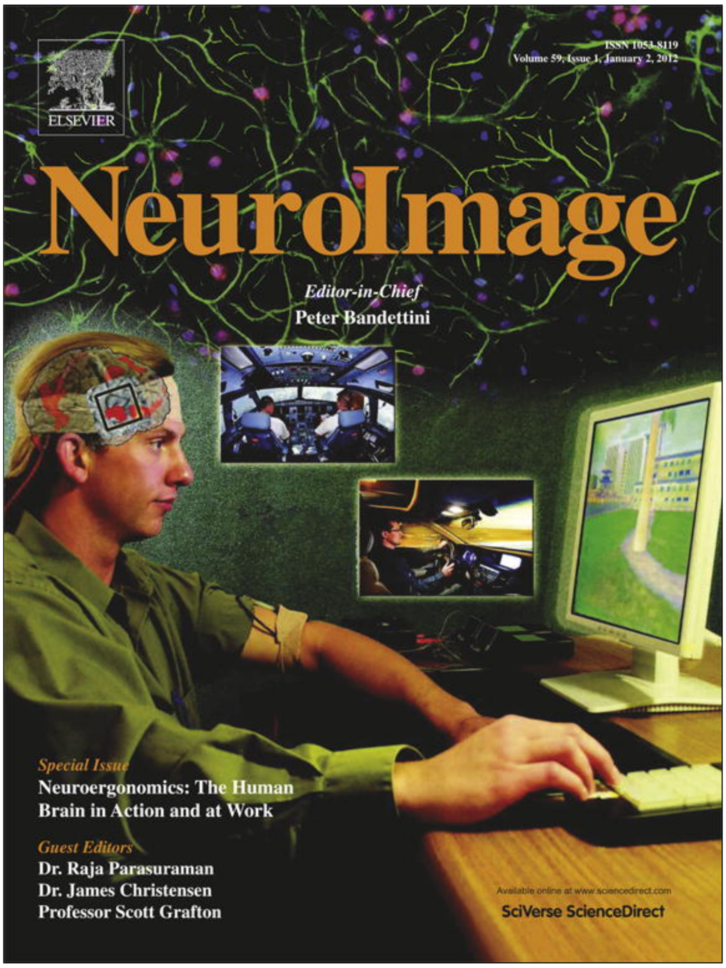
ISSN 1053-8119
Volume 59, Issue 1, January 2, 2012
ELSEVIER
NeuroImage
Editor-in-Chief
Peter Bandettini
Special Issue
Neuroergonomics: The Human Brain in Action and at Work
Guest Editors
Dr. Raja Parasuraman
Dr. James Christensen
Professor Scott Grafton
Available online at www.sciencedirect.com
SciVerse ScienceDirect

NeuroImage 59 (2012) 718–727

Contents lists available at ScienceDirect

NeuroImage

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

Decoding the neural representation of affective states

Laura B. Baucom, Douglas H. Wedell, Jing Wang, David N. Blitzer, Svetlana V. Shinkareva *

University of South Carolina, Columbia, SC, USA

ARTICLE INFO

Article history:
Received 30 December 2010
Revised 21 June 2011
Accepted 12 July 2011
Available online 23 July 2011

Keywords:
Affective states
Valence
Arousal
Multi-voxel pattern analysis
INDSCAL

ABSTRACT

Brain activity was monitored while participants viewed picture sets that reflected high or low levels of arousal and positive, neutral, or negative valence. Pictures within a set were presented rapidly in an incidental viewing task while fMRI data were collected. The primary purpose of the study was to determine if multi-voxel pattern analysis could be used within and between participants to predict valence, arousal and combined affective states elicited by pictures based on distributed patterns of whole brain activity. A secondary purpose was to determine if distributed patterns of whole brain activity can be used to derive a lower dimensional representation of affective states consistent with behavioral data. Results demonstrated above chance prediction of valence, arousal and affective states that was robust across a wide range of number of voxels used in prediction. Additionally, individual differences multidimensional scaling based on fMRI data clearly separated valence and arousal levels and was consistent with a circumplex model of affective states.

Introduction

The representation and processing of emotional states in the brain has become a fundamental area of study within cognitive neuroscience. Two distinct approaches to understanding affective states have come to the forefront of the study of emotion. The categorical approach builds on the finding of a core set of distinct basic emotions, as demonstrated by studies of the perception of human facial emotional expressions and basic physiological responses of the autonomic nervous system to emotional stimuli (Ekman, 1992a,b). Moreover, these basic emotions, such as anger, fear, disgust, sadness and joy, are thought to be represented by different neural systems (Panksepp, 1992, 1998).

An alternative to the categorical approach is to consider the underlying structure of emotions as deriving from two or three basic dimensions of affective processing (Posner et al., 2005; Rolls, 1999; Schlosberg, 1954; Watson and Tellegen, 1985). A widely accepted dimensional model of affect, developed using multidimensional scaling techniques, conceptualizes the affective space as a circle or circumplex with two underlying primary dimensions: valence and arousal (Russell, 1980). Valence reflects the hedonic tone of the emotional state, ranging from positive to negative, while arousal, or activation, reflects the engagement of the organism, ranging from high to low (Roberts and Wedell, 1994). The circumplex model of affect suggests that all emotions or affective states can be distinguished in terms of varying levels of valence and arousal, with two distinct neural systems mediating the representation of affective states (Barrett, 1998).

As described above, both the categorical and dimensional approaches to understanding emotional states have support from behavioral and neuroimaging studies. One way to resolve this seeming contradiction is to assume that although emotional states can be described by dimensional variation along valence and arousal, further categorical processing of states may overlay this structure and result in activation of distinct cognitive and neural components. Thus, for example, anger and disgust may both be negative and high arousal states, but their categorical processing leads to different implications, as described in appraisal theory (Lazarus, 1991, 1995). Thus, while the methods we describe in the present study build on the circumplex model of affective states, we believe they may also be applied to categorical approaches.

Traditionally, neuroimaging studies have used univariate statistical parametric mapping methods to determine which areas of the brain subserve the processing of emotional stimuli and the generation of emotional states. In a meta-analysis of 162 neuroimaging studies of emotion, Kober et al. (2008) demonstrated that medial frontal areas are co-activated with core limbic structures and that the dorsomedial prefrontal cortex may underlie the generation of emotional states. Consistent with dimensional models of emotion, neuroimaging studies have demonstrated a dissociation of valence and arousal for various stimulus modalities, such as olfactory (Anderson et al., 2003), gustatory (Small et al., 2003), picture (Anders et al., 2004; Grimm et al., 2006; Nielen et al., 2009), word (Kensinger and Corkin, 2004; Lewis et al., 2007; Nielen et al., 2009; Posner et al., 2009), and face (Gerber et al., 2008), as well as emotional experiences induced by the presentation of evocative sentences (Colibazzi et al., 2010). The results of these studies suggest that valence and arousal may be represented in separate neural circuits containing the amygdala, insula, thalamus, dorsal anterior cingulate cortex, and prefrontal regions. These regions are generally consistent with the hypothesis that responses to valence are part of

* Corresponding author at: Department of Psychology, University of South Carolina, Columbia, SC 29208, USA. Fax: +1 803 777 9558.
E-mail address: shinkareva@sc.edu (S.V. Shinkareva).

1053-8119/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2011.07.037

motivational circuitry linked to mesolimbic structures (Lang and Bradley, 2010). Response in the amygdala has been associated with both arousal (Anderson et al., 2003; Colibazzi et al., 2010; Lewis et al., 2007; Small et al., 2003) and valence (Anders et al., 2004; Gerber et al., 2008; Posner et al., 2009), suggesting this region may belong to both valence and arousal networks. Traditionally, the amygdala is thought to be a part of a subcortical pathway devoted to the processing of emotional stimuli, but a more recent view suggests the amygdala modulates multiple networks by identifying and allocating resources to biologically-relevant stimuli (Pessoa and Adolphs, 2010).

