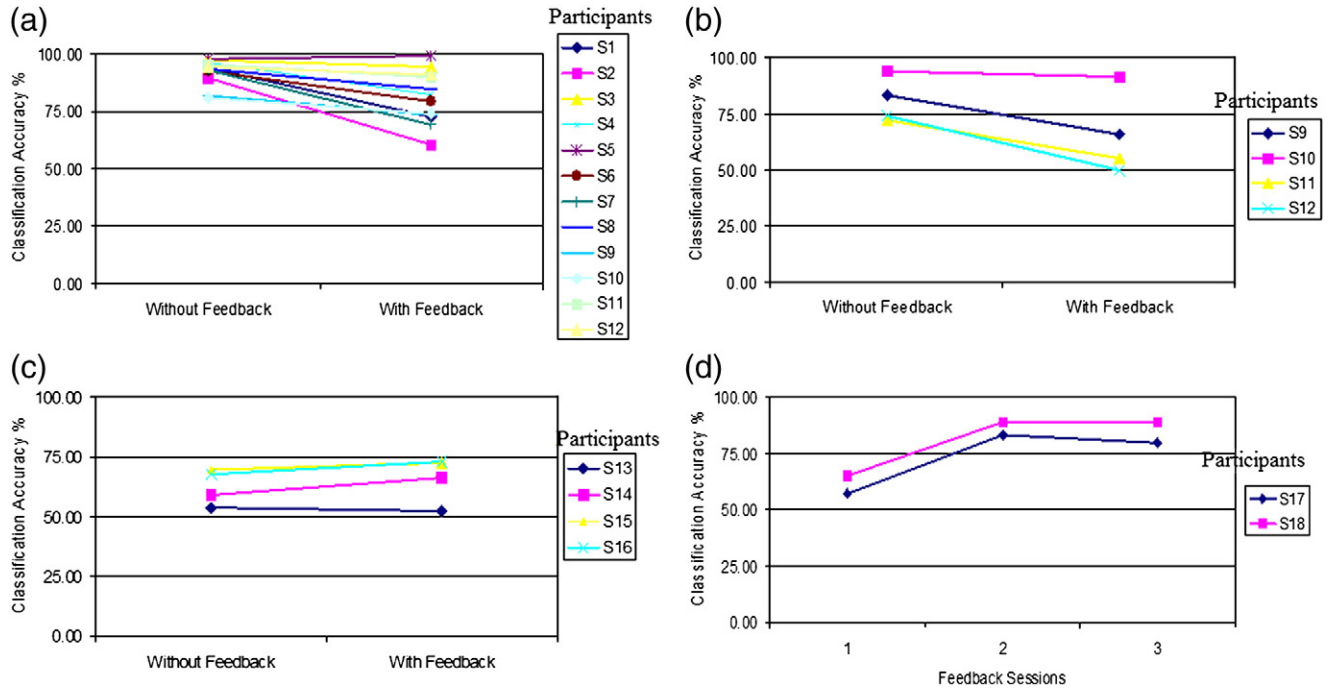

Fig. 4. Online SVM classification performance. (a) Binary classification (happy vs. disgust) accuracies (chance level = 50%) in 12 participants who used the same emotion imageries during both SVM training and testing sessions. (b) Binary classification accuracies in 4 participants who used different emotion imageries between SVM training and testing sessions. (c) Multiclass classification (happy vs. disgust vs. sad) accuracies (chance level = 33%) in 4 participants who used same emotion imageries in SVM training and testing sessions. (d) Classification accuracies (chance level = 50%) of extended feedback training on 2 participants on two-class (happy vs. disgust) classification.

feedback (62% ± 14%) and with feedback (60% ± 16%). Classification performance indicates the beneficial effects of online realignment, feature selection and optimized experimental procedures. Robust prediction accuracies were seen even when participants were intentionally instructed to use different strategies between SVM training and testing runs, indicating the ability of the classifier to generalize across varied emotion imagery strategies and memory recall scenarios.

