Behavioral/Systems/Cognitive

Supramodal Representations of Perceived Emotions in the Human Brain

Marius V. Peelen,[1,2,3,4] **Anthony P. Atkinson,**[5] **and Patrik Vuilleumier**[3,4]

[1]Center for Mind/Brain Sciences, University of Trento, 38068 Rovereto (TN), Italy, [2]Department of Psychology, Harvard University, Cambridge, Massachusetts 02138, [3]Swiss Center for Affective Sciences, University of Geneva, CH-1205 Geneva, Switzerland, [4]Laboratory for Neurology and Imaging of Cognition, Department of Neurosciences and Clinic of Neurology, University Medical Center, CH-1211 Geneva, Switzerland, and [5]Department of Psychology, Durham University, Science Laboratories, Durham DH1 3LE, United Kingdom

Basic emotional states (such as anger, fear, and joy) can be similarly conveyed by the face, the body, and the voice. Are there human brain regions that represent these emotional mental states regardless of the sensory cues from which they are perceived? To address this question, in the present study participants evaluated the intensity of emotions perceived from face movements, body movements, or vocal intonations, while their brain activity was measured with functional magnetic resonance imaging (fMRI). Using multivoxel pattern analysis, we compared the similarity of response patterns across modalities to test for brain regions in which emotion-specific patterns in one modality (e.g., faces) could predict emotion-specific patterns in another modality (e.g., bodies). A whole-brain searchlight analysis revealed modality-independent but emotion category-specific activity patterns in medial prefrontal cortex (MPFC) and left superior temporal sulcus (STS). Multivoxel patterns in these regions contained information about the category of the perceived emotions (anger, disgust, fear, happiness, sadness) across all modality comparisons (face–body, face–voice, body–voice), and independently of the perceived intensity of the emotions. No systematic emotion-related differences were observed in the overall amplitude of activation in MPFC or STS. These results reveal supramodal representations of emotions in high-level brain areas previously implicated in affective processing, mental state attribution, and theory-of-mind. We suggest that MPFC and STS represent perceived emotions at an abstract, modality-independent level, and thus play a key role in the understanding and categorization of others' emotional mental states.

Introduction

Successful social interaction requires a precise understanding of the feelings, intentions, thoughts, and desires of other people. The importance of understanding other people's minds is reflected in the exceptional ability of humans to infer complex mental states from subtle sensory cues. In the domain of emotion, these sensory cues come from various sources, such as facial expressions, body movements, and vocal intonations. For example, fear can be recognized using visual information from the eye region of faces, from movements and postures of body parts, as well as from acoustic changes in voices (de Gelder et al., 2006; Heberlein and Atkinson, 2009). Despite their disparate sensory nature, these signals can similarly lead to the recognition of an emotional state (i.e., that the person is afraid) and activate similar emotion-specific responses in the observer (Magnée et al., 2007). Thus, when evaluating how someone feels, our representation of another person's emotional state (e.g., fear) is abstracted away from the specific sensory input (e.g., a cow-

ering body posture or a trembling voice). This raises the critical question of whether there may be brain regions that encode emotions independently of the modality through which they are perceived.

In the present study, we sought to identify modality-independent neural representations of emotional states by recording functional magnetic resonance imaging (fMRI) responses in participants while they viewed emotional expressions portrayed in three different perceptual modalities (body, face, and voice). We asked whether brain regions exist that exhibit patterns of activity that are specific to the perceived emotional states but independent of the modality through which these states are perceived.

We measured brain activity in 18 volunteers while they viewed or listened to actors expressing 5 different emotions (anger, disgust, happiness, fear, sadness). Each of these emotions could be expressed through either 1) body movements with the face obscured (body), 2) facial expressions with the body not visible (face), or 3) vocal intonations without any visual stimulus (voice). After each stimulus presentation, volunteers were instructed to assess how intensely they thought the actor was feeling the emotion just expressed (see Fig. 1 for an overview of the design). Importantly, while the emotional state of the actor could readily be inferred from each of the three stimulus modalities, different sensory cues must be used for these inferences.

fMRI data were analyzed using multivoxel pattern analysis (MVPA). This recently developed technique is sensitive to fine-grained neural representations (Haynes and Rees, 2005; Kami-

tani and Tong, 2005), and has been successfully used to reveal, for example, coding of object category in ventral temporal cortex (Haxby et al., 2001; Peelen et al., 2009), individual face identity in anterior temporal cortex (Kriegeskorte et al., 2007), as well as voice identity and voice prosody in auditory cortices (Formisano et al., 2008; Ethofer et al., 2009). Using a spherical searchlight approach (Kriegeskorte et al., 2006), we tested for regions in the brain where local activity patterns contained information about the emotion categories (e.g., fear, anger) independent of stimulus modality (body, face, voice).

Materials and Methods

Participants. Eighteen healthy adult volunteers participated in the study (10 women, mean age 26 years, range 20–32). All were right-handed, with normal or corrected-to-normal vision, and no history of neurological or psychiatric disease. Participants all gave informed consent according to ethics regulations.

Stimuli. Movies of emotional facial expressions were taken from the set created by Banse and Scherer (1996), and used by Scherer and Ellgring (2007). The movies were cropped such that non-facial body parts were not visible. Four actors (2 male, 2 female) expressed each emotion.

Movies of emotional body expressions were taken from the set created and validated by Atkinson et al. (2004; 2007), and used in Peelen et al. (2007). The actors wore uniform dark-gray, tight-fitting clothes and headwear so that all body parts were covered. Facial features and expressions were not visible. Four actors (2 male, 2 female) expressed each emotion.