As a complement to traditional univariate analyses, examining whole brain processing of states may prove valuable in identifying emotions from neuroimaging and in determining the representational structure of those affective states. The techniques entailed in multi-voxel pattern analysis (MVPA) are well-suited to the study of affect, as the pattern-based approach of MVPA detects cognitive states by jointly investigating information in multiple voxels and is more sensitive compared to univariate statistical parametric mapping (Haynes and Rees, 2006; Norman et al., 2006; O'Toole et al., 2007). Likewise, the results of traditional univariate analyses may be informative for MVPA techniques, as they suggest a core group of brain structures that may contribute to whole brain patterns of activity. Previous neuroimaging studies have investigated the representation of affect with pattern-based approaches, successfully decoding affective states from patterns of brain activity located in specific regions of interest as well as from patterns of whole brain activity. Peelen et al. (2010) used MVPA to investigate which brain regions encode emotions independently of the modality (e.g., body, face, voice) through which they are perceived. Other neuroimaging studies employing MVPA methods have investigated modality-specific encoding of discrete emotional states using emotionally-evocative voice recordings (Ethofer et al., 2009), facial expressions (Said et al., 2010), faces showing expressions of fear (Pessoa and Padmala, 2007), and recall of emotional situations (Sitaram et al., 2011).

The MVPA techniques may also prove useful when the affective state is subtly manipulated, as when incidental rather than intentional methods of inducing affective states are used. In intentional methods participants are typically told to think about the emotional consequences of a particular affective cue, such as a picture, video or story. As such, researchers may be analyzing neural imaging signals related to those voluntary thought processes rather than to the emotional state itself. In incidental methods the affective content of the stimuli is not explicitly processed. While a few studies have used incidental presentation of affective cue stimuli (Ethofer et al., 2009), these procedures may produce a weaker neural signal that reduces the power of univariate approaches. The MVPA approach is well suited to capturing subtle changes in neural processing distributed throughout the brain, and thus may be ideal for studying emotional responses in incidental exposure tasks.

The current study used an incidental affect inducement approach in which affectively scaled pictures were rapidly presented at a 200 ms rate, and the participant's task was simply to maintain focus on a central fixation point. Each picture set consisted of 20 photographs and represented one of five affective conditions resulting from the combination of low and high arousal with positive, neutral and negative valence. Our procedure is incidental in that participants were not told the nature of the study and were not asked to make any evaluations regarding the pictures. Prior scaling of these pictures (Lang et al., 2008) demonstrates that when people perceive each picture they have a reliable affective reaction that can be measured primarily along dimensions of valence and arousal. Our procedure does not disentangle the perception of affect from the experiencing of affect.

The purpose of our study was twofold. First, we explored whether MVPA methods could be used to identify affective states within each individual by decoding functional patterns of whole brain activity, thus extending previous MVPA studies of affect to specifically examining valence and arousal dimensions. Consistent with the circumplex model (Russell, 1980), we hypothesized that functional patterns of whole brain activity would contain information discriminating the states in terms of valence and arousal levels, as elicited by viewing the visual stimuli. Thus, we tested for classification of positive and negative valence, high and low arousal, and finally the four separate states. Our statistical approach within individuals was to train our classifiers on all but one trial of each type and cross-validate the pattern analysis on the remaining trials. We also used MVPA to predict affective states across individuals, by training the classifier on all but one participant and then predicting the emotion state of the excluded participant.

A second purpose of our study was to extract the internal representation of affective states elicited by viewing emotionally-evocative pictures from fMRI data and compare it with predictions from the circumplex model of affect. While our predictions for valence- and arousal-based classification are consistent with the circumplex model of affect, these predictions could also be accounted for by a categorical model that posits four distinct brain areas for the emotional states we induce. However, the categorical approach does not predict recovery of a circumplex relationship from similarity metrics derived from the fMRI data. We used individual differences multidimensional scaling (INDSCAL) to explore the lower dimensional representation of affective states from functional patterns of whole brain activity. The internal representation of affect has been previously shown to separate on the dimension of valence with electroencephalogram (EEG) data (Onton and Makeig, 2009). In this work we extended these findings by demonstrating that the internal representation of affect derived from fMRI data can be separated on both dimensions of valence and arousal, providing additional support for the circumplex model of affect. Thus unlike studies that focus on specific regions of interest, our focus was to utilize the distributed representations elicited by affective stimuli to capture the underlying affective states.

Method

Participants

Thirteen right-handed volunteer adults (12 females) from the University of South Carolina community with normal or corrected to normal vision participated and gave written informed consent in accordance with the Institutional Review Board at the University of South Carolina.

Materials

Participants viewed a series of color photographs obtained from a database provided by the National Institute of Mental Health Center for the Study of Emotion and Attention. Pictures were selected based on normed valence and arousal ratings from the International Affective Picture System (IAPS), which provides ratings of emotionally-evocative color photos that contain subject matter from a variety of semantic categories. The photographs were rated along the dimensions of valence (ranging from unpleasant to pleasant), arousal (ranging from calm to excited), and dominance (ranging from in control to dominated). Ratings were obtained using the 9-point Self-assessment Manikin (SAM) and based on normative studies involving both adult and child populations (Lang et al., 2008).

Stimuli were matched on hue, saturation, and intensity values in MATLAB by setting the values for each photograph to the overall mean for all photos. Five stimulus sets were constructed that consisted of 20 pictures each and that varied in valence and arousal: high arousal negative valence (HN), low arousal negative valence (LN), low arousal neutral valence (L0), low arousal positive valence (LP), and high arousal positive valence (HP). We did not include a high arousal

Table 1
Picture sets.

	Sets				
	HN	LN	L0	LP	HP
IAPS identification numbers	2811, 3500 3530, 6230 6250.1, 6260 6300, 6312 6370, 6510 6540, 6821 8485, 9050 9250, 9252 9410, 9600 9635.1, 9921	1111, 1275 2278, 2490 2590, 2700 2715, 9000 9001, 9046 9220, 9331 9341, 9470 9471, 9520 9560, 9571 9911, 9912	2038, 2190 2200, 2221 2393, 2440 2480, 2580 2620, 2745.1 2840, 5390 5510, 5530 5731, 7140 7205, 7224 7700, 9210	1600, 1603 1604, 1670 2000, 2370 2501, 2655 5000, 5010 5030, 5200 5711, 5750 5760, 5800 5811, 5870 5891, 7545	5470, 5621 5626, 5629 8030, 8034 8080, 8161 8170, 8179 8180, 8185 8186, 8190 8200, 8300 8370, 8400 8490, 8499
Valence	M = 2.30 SD = 0.32	M = 2.91 SD = 0.55	M = 5.03 SD = 0.48	M = 6.92 SD = 0.46	M = 7.28 SD = 0.41
Arousal	M = 6.70 SD = 0.28	M = 4.62 SD = 0.63	M = 2.90 SD = 0.29	M = 3.22 SD = 0.45	M = 6.54 SD = 0.41
Dominance	M = 2.96 SD = 0.45	M = 4.24 SD = 0.59	M = 5.84 SD = 0.56	M = 6.40 SD = 0.66	M = 5.49 SD = 0.51
Pictures with faces	12	10	10	4	7
Pictures with human bodies	16	9	10	3	18
General description	Assault scenes, disasters scenes, weapons aimed at viewer	Dead animals, wrecked cars and buildings, sad or downcast people	People standing and sitting, objects, rooms, plants	Animals, smiling people, flowers, gardens, sky scenes	Sky diving, hang gliding, roller coasters, skiing, boating

Note: IAPS = International Affective Picture System; HN = High Arousal Negative Valence, LN = Low Arousal Negative Valence, L0 = Low Arousal Neutral Valence, LP = Low Arousal Positive Valence, HP = High Arousal Positive Affect; Statistics summarize 9-point IAPS ratings of pictures in each set.

neutral valence condition because it was difficult to identify these pictures in preliminary rating tasks. Pictures for the five sets were selected based on valence and arousal ratings, with the goal to maximize differences in valence for positive versus negative sets and maximize differences in arousal for low versus high arousal sets, while attempting to match levels on the shared dimensions across sets. Because arousal and valence values are not independent, precise matching was not possible. Table 1 presents a summary of the picture sets used in this study, including the IAPS identification numbers, summary statistics of the arousal, valence and dominance ratings for pictures within each set, number of pictures that included human faces and human bodies, and a short description of each set. Note that the positive sets differ strongly from the negative sets in mean valence ratings ($M_{Positive} = 7.10$ and $M_{Negative} = 2.61$), with the neutral set approximately at the midpoint of the valence scale. Similarly low arousal and high arousal sets differed strongly in mean arousal ratings ($M_{Low} = 3.58$ and $M_{High} = 6.67$). Whereas valence ratings were fairly closely matched for sets of the same valence, arousal ratings were somewhat more variable, especially for the low arousal sets. All sets included pictures of human faces and human bodies.[1]