Significantly higher accuracies of the SVM classifier were found in the sessions without feedback (median = 88.5) compared to the sessions with feedback (median = 76), $z = -2.02$, $p < .05$, $r = -.37$ (Wilcoxon signed-rank test). To assess the effect of extended feedback training on participants' performance, we trained 2 new participants for 3 sessions of emotion recall in the presence of feedback and incremental retraining of the classifier. Our results show that classification performance improves already in the second feedback session and classification accuracy continues to be maintained in the third session (see Fig. 4d). However, by the end of 5 training sessions, participants were unable to continue due to tiredness and hence feedback training had to be stopped. Nevertheless, our data suggests that feedback training with incrementally retrained classifier could enhance participant's performance.

SVM outputs for binary and multiclass classification

Supplementary Figs. 3a and b show the online SVM output for a participant before and after bias removal. Examination of the SVM outputs across different participants for multiclass classification showed robust classifier performance on a block-by-block basis within each testing session (see Supplementary Fig. 3c).

Effect maps and ROI classification performance

Extensive results of the single-subject whole brain classification and Effect maps generated thereof (in an offline analyses performed after the real-time experiments to reveal the discriminating brain regions) have been documented in the Supplementary information (Supplementary Sections 2.2 and 2.3, including supplementary Table 3 (a–k), Supplementary Fig. 1 (a–k) and Table 4 (a–c)). The Supplementary Tables also provide, for each brain region, the effect value (EV) and the cluster size, which together indicate the degree of importance of the cluster in discriminating between the emotions. Much inter-subject variability was observed in brain activations, possibly due to the diverse emotion imageries employed by the participants (see Table 1 for list of emotion imageries). The following brain regions were also observed as commonly activated among the participants: orbital medial frontal cortex (oMFC), anterior rostral medial frontal cortex (arMFC) and posterior rostral medial frontal cortex (prMFC) (these anatomical segregation were based on Amodio and Frith (2006)); superior and lateral frontal cortices, anterior cingulate cortex (ACC), insula, superior temporal gyrus (STG), primary and association areas of the visual cortex, posterior cingulate cortex (PCC) and precuneus. In terms of Brodmann areas, the following areas were commonly observed: BA10, 11, 14 and 25 belonging to the oMFC, BA32 to arMFC, BA8, 9 to the prMFC; and BA17, 18 and 19 belonging to the visual cortex. Fig. 5 shows exemplary Effect maps depicting the 3 contrasts of (happy vs. disgust), (happy vs. sad) and (disgust vs. sad) emotions.

To better identify the most important brain regions for emotion imagery and regulation, we performed offline ROI classifications, based on fMRI signals extracted from 24 different brain masks (see Supplementary Fig. 2) previously implicated in emotion processing (Amodio and Frith, 2006; Ochsner et al., 2005, 2004; Phan et al., 2002). This analysis was conducted separately on 12 subjects who had performed happy and disgust imagery (see Fig. 3, Table 1). Our results (see Fig. 5d) reveal that high classification performance (>75%) was obtained in the following brain regions: middle frontal gyrus (MiFG), superior frontal gyrus (SFG), ventrolateral prefrontal cortex (VLPFC), STG and precuneus. Other ROIs that showed high prediction rates included dorsolateral prefrontal cortex (DLPFC), ACC and the insula.