Emotional voice stimuli were taken from the Montreal Affective Voices set, created and validated by Belin et al. (2008). These stimuli consisted of short (~1 s), nonlinguistic interjections ("aah") expressing different emotions. Four actors (2 male, 2 female) expressed each emotion.

All stimuli (in all three modalities) were selected to reliably convey the target emotions, as indicated by prior behavioral studies with the same material.

Design and procedure. Participants performed 6 fMRI runs, each starting and ending with a 10 s fixation baseline period. Within each run, 36 trials of 8.5 s were presented in 3 blocks of 12 trials (Fig. 1). These blocks were separated by 5 s fixation periods. The 3 blocks differed in the type of stimuli presented (bodies, faces, or voices), and each of them included 2 different clips (of different actors) for each of the 5 emotion conditions (anger, disgust, fear, happiness, sadness) and a neutral condition (not included in the analyses, see below). Trials were presented in random order within each block, such that the emotion category itself could not be predicted before stimulus onset. The order of blocks was counterbalanced across runs. Each run lasted ~5 min 40 s.

Each trial consisted of a 2 s fixation cross, followed by a 3 s movie or sound clip, a 1 s blank screen, and a 2.5 s response window. Because the sound clips were shorter than the movie clips, an extra empty time was added after the sound clips to equate the timing across conditions. After the presentation of each stimulus, participants were asked to rate the intensity of the emotional expression on a 3-point scale. This task was chosen to encourage participants to actively evaluate the perceived emotional states, because explicit judgments may increase activity in key brain regions involved in social cognition and mental state attribution (Adolphs, 2009; Lee and Siegle, 2009; Zaki et al., 2009). In addition, in
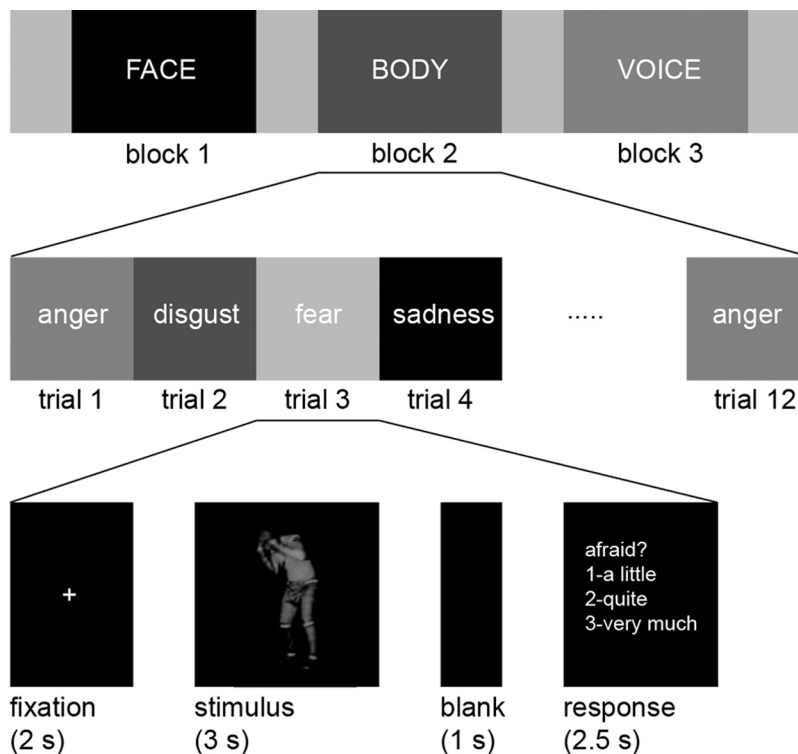

Figure 1. Schematic overview of the design. Each run consisted of 3 blocks of 12 trials each. Blocks differed in the type of stimuli presented (bodies, faces, or voices). On each trial, one of five different emotions could be expressed in a given modality. These trials were presented in random order within each block, whereas the order of blocks was counterbalanced across runs. Each trial consisted of a 2 s fixation cross, followed by a 3 s stimulus (movie or sound clip), a 1 s blank screen, and then a 2.5 s response window. After the presentation of each stimulus, participants were asked to rate the intensity of the perceived emotion on a 3-point scale. Not drawn to scale.

contrast to a direct emotion categorization task, the intensity judgment task has the advantage that it decouples the button press responses from the emotion categories since it requires the same responses for all emotion categories. For example, after an angry body movie, the response was cued by the following text display: "Angry? 1—a little, 2— quite, 3—very much," where "1", "2", "3" referred to 3 response buttons [from left to right; for half the participants this response mapping was reversed (from right to left)] on a keypad held in the right hand. Hence, the planning and execution of motor responses were fully orthogonal to the emotion categories (i.e., the same responses had to be selected for all categories and all modalities), such that motor selection processes could not confound any differential activity across emotions. For the neutral condition, participants rated how "lively" the movie was. The neutral conditions were not further analyzed due to a relative ambiguity of these expressions in this context, and the low intensity ratings for the neutral conditions (mean intensity rating: 1.76) relative to the emotion conditions (mean intensity rating: 2.24). Due to technical problems, responses could not be recorded in one participant.

Data acquisition. Scanning was performed on a 3T Siemens Trio Tim MRI scanner at the Center for Bio-Medical Imaging, University Hospital of Geneva. For functional scans, a single-shot echoplanar imaging sequence was used (T2* weighted, gradient echo sequence), with the following parameters: TR (repetition time) = 2490 ms, TE (echo time) = 30 ms, 36 off-axial slices, voxel dimensions: 1.8 × 1.8 mm, 3.6 mm slice thickness (no gap). Anatomical images were acquired using a T1-weighted sequence, with parameters: TR/TE: 2200 ms/3.45 ms; slice thickness = 1 mm; in-plane resolution: 1 × 1 mm.