In a behavioral study with a separate group of participants ($n = 15$), these five picture sets were shown to differ on the dimensions of valence and arousal. The behavioral study used the same picture sets and presentation rate used in the fMRI study, but participants were asked to rate their emotional states along one of nine dimensions after each sequence of pictures. The dimensions reflected the degree to which the participant reported feeling negative, angry, anxious, sad, calm, relaxed, excited, happy, and positive. The dimensions of valence and arousal were not rated directly but were derived from ratings along the other nine dimensions. A 5×5 correlation matrix of the five states across the nine ratings was constructed for each individual and INDSCAL was used to analyze the combined data (Borg and Groenen, 1997). A two-dimensional solution described the data well, with the results of the behavioral validation of standardized stimuli displayed in Fig. 1. As shown, the predicted circumplex relationship was obtained. The INDSCAL procedure also provides indicators of how well each participant fits the common model and the relative weighting of dimensions. Eight of the 15 participants fit the joint configuration with R^2 greater than 0.90, three with R^2 between 0.60 and 0.90, and four with R^2 between 0.37 and 0.60. Thus, while the majority of participants' data was well explained by the model, there were individual differences and some participants were not well described by the model.

To conduct statistical comparisons between the five picture sets, valence and arousal scores were derived for each participant. These scores were determined by two dimensional nonlinear MDS solutions

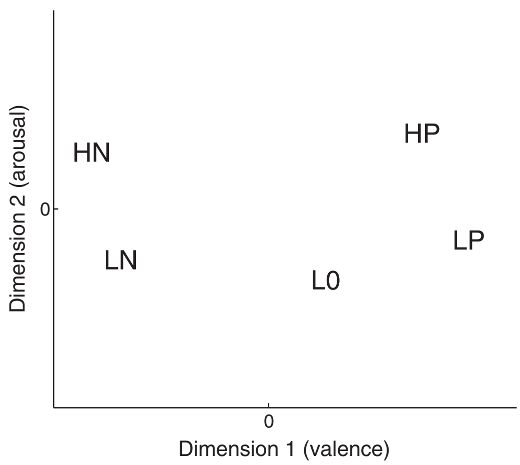

Fig. 1. Lower dimensional representation of affective space based on behavioral data. A two-dimensional solution described the data well, providing behavioral validation of the predicted circumplex relationship. LP — low arousal, positive valence; LN — low arousal, negative valence; L0 — low arousal, neutral valence; HP — high arousal, positive valence; HN — high arousal, negative valence.

[1] Overall the percentage of human faces in our sets (43%) was close to that in the IAPS database (approximately 48%). Pictures with faces and bodies were included in all conditions but were somewhat underrepresented in the LP condition, which was representative of the data base itself. Although inclusion of fewer of these features could be used to aid classification of this condition when decoding across the full set of emotion sets, it would not help in classifying into high and low arousal categories. Ultimately, features unique to sets (such as flowers in LP or guns in HN) may contribute to distinguishing sets, but these should be less helpful for the broader classification of arousal and valence. More generally, our inclusion of many pictures for each set, with very short exposure to each picture (200 ms), was designed to minimize classification based on specific features and maximize classification based on the emotional response to the pictures.

for each participant, transformed using a Procrustes rotation to the design matrix orientation to reflect degree of valence and arousal. A one-way repeated measures Analysis of Variance (ANOVA) conducted on the valence scores across the five sets revealed a large main effect of emotion set, $F(4,56) = 107.1$, $p < 0.001$, ($M_{HN} = -1.05$, $M_{LN} = -0.94$, $M_{L0} = 0.24$, $M_{LP} = 0.93$, and $M_{HP} = 0.82$). Bonferroni corrected paired comparisons indicated that all pairwise differences were significant (at $p < 0.05$), except for the differences between the two negative valence sets (HN and LN) and the two positive valence sets (HP and LP). A one-way repeated measures ANOVA conducted on the arousal scores across the five sets revealed a large main effect of emotion set, $F(4,56) = 14.7$, $p < 0.001$, ($M_{HN} = 0.49$, $M_{LN} = -0.14$, $M_{L0} = -0.24$, $M_{LP} = -0.34$, and $M_{HP} = 0.24$). Bonferroni corrected paired comparisons ($p < 0.05$) revealed no significant differences between sets within the same arousal level and significant differences between sets at different arousal levels, except for a nonsignificant difference between HP and L0 sets. The overall pattern of results provides additional validation of the manipulated differences in affect states for the five sets.

Experimental paradigm

Participants viewed sets of color photographs depicting various scenes designed to elicit a specific affective state (HN, LN, L0, LP, and HP). All images were 640 × 480 pixels and were presented in 32-bit color using E-prime software (Psychology Software Tools, Sharpsburg, PA).

Functional data were collected in one scanning session. For each of five experimental conditions there was one set of 20 photographs. There were 12 presentations of each set, with randomized order of individual photographs within each set. A fixation cross was superimposed on each photograph, and participants were instructed to maintain fixation throughout the scanning session. Each set of 20 photographs was presented at a rate of 200 ms per photograph, to minimize the effect of semantic content, and was followed by an 8 s rest period, during which participants were instructed to maintain fixation on a fixation cross in the center of the screen (Fig. 2). Thus the experiment consisted of 60 trials (5 conditions, 12 presentations), and each trial lasted 12 s.

fMRI procedure

Functional images were acquired on a Siemens Magnetom Trio 3.0T scanner (Siemens, Erlangen, Germany) at the McCausland Center for Brain Imaging at the University of South Carolina, using a gradient echo EPI pulse sequence with TR = 2198 ms, TE = 30 ms and a 90° flip angle. Thirty-six 3 mm thick oblique-axial slices were imaged with no gap. The acquisition matrix was 64 × 64 with 3 × 3 × 3 mm voxels.

fMRI data processing and analysis

Data processing and statistical analysis were performed using Statistical Parametric Mapping 8 software (Wellcome Department of Cognitive Neurology, London, UK). The data were corrected for slice timing, motion, and linear trend, and a high-pass filter was applied (0.008 Hz cut off). Images were spatially normalized to MNI space using a 12-parameter affine transformation. Only voxels common to all participants were included in the analysis. The data preprocessing steps and MVPA analysis employed in this work are similar to those that have been successfully used in other MVPA studies (Mitchell et al., 2008; Shinkareva et al., 2008, 2011). The percent signal change (PSC) relative to the average activity in a voxel was computed for each voxel in every volume. The mean PSC of two volumes, offset 4.4 s from the stimulus onset (to account for the delay in hemodynamic response), was used as the input for further analyses (Fig. 2). Furthermore, the mean PSC data for each voxel was standardized to have a mean of zero and variance of one.

Pattern classification methods

Classifiers were trained to identify cognitive states from the pattern of brain activity (mean PSC) elicited by viewing pictures from four affective categories. Two-category or four-category classification was performed to identify cognitive states associated with valence (*positive* and *negative*), arousal (*low* and *high*), or four affective states (HN, LN, LP and HP). For classification, classifiers were defined as a function f: $mean\_PSC \rightarrow Y_j$, $j = \{1, \ldots, k\}$, where k was the number of categories used for classification, Y_j were categories of valence, arousal or affect and where $mean\_PSC$ was a vector of mean PSC voxel activations.

Prior to classification, trials were divided into training and test sets, and relevant features (voxels) were extracted (see below for feature selection method) from the training set only. The classifier was constructed using the selected features from the training set. The classifier was applied subsequently to the unused test set and classification performance was evaluated with cross-validation.