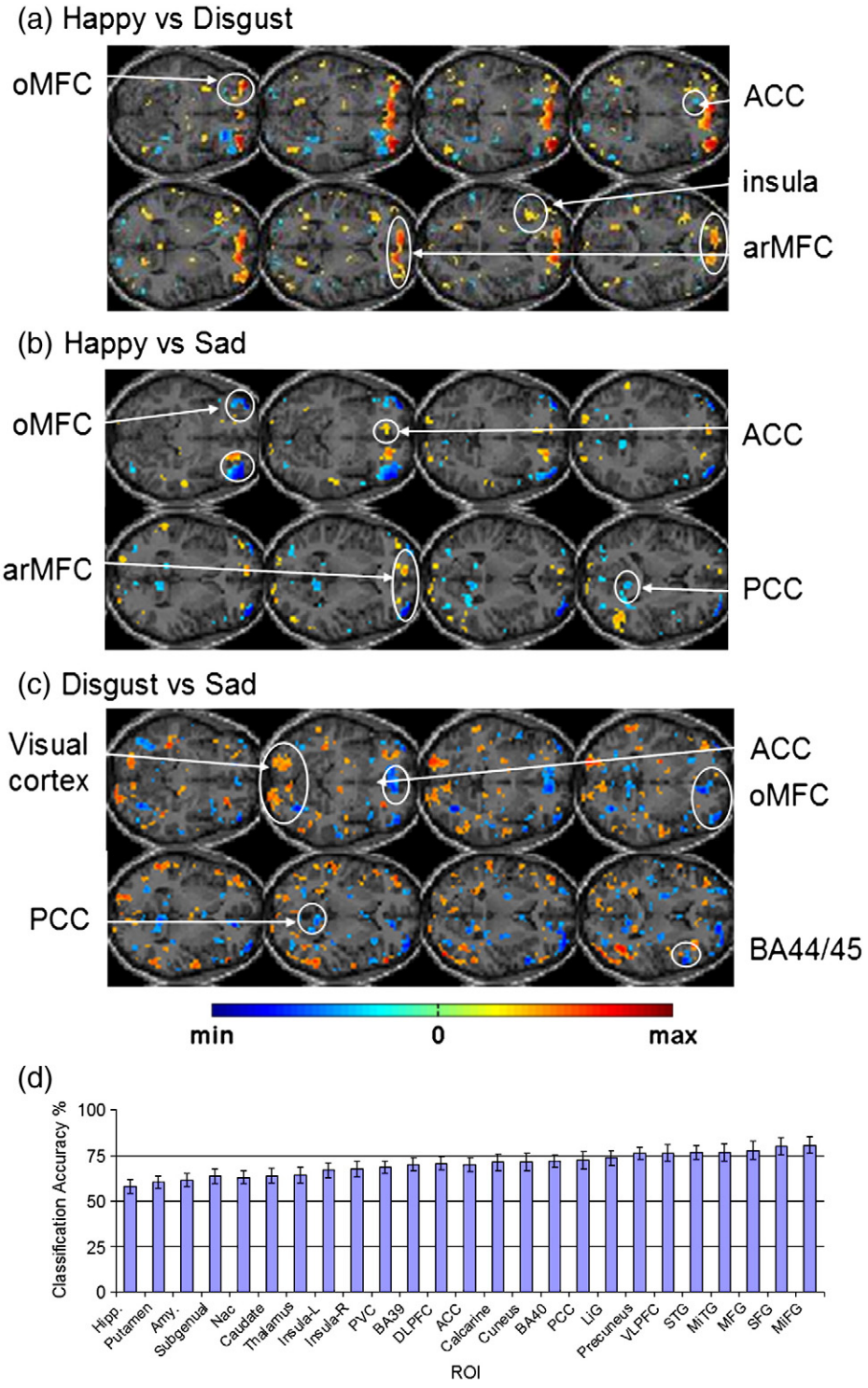

Fig. 5. Effect maps generated from single-subject whole brain SVM classification showing discriminating voxels for: (a) happy vs. disgust classification (for participant S1), (b) happy vs. sad classification and (c) disgust vs. sad classification (for participant S13). Second level threshold applied was 0.5, slices shown are in the range, Z(MNI coordinates) = −11 to +10 in steps of 3 units. Brain regions: oMFC—orbital medial frontal cortex, arMFC—anterior rostral MFC (based on Amodio and Frith, 2006); OFC—orbitofrontal cortex, ACC—anterior cingulate cortex, PCC—posterior cingulate cortex. (d) Prediction accuracy of 24 regions of interest (ROIs) in a single-subject ROI classification from 12 participants performing happy and disgust motor imagery. ROIs: Subgenual—subgenual cingulate (BA25), Nac—nucleus accumbens, PVC—primary visual cortex, DLPFC—dorsolateral prefrontal cortex, LiG—lingual gyrus, VLPFC—ventrolateral prefrontal cortex, STG—superior temporal gyrus, MiTG—middle temporal gyrus, MFG—medial frontal gyrus, SFG—superior frontal gyrus, MiFG—middle frontal gyrus.

Discussion

The present implementation of brain state decoding and fMRI-BCI is based on real-time SVM classification of fMRI signals after online head motion correction, spatial smoothing and feature selection based on our new approach called Effect mapping (EM) (Lee et al., 2010). EM is a method of generating functional activations by multiplying two types of information at each voxel of the brain: (1) SVM trained weight value at the voxel, and (2) and the mutual information (MI) between the BOLD value at the voxel and the SVM output. Activation

maps so derived represent a hybrid of univariate and multivariate functional mapping, resulting at once in spatially distributed activations as well as in large clusters centered about brain regions presumably important for discriminating between the brain states. In the present study, EM was used as the basis for feature selection to reduce the multidimensionality of fMRI data, and to choose the most discriminating clusters of voxels for classification. Our results show that the application of feature selection results in consistently good predictions of emotions for binary classification as well as for multiclass classification.