Data preprocessing. Data were analyzed using the AFNI software package and MATLAB (The MathWorks). All functional data were slice-time corrected, motion corrected, and low-frequency drifts were removed with a temporal high-pass filter (cutoff of 0.006 Hz). Data were spatially smoothed with a Gaussian kernel (4 mm full-width half-maximum). The

three-dimensional anatomical scans were transformed into Talairach space, and the parameters from this transformation were subsequently applied to the coregistered functional data, which were resampled to 3 mm isotropic voxels. All analyses were restricted to voxels that fell within the anatomical brain volume.

fMRI data analysis. Information-based brain mapping by means of a spherical searchlight was used to analyze the fMRI data (Kriegeskorte et al., 2006). This technique exploits differences in multivoxel activity patterns to determine which brain areas differentiate among the experimental conditions. Multivoxel pattern analysis was performed on parameter estimates (βs), which were extracted using standard regression analysis [generalized linear model (GLM)]. For the GLM, events were defined as the 4 s period starting with the onset of the movie or sound clip, and thus excluding the part of the trial where the response screen was shown (a control analysis was performed with conditions modeled relative to the response screen; see Results). These events were convolved with a standard model of the hemodynamic response function. A general linear model was created with one predictor for each condition of interest. Regressors of no interest were also included to account for differences in mean magnetic resonance (MR) signal across scans and participants. Regressors were fitted to the MR time-series in each voxel and the resulting β parameter estimates were used to estimate the magnitude of response to the experimental conditions.

For each voxel in the brain we extracted, for each stimulus condition, the multivoxel pattern of activation in a sphere of 8 mm radius (corresponding to 81 resampled voxels) around this voxel. In other words, each of the 15 conditions (5 emotions × 3 modalities) was associated with a unique multivoxel pattern of βs in each sphere. These 15 patterns were then correlated with each other, resulting in a symmetrical 15 × 15 correlation matrix. These correlations reflect the similarity between the activation patterns of each pair of experimental conditions. The correlation values from each sphere were Fisher transformed ($0.5 * \log_e((1 + r)/(1 - r))$) and assigned to the center voxel of the sphere. To test for modality-independent emotion representations, we analyzed the correlations that resulted from the pairing of the different modalities (face–voice, body–voice, body–face). For the whole-brain searchlight analysis, the within-emotion correlations (e.g., fearful body—fearful voice) were averaged to provide one estimate of within-emotion similarity. Similarly, the between-emotion correlations (e.g., fearful body—happy voice) were averaged to provide one estimate of between-emotion similarity. Finally, a random-effects whole-brain group analysis (N = 18) was performed using spatially normalized data to test for spheres with significantly stronger correlations within emotion categories than between emotion categories. Importantly, the sensory dissimilarity between these two comparisons is identical, and any significant differences between these correlations were therefore due to the emotional state perceived from the stimuli.

The appropriate statistical threshold was determined using Monte Carlo simulation. 100 simulations were performed on the same data using the same searchlight analysis as described above, but with the 15 condition labels shuffled before computing the correlations. At an uncorrected threshold of $p < 0.001$ and a cluster-size threshold of 20 (resampled) contiguous voxels, 100 simulations revealed a total of 2 significant activation clusters, conservatively estimating the false-positive probability (2/100) at $p < 0.05$.

Results

Behavioral results

Participants rated the intensity of the perceived emotions on a 3-point scale, ranging from 1 ("a little") to 3 ("very much"). The average rating for the five emotion conditions was 2.24 (Fig. 2), and comparable for the three modalities (face: 2.13, body: 2.31; voice: 2.27). A 5 (Emotion) × 3 (Modality) ANOVA on these ratings showed a significant interaction between Emotion and Modality ($p < 0.001$), indicating that differences between the perceived intensities of emotions depended on modality. For example, disgust was perceived as more intense in the body condition (2.44) than the voice condition (1.95; $p <$
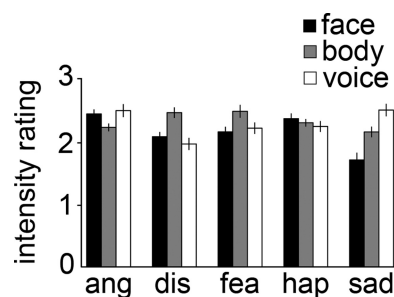

Figure 2. Average intensity ratings for each experimental condition. ang, Anger; dis, disgust; fea, fear; hap, happiness; sad, sadness.

0.005). Conversely, anger was perceived as more intense in the voice (2.48) than the body condition (2.21; $p < 0.01$). The Emotion × Modality interaction indicates that differences in the perceived intensity of the emotions were not consistent across modality, and thus that any supramodal emotion-specific effects in subsequent fMRI analyses were unlikely to be due to systematic differences in perceived intensity. To confirm the absence of systematic emotion-related differences in perceived intensity, the emotion-related variations in the ratings were correlated across modalities. If some emotions were consistently (i.e., for multiple modalities) rated as more or less intense than the other emotions, this would be reflected in positive correlations between the ratings for different modalities. Correlations were computed for each participant separately, Fisher transformed, and tested against zero. None of the correlations were significantly positive (body–face: $r = 0.15$, $p > 0.2$; body–voice: $r = -0.29$, $p = 0.07$; face–voice: $r = 0.08$, $p > 0.6$). In summary, there was no evidence for emotion-specific variations in intensity ratings that were consistent across modalities. This excludes the possibility that differences in perceived intensity (or other rating-related differences such as the planning and execution of motor responses) could provide an alternative explanation for supramodal emotion-specific fMRI responses.