Feature selection

To reduce the size of the data, only gray matter voxels with the most stable responses across multiple presentations of the same conditions were selected (Mitchell et al., 2008; Shinkareva et al., 2008). Voxel stability scores were computed by averaging pairwise correlation coefficients between vectors of presentations of all conditions in the training set, thus assigning higher scores to voxels with more consistent variation in activity across conditions. Voxels were then ordered within each individual's data set from highest to lowest stability. We explored the impact of retaining different numbers of voxels on each analysis, rather than deciding upon an arbitrary threshold.

Classification

A logistic regression (multinomial logistic regression for four-way classification) classifier was used for classification of affective states (Bishop, 2006). Logistic regression is a widely used classifier that learns the function f: $P(Y|X)$, where Y is discrete dependent variable, and X is a vector containing discrete or continuous variables. By using the maximum likelihood estimation, this algorithm estimates the probability of the given data belonging to an output category and classifies the data into the most probable category. As a classifier, logistic regression

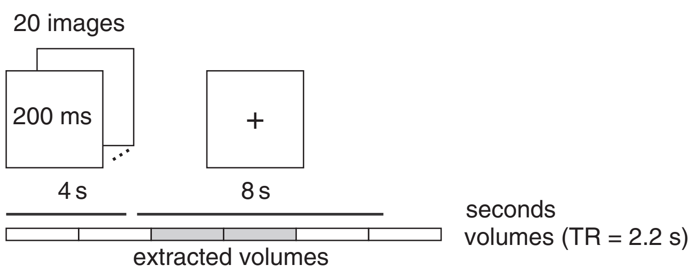

Fig. 2. A schematic representation of the presentation timing and data extraction for a single trial.

directly estimates its parameters from the training data. Twelve-fold cross-validation was used to evaluate classification performance, where each fold corresponded to one presentation of each of the four conditions. Thus, the classifier was trained on 11 presentations and tested on one presentation. For binary classification, trials of the same condition were averaged for the test set. Classification was repeated iteratively until each presentation served as the test set once. Classification accuracies were computed based on the average classification accuracy across test folds. As a result, classification accuracy was always based upon the test data only, which remained disconnected from the training data.

If classification is successful, accuracies should be significantly different from the chance level accuracy, i.e. the accuracy of guessing. The significance of classification accuracy was evaluated based on the binomial distribution $B(n, p)$, where n is the number of trials of each classification computation and p is the probability of correct classification when the exemplars are randomly labeled (Pereira et al., 2009).

Visualization of voxel locations

To investigate the location of voxels that contributed most to decoding valence and arousal (henceforth, informative voxels), voxels with the highest and lowest 5% of logistic regression weights were identified for each cross-validation fold. A union of such voxels across cross-validation folds was visualized for each participant. To investigate the consistency of informative voxel locations across individuals, a voxel location probability map was generated across participants after convolving each voxel with a 4 mm Gaussian kernel (Kober et al., 2008). The probability map was further thresholded by a simulated null hypothesis distribution to control for multiple comparisons (FWE = 0.05).

Cross-participant analysis

To establish commonalities between participants' neural representations of affective states, we conducted cross-participant classification. Data from all but one participant were used to train a classifier to distinguish affective states. The classifier was then tested on the data of the left-out participant. Trials of the same condition were averaged for the test set. Classification was repeated iteratively until each participant's data served once as the test set. Entropy-based feature selection (Poldrack et al., 2009) was conducted on the combined data of all participants except the one left out for testing. The significance of classification accuracy was evaluated based on the binomial distribution.

Lower dimensional representation

The INDSCAL model was used to investigate the lower dimensional representation of affective space from the pattern of brain activity (mean PSC) elicited by viewing photographs from each of the five affective categories. Prior to analysis the top 400 most stable voxels were selected for each individual using the feature selection method described above (that was not informed about the affective categories). The conditions-by-voxel mean PSC matrices for each individual were averaged across six alternating presentations of each condition to obtain two data points, or exemplars, for each experimental condition. Pairwise correlations were computed between exemplars, which resulted in a single 10 × 10 exemplar-by-exemplar matrix for each individual. The 13 correlation matrices were then analyzed by INDSCAL, applying a monotone function relating similarities to model distances. INDSCAL fits a common configuration but allows for the expansion or shrinking of dimensions for individuals as given by dimensional weight, thought to reflect attention to that dimension. The analysis also indicates how well each individual is fit by the common configuration.

Results

Category identification of affective states within participants

A classifier was trained for each participant to determine if it was possible to identify the four affective categories based on whole brain activation elicited by picture stimuli. Classification accuracies for classification of the four affective states significantly exceeded the chance level (0.25) for all levels of the most stable voxels ($p<0.05$), for the majority of participants (Fig. 3). The highest classification accuracy obtained for a single participant was 0.77. Accurate classification was robust across the wide range of voxels used (from 25 to 2000). The average accuracy increased with number of voxels included up to an intermediate number and then declined with use of additional voxels. The voxels selected based on stability were distributed throughout the brain. Using the 400 voxels with most stable responses, twelve of 13 participants showed significant classification accuracies ($p<0.05$).

Next, a classifier was trained for each participant to determine if it was possible to identify positive and negative valence based on brain activation elicited by picture stimuli. Classification accuracies for the two valence levels significantly exceeded the chance level (0.50) for all levels of the most stable voxels ($p<0.05$) for the majority of participants (Fig. 4). The highest classification accuracy obtained for a single participant was 0.92. Using the top 400 voxels with most stable responses, all 13 participants showed significant classification accuracies ($p<0.05$). Once again, average classification accuracy showed the typical concave function across the number of voxels included in the analysis.

Finally, a classifier was trained for each participant to determine if it was possible to identify high or low arousal based on brain activation elicited by picture stimuli. Classification accuracies for the two levels of arousal significantly exceeded the chance level (0.50) for all levels of the most stable voxels ($p<0.05$) for the majority of participants (Fig. 5). The highest classification accuracy obtained for a single participant was 0.92. Using the 400 most stable voxels, twelve of 13 participants showed significant classification accuracies ($p<0.05$). As in the other two classification analyses, average classification accuracy rose with inclusion of more voxels, peaked at 400 voxels and then slowly declined but remained significant with inclusion of up to 2000 voxels.

The locations of voxels with largest classifier weights for identification of valence or arousal were distributed throughout the brain (Fig. 6). The informative voxels were distributed across the brain for a wide

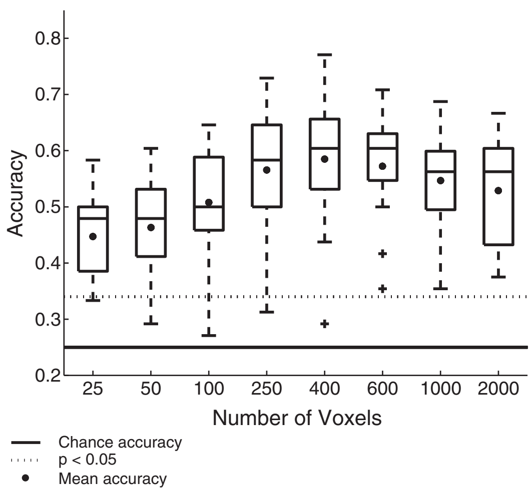

Fig. 3. Within-participant accuracies for classifying affective states. Classification accuracies across the 13 participants, summarized by box plots, are shown for different subsets of the most replicable gray matter voxels.

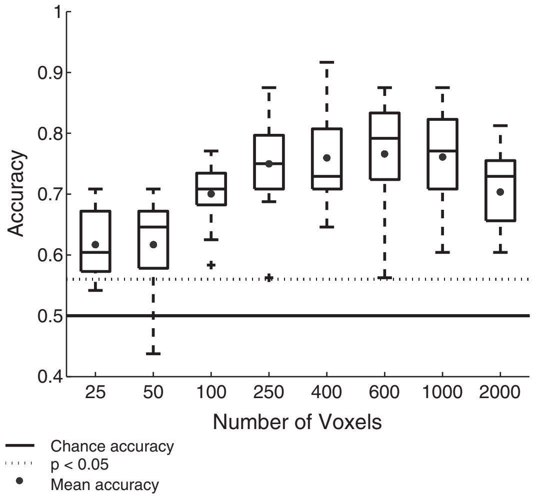

Fig. 4. Within-participant accuracies for classifying positive and negative valence. Classification accuracies across the 13 participants, summarized by box plots, are shown for different subsets of the most replicable gray matter voxels.

range of voxels considered by the feature selection method. Moreover, the locations of informative voxels were similar across participants (as summarized in Fig. 6B). Informative voxel location clusters that were robustly identified across participants (based on 2000 voxels) and were critical for both valence and arousal decoding included the inferior temporal gyrus, lentiform nucleus, medial prefrontal cortex, middle occipital gyrus, middle temporal gyrus, parahippocampus, postcentral gyrus, and precuneus. Voxel locations specifically informative for decoding of valence included the anterior cingulate, fusiform, and inferior frontal gyri, superior parietal lobule, and ventrolateral prefrontal cortex. Voxel locations specifically informative for decoding arousal included dorsolateral prefrontal cortex, inferior parietal lobule, medial superior temporal gyrus, posterior superior temporal sulcus, and insula. Notably, voxels in bilateral amygdala were also found informative in both valence and arousal decoding, although they did not form clusters.