Subjective reports collected from participants ascertained that participants performed emotion imagery and that the online classifier decoded emotions and not arbitrary states of the brain. This study also shows a relationship between the participants' affect scores as measured by PANAS and the subjective ratings of their performance in the emotion imagery task, indicating that participants who report greater negative affect rate themselves relatively lower on the imagery tasks.

One may expect that higher subjective reports of success in recalling emotions should correlate with higher prediction accuracies. However, one should note that pattern classification works well and shows higher accuracy based on how similar the samples in the test set are in comparison to the training set. The greater the similarity, the greater is the accuracy. However, if the test set has different brain signals generated due to higher success in recall than in the training set, then the similarity of patterns is reduced and so is the classification accuracy. This means that improved success in recalling emotions may not result in greater decoding accuracy. This is in contrast to the situation with single ROI real-time feedback experiments where improved success in recall may correlate with higher % BOLD in the target region of interest.

This may then lead one to ask the question: How then can a pattern classifier provide contingent feedback in a self-regulation application where participants, including patients, are normally expected to improve their ability (with more training) to induce themselves into desired brain states? One way to approach this would be by adapting the classifier to participant's learning by incrementally retraining the classifier with new samples at the end of each session of feedback training. Another way could be to carry out feedback training of patients or beginners using a brain state classifier which has been previously trained on data from experts or individuals who are already good at self-regulation. In this approach, the classifier would show improved performance as the trainees achieve brain states similar to the experts. In such a classifier, one can expect improved prediction accuracy with greater success in recall or regulation. In any method, the experimenter has the prerogative to ascertain whether the participant is performing self-regulation as desired and whether the relevant brain regions are activated by scrutinizing the multivariate functional activations.