fMRI results

The searchlight analysis revealed two clusters that showed significant supramodal emotion information ($p_{corr} < 0.05$, corrected for multiple comparisons; see Materials and Methods). The first was located in rostral MPFC [peak coordinates (x, y, z): 11, 48, 17; $t_{17} = 6.0$, $p = 0.00001$] (Fig. 3). The second cluster was located in left posterior superior temporal sulcus [STS; peak coordinates (x, y, z): $-47, -62, 8$, $t_{17} = 6.4$, $p < 0.00001$; Fig. 3]. In both areas, the patterns of neural activity associated with the same emotions perceived from different, nonoverlapping, stimulus types (body, face, voice) were more similar to each other than were the patterns associated with different emotions, again perceived from different stimulus types (Fig. 3; supplemental Fig. 1, available at www.jneurosci.org as supplemental material). In other words, these regions showed remarkably consistent activity patterns that followed the emotional content that was inferred from the stimuli independent of the specific sensory cues that conveyed the emotion. Finally, ANOVAs revealed that in both MPFC and STS, the strength of this effect was equally strong for the three modality comparisons (face–voice, body–voice, body–face; $p > 0.1$, for both regions), and the five emotion categories ($p > 0.07$, for both regions).

To test whether supramodal emotion information (again defined as the average within-emotion across-modality correlation versus the average between-emotion across-modality correla-
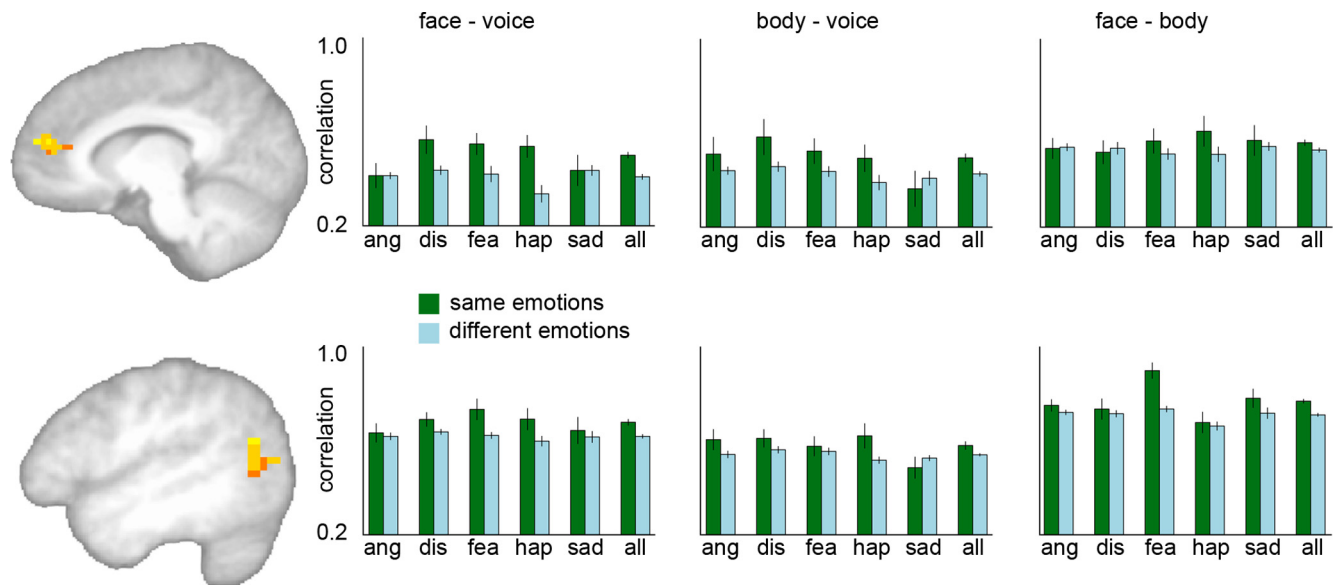
Figure 3. Results of a whole-brain searchlight analysis showing clusters with significant emotion-specific activity patterns across modality. Similarity of activity patterns was expressed as a correlation value, with higher correlations indicating higher similarity. Patterns were generally more similar within emotion categories (green bars) than between emotion categories (blue bars). Top row: Cluster in MPFC; peak (x, y, z): 11, 48, 17; $t_{17} = 6.0$, $p = 0.00001$, and average Fisher-transformed correlation values in MPFC cluster for the three cross-modality comparisons. Correlations were higher for within- than between-emotion comparisons for all three modality combinations ($p < 0.05$, for all tests). Bottom row: Cluster in left STS; peak (x, y, z): −47, −62, 8, $t_{17} = 6.4$, $p < 0.00001$, and average Fisher-transformed correlation values in STS cluster for the three cross-modality comparisons. Correlations were higher for within- than between-emotion comparisons for all three modality combinations ($p < 0.05$, for all tests). ang, Anger; dis, disgust; fea, fear; hap, happiness; sad, sadness; all, average across the five emotions. Error bars indicate within-subject SEM.

tion) in the peak spheres of MPFC and STS were modulated by the perceived intensity of the emotions, the multivoxel pattern analysis was repeated on two halves of the data, divided based on the participants' intensity ratings. For each participant and stimulus condition separately, the trial that was rated the lowest in each run was assigned to the "low rated" condition whereas the trial that was rated the highest was assigned to the "high rated" condition. (Each condition (e.g., fearful bodies) appeared twice in a run, expressed by different actors, so that high- and low-rated trials appeared with a similar history and distribution over the entire experiment.) When both trials were rated equally (or if no response was collected for one or both trials) the trials were randomly assigned to the conditions, which was true for, on average, 53% of the trials. The mean intensity rating of the high-rated condition (2.46) was indeed substantially higher than the mean intensity rating of the low-rated condition (1.85). Despite this difference in perceived intensity, the difference between the supramodal emotion information in multivoxel activity patterns for high-rated versus low-rated trials was not significant in either region ($p > 0.3$, for both spheres), indicating that both the low- and high-rated trials contributed to the overall emotion information observed. However, it should be emphasized that our stimulus set was purposefully chosen to a have relatively "high" expressivity and thus be reliably recognized, such that the overall variability in ratings might have been insufficient to reveal intensity-related differences in this subsidiary analysis.