In summary, classifiers trained on single-participant data were able to identify affective states, valence, and arousal reliably above chance. Thus, we conclude that information about valence and arousal is represented in whole brain activation patterns elicited by viewing photographs within each participant.

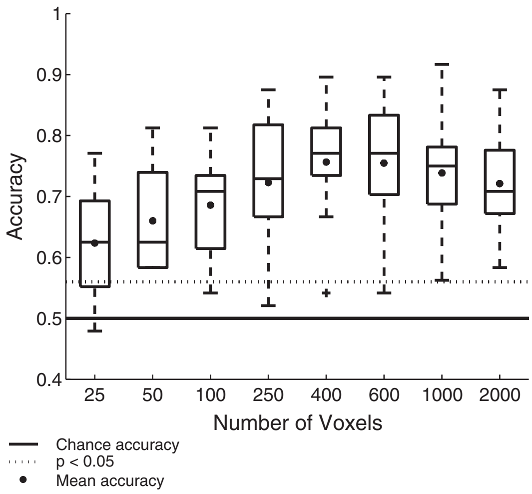

Fig. 5. Within-participant accuracies for classifying high and low arousal. Classification accuracies across the 13 participants, summarized by box plots, are shown for different subsets of the most replicable gray matter voxels.

Category identification of affective states across participants

To examine the consistency of the neural representations of affect across participants, the whole brain activation data from all but one participant were used to identify the affective category of stimuli presented to the left-out participant. A classifier was trained on the data from all but one participant and tested on the data from the left-out participant. The highest accuracy for classifying the four affective states obtained for any voxel level was 0.75 (compared to 0.25 chance level). Classification accuracies for the four affective states were significant for all levels of the most stable voxels we tested ($p<0.05$) for the majority of participants (Fig. 7).

Next, a classifier was trained on the combined data from all but one participant to identify levels of valence in the left out participant. The highest accuracy for classifying positive and negative valence obtained for any voxel level was 1.0. Classification accuracies for the two levels of valence were significant for all levels of the most stable voxels ($p<0.05$) for the majority of participants (Fig. 8).

Finally, a classifier was trained on the combined data from all but one participant to identify levels of arousal in the left out participant. The highest accuracy for classifying high and low arousal obtained for any voxel level was 1.0. Classification accuracies for the two levels of arousal were significant for all levels of the most stable voxels ($p<0.05$) for the majority of participants (Fig. 9).

The successful classification of the emotion related properties of stimuli of each individual from a classifier trained on the other 12 participants implies that the neural activation patterns elicited by affective categories are highly consistent across individuals. This similarity in patterns of brain activation suggests that the neural representation of affect is similar from one individual to another. These results are consistent with the similarity of locations of informative voxels across participants (Fig. 6). Note that cross-participant prediction required recruiting higher numbers of voxels than within-participant prediction.

Lower dimensional representation of affective space

To examine the lower dimensional representation of affective space, the correlation matrices for odd and even versions of each of the five states were generated for each participant based on the most stable 400 voxels for each participant. The 10×10 correlation matrices were then input into INDSCAL and a common configuration was abstracted, with fit indices and dimensional weights for each participant. In INDSCAL, dimensional weights for each individual are only interpretable when dimensions are left unrotated.

The two dimensional solution had an overall stress value of 0.27. Although this is a fairly high value of stress, it represents a strong reduction from the one dimensional solution (stress = 0.42). The representation is consistent with the circumplex model, with the first dimension reflecting arousal in that it separates the low and high arousal conditions, and the second dimension reflecting valence in that it separates the positive and negative valence conditions (Fig. 10). The neutral L0 replicates tend to cluster with the negative valence replicates in this configuration. Note that replicates of each state tended to be closer to each other in the space than to other states, reflecting reliability in classification.

As a fit index, two participants had an R^2 greater than 0.70, three had R^2 values between 0.50 and 0.70, four had R^2 values between 0.40 and 0.50, and four had R^2 values below 0.40, suggesting these four were not well fit by the model. Individual differences can be represented in INDSCAL by differences in the dimensional weights. While the dimension corresponding to the greatest variation was arousal (Dimension 1), eight of the 13 participants were shown to weight valence more than arousal. Overall, the recovered solution provides support for the circumplex model and appears to be representative of the majority of participants.