In our experiments, the performance of the classifier in the presence of feedback was lower than in its absence although a few participants performed better in the presence of feedback than in its absence. This lower accuracy of the classifier when the feedback was introduced could be due to as yet unlearned ability of the participants to perform consistent emotion imagery while simultaneously monitoring the thermometer. These data are in line with self-reports by participants after scanning sessions, about the difficulties they had in performing emotion regulation when feedback was introduced for the first time. It has been known that the introduction of feedback initially creates a task conflict between attention to the imagery task and the visual feedback (Rota et al., 2008), which could be overcome with longer training (Caria et al., 2007; deCharms, 2008; Rota et al., 2008). A second reason for the reduction in classification performance could be the extra activation of the visual cortex due to the graphical animation of the thermometer that was previously absent during the classifier training. This additional BOLD activation could act as noise or confound to the classifier input, reducing its prediction accuracy. To overcome the above limitations and to assess the effect of extended feedback training on participants' performance, we recruited two more participants and trained them on two-class (happy vs. disgust) classification, with 3 sessions of feedback training by incorporating an incremental method of classifier retraining. Our results show that participants did learn to improve their performance in the presence of feedback. An elaborate study on the effects of feedback training is beyond the scope of this work. Future studies could be dedicated to the systematic evaluation of the effect of feedback training on self-regulation performance, behaviour and changes in relevant neural network.

We performed additional offline pattern classification and multivariate statistical mapping based on the EM (Lee et al., 2010) approach to identify brain regions for discriminating emotion under consideration. Our results are in line with previous findings. Firstly, the most observed involvement of the medial frontal cortex in our clusters of activation and Effect maps is of interest. A theoretical review of the involvement of medial frontal cortex (MFC) by Amodio and Frith (2006) delineates the functional roles of the 3 regions of MFC, namely, arMFC, prMFC and oMFC. The authors associated the activation of the arMFC (the most anterior part of the frontal cortex including BA10, BA32 and rostral ACC) with the monitoring of one's own emotional state, such as rating one's emotions in response to pictures of varying valence. The more posterior region, the rostral MFC (prMFC), including BA8, BA9 and dorsal ACC, is activated by cognitive tasks such as action monitoring and attention. The orbital MFC (oMFC), including BA10, BA14 and BA25, was linked to the monitoring of task outcomes associated with punishments or rewards. Activations of the prefrontal cortex, ACC and insula found in our analyses are in line with an influential meta-analysis by (Phan and colleagues (2002) which states that the MFC has a general role in emotional processing and that induction by emotion imagery recruits the anterior cingulate and insula. Our offline analyses serve to confirm that participants performed emotion imagery and that the SVM results are reliable. Additionally, to our knowledge, our study provides the first objective comparison, through pattern classification, of the degree of involvement of different brain regions (see Fig. 5 and Supplementary Table 4) in emotion imagery and regulation.

In our real-time implementation of the SVM classifier, we programmed the online SVM classifier and the BCI system in a combination of C programs and Matlab scripts on a dedicated workstation. This design decision was beneficial in the following ways: (1) it allowed us greater flexibility in the repeated software cycle of source code modification and testing, (2) it helped us avoid interfering with the normal operation of the scanner and (3) it resulted in faster SVM training and online SVM online classification. This modular approach to software design has an added advantage in that when future versions of the scanner operating software (e.g., *Syngo* version VB15) are implemented, it is no longer necessary for the user to modify, recompile and relink the SVM classification source code.

Our decision to use was SVM based on recent developments in the pattern classification of neuroimaging data (LaConte et al., 2003, 2005; Martinez-Ramon et al., 2006; Shaw et al., 2003; Strother et al., 2004). SVM has been shown to have certain advantages in the classification of fMRI signals in comparison to other methods such as linear discriminant analysis (LaConte et al., 2005) and multilayer neural network. SVM is less sensitive to preprocessing (LaConte et al., 2005), better capable of handling large data sizes (Haynes and Rees, 2006), techniques developed to perform multiclass classification and most importantly produces unique optimal solutions (Collobert et al., 2006). Although, SVM model training is computationally more intensive, current availability of faster yet cheaper processors more than compensate for this drawback (LaConte et al., 2007).