In the above analyses, conditions were modeled relative to stimulus presentation, excluding the part of the trial where the response screen was shown. However, given the close temporal proximity of stimulus and response screen, it cannot be excluded that the emotion-specific activity patterns observed may have been partly related to the processing of the response screen (e.g., related to the written labels of the emotions). To test for this possibility, we ran a control analysis, now specifically modeling the conditions relative to the 2.5 s presentation of the response

screen. The same contrast that revealed strong emotion-specific information when conditions were modeled relative to stimulus presentation (Fig. 3) failed to reveal any significant clusters when conditions were modeled relative to response-screen presentation. This was true even when the threshold was lowered to $p < 0.01$ (uncorrected for multiple comparisons). Furthermore, the peak spheres in MPFC and STS, which showed highly significant emotion-specific information during stimulus presentation, did not contain any such information during the presentation of the response screen ($p > 0.3$, for both spheres). These analyses show that the effects in MPFC and STS were related to the processing of the emotional body, face, and voice stimuli rather than to the processing of the response screen.

Mean activation in searchlight clusters

Mean activation levels were extracted from the searchlight clusters in MPFC and STS. None of the stimulus conditions activated the clusters significantly above baseline (Fig. 4), which is commonly observed due to high activation in these regions during resting states (Mitchell et al., 2002). Nevertheless, in both regions, a 5 (Emotion) × 3 (Modality) ANOVA on parameter estimates of activation showed a significant interaction between Emotion and Modality ($p < 0.05$, for both regions), indicating that the activation differences between the emotions differed for the 3 modalities. To test whether there were emotion-specific activation differences that were consistent across modality, we tested whether the mean activations (as estimated by β values) in MPFC and STS were more similar within than between emotion categories. First, for each participant, the response magnitudes in the three modalities were mean-centered such that the mean responses to the three modalities were identical, leaving only emotion-related variability. Subsequently, the absolute difference between responses to the same emotion across modalities (e.g., fearful faces vs fearful bodies) were computed, and compared with the absolute difference between responses to different
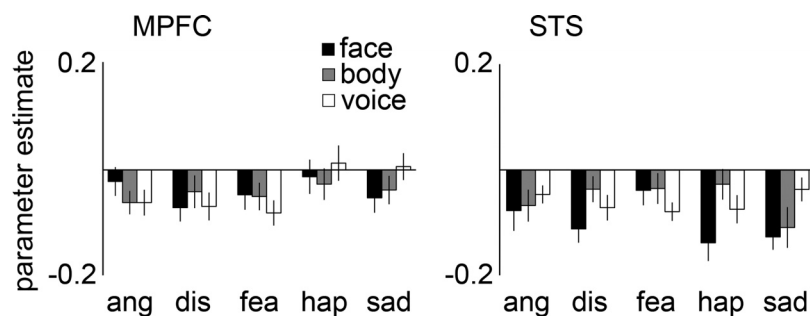
Figure 4. Mean activation (parameter estimates) in MPFC (left) and STS (right) searchlight clusters for each experimental condition. Abbreviations are as in Figure 2.

Table 1. Regions showing univariate emotion-selective responses across the 3 modalities (F + B + V), and significance for the same contrasts within these regions for each of the modalities separately

	Talairach coordinates	F + B + V	Face	Body	Voice
Disgust vs other emotions					
Brainstem	1, −35, −24	$p < 0.001$	$p < 0.001$	NS	$p < 0.01$
Fear vs other emotions					
Left MFG	−43, 14, 38	$p < 0.001$	$p < 0.05$	$p < 0.05$	$p < 0.01$
Happiness vs other emotions					
Right ACC	5, 43, −4	$p < 0.001$	NS	$p < 0.005$	$p < 0.05$
Right Occipital	18, −92, 2	$p < 0.001$	NS	$p < 0.001$	NS
Right ATL	38, 1, −21	$p < 0.001$	NS	$p < 0.05$	$p < 0.05$

F + B + V, Face + body + voice; MFG, middle frontal gyrus; ACC, anterior cingulate cortex; ATL, anterior temporal lobe.

emotions across modalities (e.g., fearful faces vs disgusted bodies). If the mean β values carried emotion-specific information, the average within-emotion response difference should be smaller than the average between-emotion response difference. This was, however, not the case. The average within-emotion response difference was equivalent to the average between-emotion response difference for all three modality pairs in both the MPFC and STS ($p > 0.05$, for all tests). Thus, in contrast to the multivariate measure (multivoxel correlation), which carried reliable emotion-specific information across modality, the univariate measure (based on mean response amplitude) did not provide such modality-independent emotion information.

Overlap of STS searchlight cluster with face- and body-responsive STS

The low overall activity in the left STS cluster (Fig. 4) suggests that it may differ from those STS regions previously found to show strong stimulus-driven responses to moving faces and bodies (for review, see Allison et al., 2000). To directly test for any overlap between face- and body-responsive voxels in left STS and the searchlight STS cluster showing emotion-specific patterns [peak coordinates (x, y, z): −47, −62, 8] we contrasted the face and body conditions against baseline. Faces activated a cluster in left STS (−62, −34, 2, $t_{17} = 7.5$, $p < 0.00001$) that was anterior to the searchlight STS cluster. Bodies activated two regions in left STS (posterior: −50, −50, 4, $t_{17} = 5.4$, $p < 0.0001$; anterior: −53, −35, −1, $t_{17} = 4.8$, $p < 0.0005$) that were also both located anterior to the searchlight cluster. Finally, there was no overlap between the searchlight STS cluster and any of the face- or body-responsive clusters, even at an uncorrected threshold of $p < 0.01$.