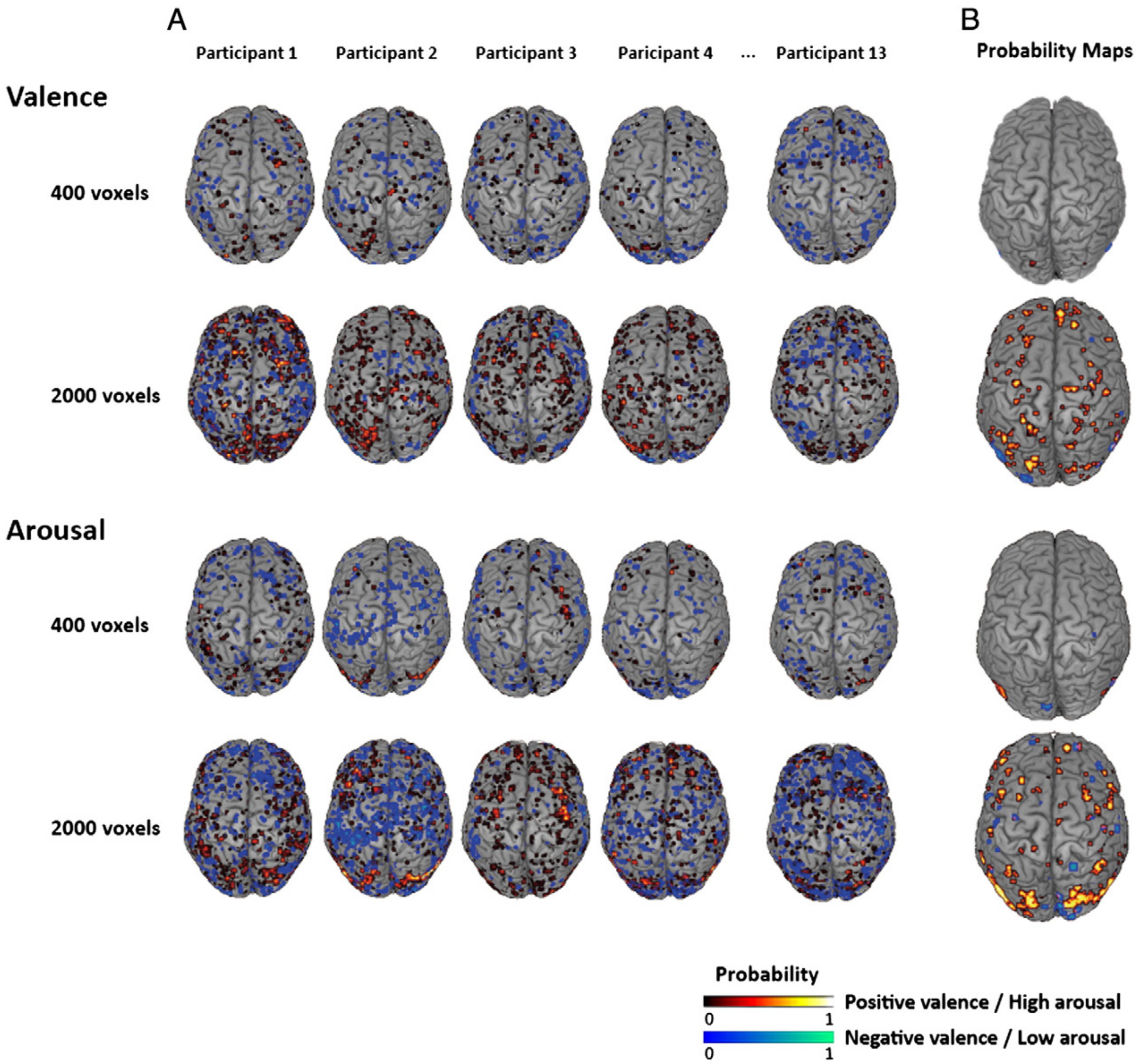

Fig. 6. Most informative voxels for decoding of valence and arousal are shown on a surface rendering. *Panel A* shows the union of most informative voxels across each cross-validation fold for 5 out of the 13 participants for 400 and 2000 voxels. The hot color map indicates the probability of the top 5% of voxels that were most informative for identifying positive valence or high arousal. The cold color map indicates the probability of the top 5% of voxels that were most informative for identifying negative valence or low arousal. *Panel B* shows the thresholded probability maps (FWE = 0.05) of the informative voxels that were consistently identified across all 13 participants.

To examine the dependence of this solution on the specific set of voxels selected, we conducted INDSCAL analyses on each of 20 successive blocks of 50 voxels. Indices of the correspondence of the solution to the predicted circumplex were obtained by rotating each solution to the design matrix orientation using a Procrustes rotation and computing the correlations between the rotated dimension values and the design values for valence and arousal. As a point of comparison, the 400 voxel solution of Fig. 10 resulted in a correlation of $r = 0.93$ with valence and $r = 0.93$ with arousal when rotated. The correlations for each successive block of 50 voxels are shown in Fig. 11. The correlations obtained for three of the first four blocks of 50 voxels are very comparable to those for the 400 voxel solution. As we proceed to blocks of lower voxel stability, the correlations tend to attenuate, reflecting reduced correspondence to the circumplex. Interestingly, however, in these later blocks typically one of the correlations is high, indicating that voxels of lower stability may still be sensitive to either valence or arousal. These results support the conclusion that the good correspondence of the INDSCAL solution shown in Fig. 10 is fairly robust across different sets of voxels selected from the most stable voxels.

The results of visualizing the lower dimensional representation of affective space indicate that the representation of affect abstracted from fMRI data includes both valence and arousal. Importantly, the voxel selection procedure was blind to the dimensional structure. The underlying dimensions of valence and arousal obtained from neuroimaging data were similar to that obtained from behavioral results (Fig. 1) and lend support to the circumplex model of emotion. Furthermore, the extreme exemplars from the two arousal categories and the two valence categories were linearly separable in the dimensional space, reflecting a simple dimensional basis for classification.

Discussion

The current study explored whether MVPA methods applied to whole brain activity patterns could be used to identify the valence and arousal levels elicited by viewing of photographs and to examine the internal representation of affective states from fMRI data for comparison to dimensional models of affect. In contrast to studies that focus on specific regions of interest, our focus was to utilize the distributed

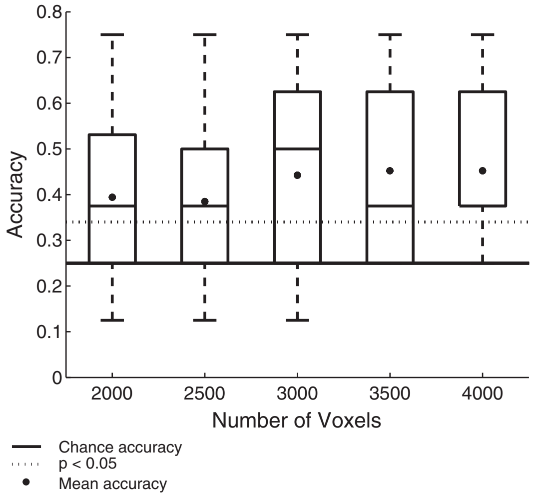

Fig. 7. Cross-participant accuracies for classifying affective states. Cross-participant classification accuracies across the 13 participants, summarized by box plots, are shown for different subsets of the most replicable gray matter voxels.

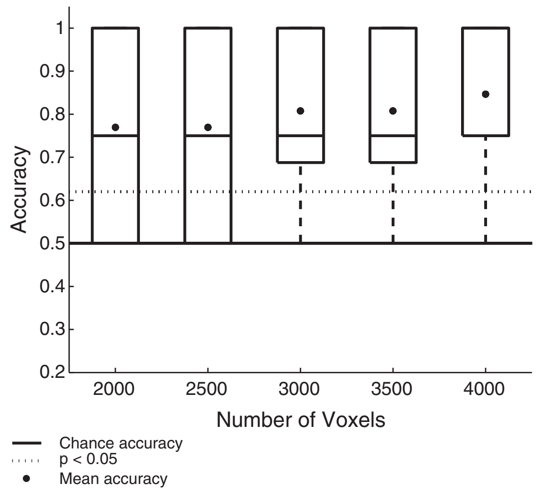

Fig. 9. Cross-participant accuracies for classifying high and low arousal. Cross-participant classification accuracies across the 13 participants, summarized by box plots, are shown for different subsets of the most replicable gray matter voxels.

representations elicited by affective stimuli to capture the underlying affective states. Categories of affect, valence, and arousal were successfully decoded from patterns of brain activation within participants. The within-participant decoding results demonstrate that information unique to valence and arousal lies within distributed patterns of brain activation across the whole brain and can be used to predict which valence and arousal levels a participant was experiencing as elicited by viewing of affect-related photographs. Indeed, as shown in Figs. 3, 4 and 5, within-participant classification was above chance for the majority of the participants when based on as few as the 25 most stable responding voxels. However, when classifying across participants, a much greater number of voxels was required, indicating that the most stable voxels selected for the within-participant classification may be fairly unique to individuals.

Consistent with the larger number of voxels needed to classify across participants, the probability maps based on the most informative of the 2000 most stable voxels showed several distinct clusters. The informative voxels included regions of considerable overlap with previous MVPA literature of emotion decoding, including the fusiform gyrus, inferior frontal gyrus, insula, medial prefrontal cortex, anterior cingulate cortex, and superior temporal gyrus and sulcus (Ethofer et al., 2009; Kanske and Hasting, 2010; Peelen et al., 2010; Pessoa and Padmala, 2007; Said et al., 2010). These clusters of informative voxels, along with a cluster implicating the occipital cortex, have also been consistently found in neuroimaging studies using statistical parametric mapping (Bush et al., 2000; Colibazzi et al., 2010; Gerber et al., 2008; Hagan et al., 2009; Killgore and Yurgelun-Todd, 2004; Narumoto et al., 2001; Posner et al., 2009; Robins et al., 2009; Wager et al., 2008). The distributed nature of brain activity containing information about the valence and arousal of emotionally-evocative pictures may be adaptive, as quick, distributed processing of emotional stimuli enhances an organism's ability to respond appropriately (Pessoa and Adolphs, 2010).

In the current study participants were not explicitly evaluating the affective content. We examined the valence and arousal levels elicited by viewing of photographs, with the participant's task simply to maintain fixation on a fixation cross. Thus, the current study contributed to literature that uses incidental methods of inducing affective states (Ethofer et al., 2009; Nielen et al., 2009). The experimental paradigm was designed to monitor responses that were not unique to a single photograph, such as semantic content, but common across a set of photographs. Moreover, the stimuli used in the current study were standardized to have the same mean value for hue, saturation, and intensity. This procedure reduced the ability to decode stimulus categories based on the lower level neural representations of visual

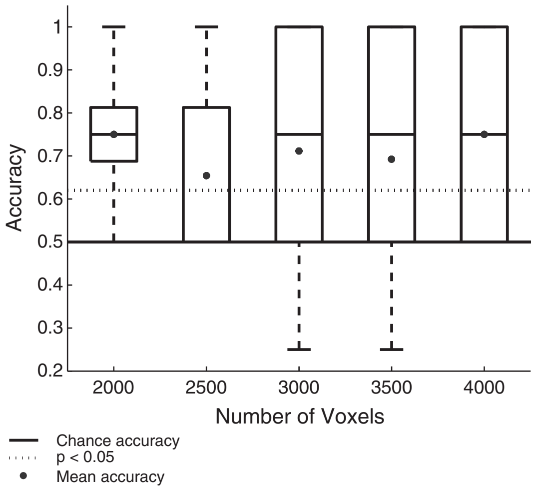

Fig. 8. Cross-participant accuracies for classifying positive and negative valence. Cross-participant classification accuracies across the 13 participants, summarized by box plots, are shown for different subsets of the most replicable gray matter voxels.