Online multiclass brain state decoding has potential applications in lie detection (critical lies, non-critical lies and truths), BCIs (detecting

multiple movement states and thoughts) for communication and control, to name a few. The majority of the present implementations of brain state decoding including real-time classification, however, have two main drawbacks. Firstly, the present methods work based on spatial pattern of brain activity alone as input in discriminating different states of the brain and ignore the temporal pattern or time evolution of brain activity. Considering that brain states may differ in their temporal signature, in terms of time evolution of activity in one region and also in terms of its dynamic interaction with other regions of the brain, spatial classification of brain activity alone is largely limiting. For example, emotion regulation involves the dynamic, temporal evolution of BOLD activity connecting different parts of the limbic region that includes insula, amygdala, anterior cingulated cortex and medial prefrontal cortex. Hence, a pattern classification system that uses both spatial and temporal information but does so in a computationally efficient way so that it could be used in real-time would be of great practical importance and should be the focus of future research.

Another limitation of our present work and possible topic of future research is the lack of an implementation for real-time, subject-independent classification of brain states. Existing methods are optimized for each participant because SVM training is carried out on subject-specific data. However, these methods have two major disadvantages: (1) one needs to collect initial data for classifier training that consumes time and participant's attention and involvement, and (2) this approach is not suitable for certain applications such as real-time lie detection for forensics and security on persons from whom data for SVM training might not be available. An adaptive classifier needs to be developed similar to those previously developed in the pattern recognition field for subject-independent speech and character recognition. Such a subject-independent brain state classifier could then be applied to new subject data without prior classifier training, and in addition, could be adapted to the idiosyncrasies of every individual's brain size, shape and activation patterns. An essential technical improvisation to be achieved in this regard is the real-time coregistration and normalization of functional images to a standard brain so that inter-subject variations in brain size, shape and activations could be overcome. We anticipate that a subject-independent classifier could find use in clinical rehabilitation, where patients with brain abnormalities pertaining to motor, cognitive or emotion processing could be retrained to achieve normal level of functioning by providing feedback from a real-time pattern classifier that is trained on healthy subjects. By repeated operant training with contingent reward from the classifier, patients could perhaps learn to mimic brain activation of healthy individuals in order to ameliorate their problem.

Acknowledgments

This work was supported by grants from the SFB 437 “Kriegserfahrungen”, DFG BI 195/59-1 and DFG BI 195/56-1 of the Deutsche Forschungsgemeinschaft (DFG) and DFG-grant KU 1453/3-1. Author S. L. is grateful to DAAD (German Academic Exchange Service) for supporting this research. Author S.R is thankful to the Faculty of Medicine, Pontificia Universidad Católica de Chile, and Conycit for supporting this research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.neuroimage.2010.08.007.