This finding suggests that the supramodal emotion-specific STS area identified by the searchlight analysis did not correspond to movement-selective regions observed in studies of face or body perception.

Univariate supramodal emotion specificity

A whole-brain univariate group analysis was performed to test for regions in which overall activity was selective for one of the five emotions independent of modality. Each of the 5 emotions was contrasted with the average of the 4 others, initially combining the data from the 3 modalities. At $p < 0.001$ (uncorrected, cluster size >20 voxels), no regions were selectively activated by any of the 5 emotions (averaged across the 3 modalities). At a more lenient threshold of $p < 0.005$ (uncorrected, cluster size >10 voxels), several regions showed emotion-specific activity for disgust, fear, or happiness (Table 1). Next, we tested whether the emotion specificity in these regions was driven by one of the modalities, or was independent of modality and significant for each of the 3 modalities separately. Of these regions, only the fear-specific region in left middle frontal gyrus showed significant emotion specificity (at $p < 0.05$) for all three modalities separately (Table 1). The absence of amygdala activity for the contrast 'fear vs other emotions' is in line with previous work showing that this region responds to multiple emotion categories when these are matched for perceived intensity (Winston et al., 2003). These data therefore indicate that category-specific modality-independent activations could not be reliably observed in standard univariate subtractive analysis, with the possible exception of fear-specificity in left middle frontal gyrus.

Discussion

Our results show that multivoxel patterns of activity in two cortical brain regions, MPFC and left STS, carried information about emotion categories regardless of the specific sensory cues (body, face, voice). No consistent differences were found between the perceived intensity of different emotion categories, ruling out the possibility that the emotion-specific activity patterns in MPFC and STS were related to systematic differences in the intensity of the different emotions. Furthermore, there were no consistent differences between emotion categories in the average magnitude of activity in these regions, indicating a similar overall recruitment of MPFC and STS in processing the different types of emotions. Critically, activity patterns in MPFC and STS distinguished between the emotion categories even across different visual cues (body versus face) and different sensory modalities (body/face versus voice). These findings therefore provide evidence for modality-independent but emotion-specific representations in MPFC and STS, regions previously implicated in affective processing, mental state attribution, and theory-of-mind. Below, we discuss the implications of these new findings in the context of previous research, and propose questions for future research.

Converging evidence points to an important role for MPFC in mental state attribution, including emotions (Frith and Frith, 2003; Mitchell, 2009; Van Overwalle and Baetens, 2009). For example, MPFC activity has been reported during mentalizing tasks involving the attribution of beliefs or intentions to protagonists in stories and cartoons (Fletcher et al., 1995; Gallagher et al., 2000; Gallagher and Frith, 2003; Völlm et al., 2006), the incorporation of the thinking of others in strategic reasoning (Coricelli

and Nagel, 2009), and the attribution of emotions to seen faces (Schulte-Rüther et al., 2007). MPFC is reliably activated in studies of emotion perception using different stimulus types, as well as in studies of emotion experience (Bush et al., 2000; Kober et al., 2008; Adolphs, 2009), particularly when participants evaluate emotional states (Lane et al., 1997; Reiman et al., 1997; Lee and Siegle, 2009; Zaki et al., 2009). Finally, activity in MPFC was found to be directly related to the accuracy of attributions made about another person's emotions, suggesting an important role for this region in the evaluation and understanding of affective states (Zaki et al., 2009). Our present finding of supramodal emotion representations in MPFC provides new evidence that this region is an important part of the brain network involved in the recognition, categorization, and understanding of others' emotions. Critically, the present study is the first to show that MPFC encodes information related to the particular content of the emotional state observed in others. This finding speaks to a principal question regarding the possible role of MPFC in mental state attribution: whether this region maintains specific representations of mental states or, instead, whether its activity during mental state attribution reflects content-general processes. For example, it has been hypothesized that the key function of MPFC may be captured by particular cognitive processes such as directing attention to one's own or others' mental states (Gallagher and Frith, 2003), the evaluation of internally generated information (Christoff and Gabrieli, 2000), or projecting oneself into worlds that differ mentally, temporally, or physically from one's current experience (Mitchell, 2009). One interpretation of these process-oriented accounts is that MPFC may activate whenever the hypothesized process (e.g., directing attention to mental states) is implicated, for example to access representations of emotions stored in other brain regions. While at a macroscopic level the function of MPFC may be described in terms of particular cognitive processes, the current finding of content-specific MPFC activity patterns suggests that at a finer-grained level this region may contain distinct representations corresponding to the content used in these processes. More generally, the macroscopic organization of cortical areas may follow particular processes or computations, while the specific content involved in these computations may be represented at a finer-grained level of cortical organization within these regions (Peelen et al., 2006; Mur et al., 2009). The recent interest and development in high-resolution fMRI and multivariate analysis methods (Haynes and Rees, 2006; Norman et al., 2006) can be expected to provide increasing evidence for such fine-grained representations within larger-scale functional brain areas.