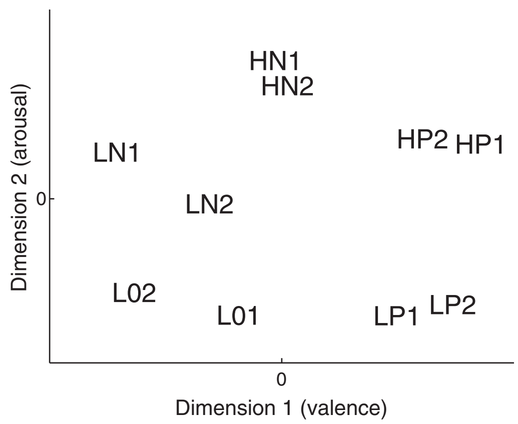

Fig. 10. Lower dimensional representation of affective space based on fMRI data. LP — low arousal, positive valence; LN — low arousal, negative valence; L0 — low arousal, neutral valence; HP — high arousal, positive valence; HN — high arousal, negative valence, 1 — odd trials, 2 — even trials.

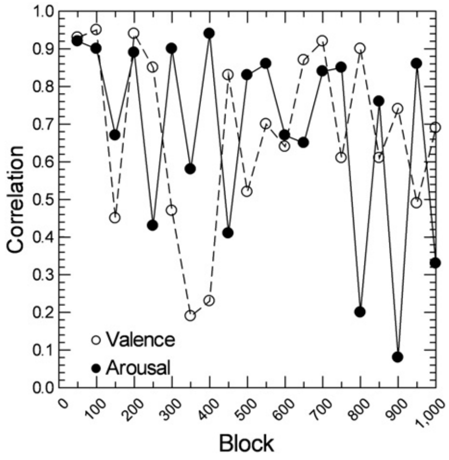

Fig. 11. Correlational indices for correspondence between MDS-derived scale values and design values for arousal and valence are shown for each successive block of 50 voxels.

properties and increased the likelihood that decoding was based on a higher level neural representation of affect. Consistent with this interpretation, the selected voxels were distributed throughout the brain.

Despite the controls we implemented there are several limitations that may be placed on interpretations of our study. First, we cannot disentangle the extent to which our analytic methods are modeling affective states or brain activity more directly associated with the perception and semantic processing of the picture stimuli. One interpretation of the widespread locations of informative voxels found in our study is that we are tapping into systems linked to affect as well as semantic and perceptual processing of the stimuli. One way to disentangle these different sources is to vary the method of inducing affective states within an experiment and look for common voxels that are predictive across methods. Second, and building on the prior point, we did not have physiological or behavioral measures for participants that indicate to what degree they may have experienced the different affective states. This is partly a consequence of using an incidental task in which affect was not the focus of processing in the scanner. The success of our analytic methods suggests that the manipulation of affect was successful, but we have no independent verification. Finally, although we have discussed some correspondences between informative voxels in our study and ROIs identified in other fMRI studies of affect, one must be very cautious in interpreting such correspondences. This is because the discriminative weights we used to identify informative voxels may not be stable across folds and will generally vary with inclusion or exclusion of other voxels. Our approach is primarily designed to use whole brain activation to identify states within and across participants, rather than to identify the key ROIs involved in processes producing those states.

Using multidimensional scaling techniques, the lower dimensional representation of affect was extracted from fMRI data. The resulting internal representation of affect showed a separation of affective stimuli based upon the dimensions of valence and arousal. The current findings lend support to previous behavioral studies investigating the circumplex model of affect. While categorical models of affect may account for the decoding of affective categories, these models do not predict the circumplex form of the recovered dimensional solution. The circumplex model of affect suggests that all emotions or affective states can be distinguished in terms of varying levels of valence and arousal that form a circular representation in the two dimensional affective space. The current findings suggest that patterns of brain activity contain information regarding valence and arousal when representing affective states. That is, affect can be distinguished in terms of varying levels of valence and arousal from both the behavioral and neuroimaging data. The lower dimensional solutions provide additional support for the utility of the dimensional approach and the circumplex model. The recovered circumplex structure constructed from fMRI data using INDSCAL is all the more impressive as it was based on incidental exposure to affective stimuli, rather than the explicit evaluation of emotional content. This result provides additional support for the circumplex model, as it is not dependent on the rating and sorting procedures used in behavioral literature (Roberts and Wedell, 1994; Russell, 1980).

The ability to identify the valence and arousal levels elicited by viewing of photographs across participants supports a common neural basis for representation of affect across people. Cross participant classification was achieved despite three major hurdles. First, individuals are known to widely differ in functional organization. Second, methodological difficulties exist in normalizing the morphological differences found among human brains. Third, individuals appear to differ in the degree to which they weight differences in arousal and valence, as shown in our INDSCAL analyses. Classification of mental states across individuals has been previously shown for visually depicted objects (Shinkareva et al., 2008), concrete nouns referring to physical objects (Just et al., 2010; Shinkareva et al., 2011), lie detection (Davatzikos et al., 2005), attentional tasks (Mourão-Miranda et al., 2005), cognitive tasks (Poldrack et al., 2009), and voxel-by-voxel correspondence across individuals has been demonstrated during movie-watching (Hasson et al., 2004). The current results demonstrate, for the first time, the ability to identify the affective category of a set of photographs viewed by a participant based on neural activation data from other participants.

In summary, the results from the current study provide support for the utility of MVPA methods for analyzing neuroimaging data related to affect. The affective manipulation was incidental and presumably resulted in subtle changes in affective states. Despite the relatively weak signal, the MVPA methods based on whole brain activity patterns were able to successfully classify states within and across participants, as well as abstract the internal representation of affective states consistent with the circumplex model of affect. When combined with growing evidence of specific neural circuitry that responds to modulations of valence and arousal (Lang and Bradley, 2010; Nielen et al., 2009; Pessoa and Adolphs, 2010), our results bolster the conclusion that neural representations of affect states may be successfully represented along the dimensions of valence and arousal.

Acknowledgments

We would like to thank John Richards, Michael Schillaci, and Jennifer Vendemia for helpful discussions at the earlier stages of the project, and the two anonymous reviewers for helpful suggestions on the earlier version of the manuscript.

References

Anders, S., Lotze, M., Erb, M., Grodd, W., Birbaumer, N., 2004. Brain activity underlying emotional valence and arousal: a response-related fMRI study. Hum. Brain Mapp. 23, 200–209.

Anderson, A.K., Christoff, K., Stappen, I., Panitz, D., Chahremani, D.G., Glover, G., Gabrieli, J.D., Sobel, N., 2003. Dissociated neural representations of intensity and valence in human olfaction. Nat. Neurosci. 6, 196–202.

Barrett, L.F., 1998. Independence and bipolarity in the structure of current affect. J. Pers. Soc. Psychol. 74, 967.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer, New York.

Borg, I., Groenen, P., 1997. Modern Multidimensional Scaling: Theory and Applications. Springer-Verlag, New York.

Bush, G., Luu, P., Posner, M.I., 2000. Cognitive and emotional influences in anterior cingulate cortex. Trends Cogn. Sci. 4, 215–222.

Colibazzi, T., Posner, J., Wang, Z., Gorman, D., Gerber, A., Yu, S., Zhu, H., Kangarlu, A., Duan, Y., Russell, J., Peterson, B.S., 2010. Neural systems subserving valence and arousal during the experience of induced emotions. Emotion 10, 377–389.

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.