References

Amodio, D.M., Frith, C.D., 2006. Meeting of minds: the medial frontal cortex and social cognition. Nat. Rev. Neurosci. 7, 268–277.
Ashburner, J., Friston, K., Penny, W., 2003. Human Brain Function, 2nd ed. .
Caria, A., Sitaram, R., Veit, R., Gaber, T., Kübler, A., Birbaumer, N., 2006. Can we learn to increase our emotional involvement? Real-Time fMRI of Anterior Cingulate Cortex During Emotional Processing. Organization for Human Brain Mapping. Florence, Italy.
Caria, A., Veit, R., Sitaram, R., Lotze, M., Weiskopf, N., Grodd, W., Birbaumer, N., 2007. Regulation of anterior insular cortex activity using real-time fMRI. Neuroimage 35, 1238–1246.
Collobert, R., Sinz, F., Weston, J., Bottou, L., 2006. Trading convexity for scalability. Proceedings of the 23rd international conference on Machine learning. ACM New York. NY, USA.
Cover, T.M., Thomas, J.A., 1991. Elements of information theory. Wiley, New York.
Davatzikos, C., Ruparel, K., Fan, Y., Shen, D.G., Acharyya, M., Loughead, J.W., Gur, R.C., Langleben, D.D., 2005. Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. Neuroimage 28, 663–668.
deCharms, R.C., 2008. Applications of real-time fMRI. Nat. Rev. Neurosci. 9, 720–729.
deCharms, R.C., Christoff, K., Glover, G.H., Pauly, J.M., Whitfield, S., Gabrieli, J.D., 2004. Learned regulation of spatially localized brain activation using real-time fMRI. Neuroimage 21, 436–443.
deCharms, R.C., Maeda, F., Glover, G.H., Ludlow, D., Pauly, J.M., Soneji, D., Gabrieli, J.D., Mackey, S.C., 2005. Control over brain activation and pain learned by using real-time functional MRI. Proc. Natl. Acad. Sci. U. S. A. 102, 18626–18631 (Epub 12005 Dec 18613).
Haynes, J.D., Rees, G., 2006. Decoding mental states from brain activity in humans. Nat. Rev. Neurosci. 7, 523–534.
Haynes, J.D., Sakai, K., Rees, G., Gilbert, S., Frith, C., Passingham, R.E., 2007. Reading hidden intentions in the human brain. Curr. Biol. 17, 323–328.
Joachims, T., 1999. Making large-scale SVM learning practical. In: Schölkopf, B., Burges, C., Smola, A. (Eds.), Advances in Kernel Methods—Support Vector Learning. MIT-Press.
Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. Nat. Neurosci. 8, 679–685.
LaConte, S., Anderson, J., Muley, S., Ashe, J., Frutiger, S., Rehm, K., Hansen, L.K., Yacoub, E., Hu, X., Rottenberg, D., Strother, S., 2003. The evaluation of preprocessing choices in single-subject BOLD fMRI using NPAIRS performance metrics. Neuroimage 18, 10–27.
LaConte, S., Strother, S., Cherkassky, V., Anderson, J., Hu, X., 2005. Support vector machines for temporal classification of block design fMRI data. Neuroimage 26, 317–329.
LaConte, S.M., Peltier, S.J., Hu, X.P., 2007. Real-time fMRI using brain-state classification. Hum. Brain Mapp. 28, 1033–1044.
Lang, P.J., Bradley, M.M., Cuthbert, B.N., 1997. International Affective Picture System (IAPS): Technical Manual and Affective Ratings. The Center for Research in Psychophysiology, University of Florida, Gainesville.
Lee, S., Halder, S., Kubler, A., Birbaumer, N., Sitaram, R., 2010. Effective functional mapping of fMRI data with support-vector machines. Hum. Brain Mapp. (Epub ahead of print).
Martinez-Ramon, M., Koltchinskii, V., Heileman, G.L., Posse, S., 2006. fMRI pattern classification using neuroanatomically constrained boosting. Neuroimage 7, 7.
Mourao-Miranda, J., Bokde, A.L., Born, C., Hampel, H., Stetter, M., 2005. Classifying brain states and determining the discriminating activation patterns: support vector machine on functional MRI data. Neuroimage 28, 980–995 (Epub 2005 Nov 2004).
Mourao-Miranda, J., Friston, K.J., Brammer, M., 2007. Dynamic discrimination analysis: a spatial-temporal SVM. Neuroimage 36, 88–99.
Mourao-Miranda, J., Reynaud, E., McGlone, F., Calvert, G., Brammer, M., 2006. The impact of temporal compression and space selection on SVM analysis of single-subject and multi-subject fMRI data. Neuroimage 33, 1055–1065.
Ochsner, K.N., Beer, J.S., Robertson, E.R., Cooper, J.C., Gabrieli, J.D., Kihsltrom, J.F., D'Esposito, M., 2005. The neural correlates of direct and reflected self-knowledge. Neuroimage 28, 797–814.
Ochsner, K.N., Ray, R.D., Cooper, J.C., Robertson, E.R., Chopra, S., Gabrieli, J.D., Gross, J.J., 2004. For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. Neuroimage 23, 483–499.
Pessoa, L., Padmala, S., 2005. Quantitative prediction of perceptual decisions during near-threshold fear detection. Proc. Natl. Acad. Sci. U. S. A. 102, 5612–5617.
Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I., 2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. Neuroimage 16, 331–348.
Polyn, S.M., Natu, V.S., Cohen, J.D., Norman, K.A., 2005. Category-specific cortical activity precedes retrieval during memory search. Science 310, 1963–1966.
Posse, S., Fitzgerald, D., Gao, K., Habel, U., Rosenberg, D., Moore, G.J., Schneider, F., 2003. Real-time fMRI of temporolimbic regions detects amygdala activation during single-trial self-induced sadness. Neuroimage 18, 760–768.
Rota, G., Sitaram, R., Veit, R., Erb, M., Weiskopf, N., Dogil, G., Birbaumer, N., 2008. Self-regulation of regional cortical activity using real-time fMRI: the right inferior frontal gyrus and linguistic processing. Hum. Brain Mapp. 30 (5), 1605–1614.
Rota, G., Sitaram, R., Veit, R., Weiskopf, N., Birbaumer, N., Dogil, G., 2006. fMRI-neurofeedback for operant conditioning and neural plasticity investigation: a study on the physiological self-induced regulation of the BA 45. Proceedings of the Cognitive Neuroscience conference, San Francisco.
Ruiz, S., Sitaram, R., Lee, S., Caria, A., Soekadar, S., Veit, R., Birbaumer, N., 2008. Learned control of insular activity using fMRI brain computer interface in schizophrenia. 1st Schizophrenia International Research Society Conference, Venice, Italy.
Shaw, M.E., Strother, S.C., Gavrilescu, M., Podzebenko, K., Waites, A., Watson, J., Anderson, J., Jackson, G., Egan, G., 2003. Evaluating subject specific preprocessing