Ekman, P., 1992a. Are there basic emotions? Psychol. Rev. 99, 550–553.

Ekman, P., 1992b. An argument for basic emotions. Cogn. Emotion 6, 169.

Ethofer, T., Van De Ville, D., Scherer, K., Vuilleumier, P., 2009. Decoding of emotional information in voice-sensitive cortices. Curr. Biol. 19, 1028–1033.

Gerber, A., Posner, J., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., Kangarlu, A., Zhu, H., Russell, J., Peterson, B.S., 2008. An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. Neuropsychologia 46, 2129–2139.

Grimm, S., Schmidt, C.F., Bermpohl, F., Heinzel, A., Dahlem, Y., Wyss, M., Hell, D., Boesiger, P., Boeker, H., Northoff, G., 2006. Segregated neural representation of distinct emotion dimensions in the prefrontal cortex—an fMRI study. Neuroimage 30, 325–340.

Hagan, C.C., Woods, W., Johnson, S., Calder, A.J., Green, G.G.R., Young, A.W., 2009. MEG demonstrates a supra-additive response to facial and vocal emotion in the right superior temporal sulcus. Proc. Natl. Acad. Sci. U. S. A. 106, 20010–20015.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. Science 303, 1634–1640.

Haynes, J.D., Rees, G., 2006. Decoding mental states from brain activity in humans. Nat. Rev. Neurosci. 7, 523–534.

Just, M.A., Cherkassky, V.L., Aryal, S., Mitchell, T.M., 2010. A neurosemantic theory of concrete noun representation based on the underlying brain codes. PLoS One 5, e8622.

Kanske, P., Hasting, A.S., 2010. Decoding modality-independent emotion perception in medial prefrontal and superior temporal cortex. J. Neurosci. 30, 16417–16418.

Kensinger, E.A., Corkin, S., 2004. Two routes to emotional memory: distinct neural processes for valence and arousal. PNAS 101, 3310–3315.

Killgore, W.D.S., Yurgelun-Todd, D.A., 2004. Activation of the amygdala and anterior cingulate during nonconscious processing of sad versus happy faces. Neuroimage 21, 1215–1223.

Kober, H., Barrett, L.F., Joseph, J., Bliss-Moreau, E., Lindquist, K., Wager, T.D., 2008. Functional grouping and cortical–subcortical interactions in emotion: a meta-analysis of neuroimaging studies. Neuroimage 42, 998–1031.

Lang, P.J., Bradley, M.M., 2010. Emotion and the motivational brain. Biol. Psychol. 84, 437–450.

Lang, P.J., Bradley, M.M., Cuthbert, B.N., 2008. International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual. University of Florida, Gainesville, FL.

Lazarus, R.S., 1991. Emotion and Adaptation. Oxford University Press, New York.

Lazarus, R.S., 1995. Vexing research problems inherent in cognitive-mediational theories of emotion-and some solutions. Psychol. Inq. 6, 183–196.

Lewis, P.A., Gritchley, H.D., Rotshtein, P., Dolan, R.J., 2007. Neural correlates of processing valence and arousal in affective words. Cereb. Cortex 17, 742–748.

Mitchell, T.M., Shinkareva, S.V., Carlson, A., Chang, K.-M., Malave, V.L., Mason, R.A., Just, M.A., 2008. Predicting human brain activity associated with the meanings of nouns. Science 320, 1191–1195.

Mourão-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.

Narumoto, J., Okada, T., Sadato, N., Fukui, K., Yonekura, Y., 2001. Attention to emotion modulates fMRI activity in human right superior temporal sulcus. Cogn. Brain Res. 12, 225–231.

Nielen, M.M.A., Heslenfeld, D.J., Heinen, K., Van Strien, J.W., Witter, M.P., Jonker, C., Veltman, D.J., 2009. Distinct brain systems underlie the processing of valence and arousal of affective pictures. Brain Cogn. 71, 387–396.

Norman, K.A., Polyn, S.M., Detre, G.J., Haxby, J.V., 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn. Sci. 10, 424–430.

O'Toole, A.J., Jiang, F., Abdi, H., Penard, N., Dunlop, J.P., Parent, M.A., 2007. Theoretical, statistical, and practical perspectives on pattern-based classification approaches to the analysis of functional neuroimaging data. J. Cogn. Neurosci. 19, 1735–1752.

Onton, J., Makeig, S., 2009. High-frequency broadband modulations of electroencephalographic spectra. Front. Hum. Neurosci. 3, 61.

Panksepp, J., 1992. A critical role for "affective neuroscience" in resolving what is basic about basic emotions. Psychol. Rev. 99, 554–560.

Panksepp, J., 1998. Affective Neuroscience: The Foundations of Human and Animal Emotions. Oxford University Press, New York.

Peelen, M.V., Atkinson, A.P., Vuilleumier, P., 2010. Supramodal representations of perceived emotions in the human brain. J. Neurosci. 30, 10127–10134.

Pereira, F., Mitchell, T., Botvinick, M., 2009. Machine learning classifiers and fMRI: a tutorial overview. Neuroimage 45, S199–S209.

Pessoa, L., Adolphs, R., 2010. Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. Nat. Rev. Neurosci. 11, 773–782.

Pessoa, L., Padmala, S., 2007. Decoding near-threshold perception of fear from distributed single-trial brain activation. Cereb. Cortex 17, 691–701.

Poldrack, R.A., Halchenko, Y.O., Hanson, S.J., 2009. Decoding the large-scale structure of brain function by classifying mental states across individuals. Psychol. Sci. 20, 1364–1372.

Posner, J., Russell, J., Gerber, A., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., Kangarlu, A., Zhu, H., Peterson, B.S., 2009. The neurophysiological bases of emotion: an fMRI study of the affective circumplex using emotion-denoting words. Hum. Brain Mapp. 30, 883–895.

Posner, J., Russell, J., Peterson, B.S., 2005. The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Dev. Psychopathol. 17, 715–734.

Roberts, J.S., Wedell, D.H., 1994. Context effects on similarity judgments of multi-dimensional stimuli: Inferring the structure of the emotion space. J. Exp. Soc. Psychol. 30, 1–38.

Robins, D.L., Hunyadi, E., Schultz, R.T., 2009. Superior temporal activation in response to dynamic audio–visual emotional cues. Brain Cogn. 69, 269–278.

Rolls, E., 1999. The Brain and Emotion. Oxford University Press, New York.

Russell, J., 1980. A circumplex model of affect. J. Pers. Soc. Psychol. 39, 1161–1178.

Said, C.P., Moore, C.D., Engell, A.D., Todorov, A., Haxby, J.V., 2010. Distributed representations of dynamic facial expressions in the superior temporal sulcus. J. Vis. 10, 1–12.

Schlosberg, H., 1954. Three dimensions of emotion. Psychol. Rev. 61, 81–88.

Shinkareva, S.V., Malave, V.L., Mason, R.A., Mitchell, T.M., Just, M.A., 2011. Commonality of neural representations of words and pictures. Neuroimage 54, 2418–2425.

Shinkareva, S.V., Mason, R.A., Malave, V.L., Wang, W., Mitchell, T.M., Just, M.A., 2008. Using FMRI brain activation to identify cognitive states associated with perception of tools and dwellings. PLoS One 3 e1394-e1394.

Sitaram, R., Lee, S., Ruiz, S., Rana, M., Veit, R., Birbaumer, N., 2011. Real-time support vector classification and feedback of multiple emotional brain states. Neuroimage 56, 753–765.

Small, D.M., Gregory, M.D., Mak, Y.E., Gitelman, D., Mesulam, M.M., Parrish, T., 2003. Dissociation of neural representation of intensity and affective valuation in human gustation. Neuron 39, 701–711.

Wager, T.D., Barrett, L.F., Bliss-Moreau, E., Lindquist, K., Duncan, S., Kober, H., Joseph, J., Davidson, M., Mize, J., 2008. The neuroimaging of emotion. In: Lewis, M., Haviland-Jones, J.M., Barrett, L.F. (Eds.), Handbook of Emotions. The Guilford Press, New York, pp. 249–271.

Watson, D., Tellegen, A., 1985. Toward a consensual structure of mood. Psychol. Bull. 98, 219–235.