choices in multisubject fMRI data sets using data-driven performance metrics. Neuroimage 19, 988–1001.
Sitaram, R., Caria, A., Veit, R., Gaber, T., Kuebler, A., Birbaumer, N., 2005. Real-time fMRI based brain–computer Interface enhanced by interactive virtual worlds. 45th Annual Meeting Society for Psychophysiological Research, Lisbon, Portugal.
Soon, C.S., Brass, M., Heinze, H.J., Haynes, J.D., 2008. Unconscious determinants of free decisions in the human brain. Nat. Neurosci. 11, 543–545.
Strother, S., La Conte, S., Kai Hansen, L., Anderson, J., Zhang, J., Pulapura, S., Rottenberg, D., 2004. Optimizing the fMRI data-processing pipeline using prediction and reproducibility performance metrics: I. A preliminary group analysis. Neuroimage 23 (Suppl 1), S196–207.
Veit, R., Lotze, M., Caria, A., Gaber, T., Sitaram, R., Birbaumer, N., 2006. Real-time fMRI of human anterior insula during emotional processing. Human Brain Mapping (Florenece, Italy).
Watson, D., Clark, L.A., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. J. Pers. Soc. Psychol. 54, 1063–1070.
Weiskopf, N., Scharnowski, F., Veit, R., Goebel, R., Birbaumer, N., Mathiak, K., 2004. Self-regulation of local brain activity using real-time functional magnetic resonance imaging (fMRI). J. Physiol. Paris 98, 357–373.
Weiskopf, N., Veit, R., Erb, M., Mathiak, K., Grodd, W., Goebel, R., Birbaumer, N., 2003. Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fMRI): methodology and exemplary data. Neuroimage 19, 577–586.
Yoo, S.S., Fairneny, T., Chen, N.K., Choo, S.E., Panych, L.P., Park, H., Lee, S.Y., Jolesz, F.A., 2004. Brain–computer interface using fMRI: spatial navigation by thoughts. Neuroreport 15, 1591–1595.
Yoo, S.S., Jolesz, F.A., 2002. Functional MRI for neurofeedback: feasibility study on a hand motor task. Neuroreport 13, 1377–1381.