A second brain area showing supramodal emotion specificity in the present study was the left posterior STS. Similar to MPFC, multivoxel activity patterns in this region carried information about perceived emotional states. Interestingly, previous research has linked the posterior STS, along with the MPFC, to mentalizing and "theory of mind". For example, patients with selective damage to left STS and left temporoparietal junction (TPJ) are impaired in understanding the mental states of others (Samson et al., 2004), indicating that this region may be critically involved in this process. Several fMRI studies have also implicated the posterior STS/TPJ in the understanding of others' minds (Gallagher et al., 2000; Saxe and Kanwisher, 2003; Decety and Grezès, 2006; Gobbini et al., 2007). This region seems functionally similar to the MPFC in that it activates when reasoning about others' mental states, but is also functionally dissimilar in that it is not recruited when reasoning about one's own mental states (Saxe et al., 2006). In accord with these notions, the left STS

region showing significant supramodal emotion specificity in our study appears to overlap with regions previously implicated in mentalizing and theory-of-mind (cf. Box 1 in Decety and Grezès, 2006). By contrast, the left STS identified here appears to differ from more anterior parts of STS that contain multisensory sectors activated by voices and faces, as well as their combined presentation (Beauchamp et al., 2004; Kreifelts et al., 2009). Likewise, other studies found selective responses in (mostly right) STS to biological movements, such as facial expressions, body movements, and point-light biological motion displays (Puce et al., 1998; Grezès et al., 1999; Allison et al., 2000; Grossman et al., 2000; Haxby et al., 2000; Vaina et al., 2001; Peelen et al., 2006; Engell and Haxby, 2007; Gobbini et al., 2007). Although the emotion-specific STS cluster in our study was located nearby these regions implicated in "social vision", it appeared to be distinct from them: first, the overall magnitude of BOLD responses to body and face movie clips in this cluster were not significantly above baseline, contrary to what is typically found in the face- and body-selective STS regions; second, directly contrasting the face and body conditions with baseline gave significant STS activation in regions anterior to the emotion-specific searchlight STS cluster and these regions did not overlap. However, it should be noted that to directly establish the anatomical and functional correspondence of the current STS region with regions previously implicated in the processing of body and face stimuli, and with those implicated in mentalizing and theory-of-mind, requires comparisons in the same group of participants, and using the same analysis methods (Gobbini et al., 2007; Bahnemann et al., 2010).

The current findings showed a high degree of emotion specificity in multivoxel fMRI patterns in MPFC and left STS. What could be the degree of specificity at the neuronal level? One possibility consistent with our findings is that these regions contain individual neurons that selectively represent emotion categories at an abstract level. A comparable degree of specificity for other categories has been found in the human medial temporal lobe (MTL) using intracranial recordings. For example, the activity of neurons in MTL selectively correlates with abstract concepts (e.g., "Marilyn Monroe") independent of how these concepts are cued (Quiroga et al., 2005; Quian Quiroga et al., 2009). It is conceivable that neurons in MPFC and/or STS may show a similarly sparse tuning for discrete emotion categories. Such tuning may be limited to the basic emotion categories used here (Oatley and Johnson-Laird, 1987; Ekman, 1992), or may possibly extend to other emotions (e.g., guilt, contempt, shame), nonemotional mental states, or even conceptual representations more generally (Binder et al., 2009). Alternatively, it is possible that neurons in MPFC and/or STS may code for particular emotion dimensions (Russell, 1980), specific emotion components (novelty, pleasantness, relevance, etc.) (Scherer, 1984), or action tendencies associated with emotions (Frijda, 1987). Future experiments are needed to dissociate between these scenarios (Anderson et al., 2003), which would in turn help to test the predictions of different psychological theories of emotion.

Our findings open up many new questions that can be directly addressed by future work using the same methodology. First, do MPFC and/or STS represent other (emotional and nonemotional) mental states than those tested in our study? Second, to what extent is emotion specificity in these areas related to the explicit evaluation of the emotions? For example, does emotion specificity persist when stimuli are presented outside the focus of attention (Vuilleumier et al., 2001), or when emotion is task irrelevant (Critchley et al., 2000; Winston et al., 2003)? Finally, are MPFC and/or STS similarly activated by perceived and experi-

enced emotions? That is, is there an overlap in the representation of perceived and self-experienced emotions (Reiman et al., 1997), such that evaluating another's emotion and experiencing this same emotion are reflected in corresponding fMRI activity patterns in these regions? These are outstanding questions with broad theoretical and philosophical implications that can now be addressed.

References

Adolphs R (2009) The social brain: neural basis of social knowledge. Annu Rev Psychol 60:693–716.

Allison T, Puce A, McCarthy G (2000) Social perception from visual cues: role of the STS region. Trends Cogn Sci 4:267–278.

Anderson AK, Christoff K, Stappen I, Panitz D, Ghahremani DG, Glover G, Gabrieli JD, Sobel N (2003) Dissociated neural representations of intensity and valence in human olfaction. Nat Neurosci 6:196–202.

Atkinson AP, Dittrich WH, Gemmell AJ, Young AW (2004) Emotion perception from dynamic and static body expressions in point-light and full-light displays. Perception 33:717–746.

Atkinson AP, Tunstall ML, Dittrich WH (2007) Evidence for distinct contributions of form and motion information to the recognition of emotions from body gestures. Cognition 104:59–72.

Bahnemann M, Dziobek I, Prehn K, Wolf I, Heekeren HR (2010) Sociotopy in the temporoparietal cortex: common versus distinct processes. Soc Cogn Affect Neurosci 5:48–58.

Banse R, Scherer KR (1996) Acoustic profiles in vocal emotion expression. J Pers Soc Psychol 70:614–636.

Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A (2004) Unraveling multisensory integration: patchy organization within human STS multisensory cortex. Nat Neurosci 7:1190–1192.

Belin P, Fillion-Bilodeau S, Gosselin F (2008) The Montreal Affective Voices: a validated set of nonverbal affect bursts for research on auditory affective processing. Behav Res Methods 40:531–539.

Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb Cortex 19:2767–2796.

Bush G, Luu P, Posner MI (2000) Cognitive and emotional influences in anterior cingulate cortex. Trends Cogn Sci 4:215–222.

Christoff K, Gabrieli JD (2000) The frontopolar cortex and human cognition: evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. Psychobiology 28:168–186.

Coricelli G, Nagel R (2009) Neural correlates of depth of strategic reasoning in medial prefrontal cortex. Proc Natl Acad Sci U S A 106:9163–9168.

Critchley H, Daly E, Phillips M, Brammer M, Bullmore E, Williams S, Van Amelsvoort T, Robertson D, David A, Murphy D (2000) Explicit and implicit neural mechanisms for processing of social information from facial expressions: a functional magnetic resonance imaging study. Hum Brain Mapp 9:93–105.

Decety J, Grezès J (2006) The power of simulation: imagining one's own and other's behavior. Brain Res 1079:4–14.

de Gelder B, Meeren HK, Righart R, van den Stock J, van de Riet WA, Tamietto M (2006) Beyond the face: exploring rapid influences of context on face processing. Prog Brain Res 155:37–48.

Ekman P (1992) Are there basic emotions? Psychol Rev 99:550–553.

Engell AD, Haxby JV (2007) Facial expression and gaze-direction in human superior temporal sulcus. Neuropsychologia 45:3234–3241.

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

Fletcher PC, Happé F, Frith U, Baker SC, Dolan RJ, Frackowiak RS, Frith CD (1995) Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. Cognition 57:109–128.

Formisano E, De Martino F, Bonte M, Goebel R (2008) "Who" is saying "what"? Brain-based decoding of human voice and speech. Science 322:970–973.

Frijda NH (1987) Emotion, cognitive structure, and action tendency. Cogn Emot 1:115–143.

Frith U, Frith CD (2003) Development and neurophysiology of mentalizing. Philos Trans R Soc Lond B Biol Sci 358:459–473.

Gallagher HL, Frith CD (2003) Functional imaging of 'theory of mind.' Trends Cogn Sci 7:77–83.

Gallagher HL, Happé F, Brunswick N, Fletcher PC, Frith U, Frith CD (2000) Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. Neuropsychologia 38:11–21.

Gobbini MI, Koralek AC, Bryan RE, Montgomery KJ, Haxby JV (2007) Two takes on the social brain: a comparison of theory of mind tasks. J Cogn Neurosci 19:1803–1814.

Grezès J, Costes N, Decety J (1999) The effects of learning and intention on the neural network involved in the perception of meaningless actions. Brain 122:1875–1887.

Grossman E, Donnelly M, Price R, Pickens D, Morgan V, Neighbor G, Blake R (2000) Brain areas involved in perception of biological motion. J Cogn Neurosci 12:711–720.

Haxby JV, Hoffman EA, Gobbini MI (2000) The distributed human neural system for face perception. Trends Cogn Sci 4:223–233.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293:2425–2430.

Haynes JD, Rees G (2005) Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat Neurosci 8:686–691.

Haynes JD, Rees G (2006) Decoding mental states from brain activity in humans. Nat Rev Neurosci 7:523–534.

Heberlein AS, Atkinson AP (2009) Neuroscientific evidence for simulation and shared substrates in emotion recognition: beyond faces. Emotion Rev 1:162–177.

Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. Nat Neurosci 8:679–685.

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

Kreifelts B, Ethofer T, Shiozawa T, Grodd W, Wildgruber D (2009) Cerebral representation of non-verbal emotional perception: fMRI reveals audio-visual integration area between voice- and face-sensitive regions in the superior temporal sulcus. Neuropsychologia 47:3059–3066.

Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. Proc Natl Acad Sci U S A 103:3863–3868.

Kriegeskorte N, Formisano E, Sorger B, Goebel R (2007) Individual faces elicit distinct response patterns in human anterior temporal cortex. Proc Natl Acad Sci U S A 104:20600–20605.

Lane RD, Fink GR, Chau PM, Dolan RJ (1997) Neural activation during selective attention to subjective emotional responses. Neuroreport 8:3969–3972.

Lee KH, Siegle GJ (2009) Common and distinct brain networks underlying explicit emotional evaluation: a meta-analytic study. Soc Cogn Affect Neurosci. Advance online publication. Retrieved March 6, 2009. doi:10.1093/scan/nsp001.

Magnée MJ, Stekelenburg JJ, Kemner C, de Gelder B (2007) Similar facial electromyographic responses to faces, voices, and body expressions. Neuroreport 18:369–372.

Mitchell JP (2009) Inferences about mental states. Philos Trans R Soc Lond B Biol Sci 364:1309–1316.

Mitchell JP, Heatherton TF, Macrae CN (2002) Distinct neural systems subserve person and object knowledge. Proc Natl Acad Sci U S A 99:15238–15243.

Mur M, Bandettini PA, Kriegeskorte N (2009) Revealing representational content with pattern-information fMRI—an introductory guide. Soc Cogn Affect Neurosci 4:101–109.

Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn Sci 10:424–430.

Oatley K, Johnson-Laird PN (1987) Towards a cognitive theory of emotions. Cogn Emot 1:29–50.

Peelen MV, Wiggett AJ, Downing PE (2006) Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. Neuron 49:815–822.

Peelen MV, Atkinson AP, Andersson F, Vuilleumier P (2007) Emotional modulation of body-selective visual areas. Soc Cogn Affect Neurosci 2:274–283.

Peelen MV, Fei-Fei L, Kastner S (2009) Neural mechanisms of rapid natural scene categorization in human visual cortex. Nature 460:94–97.

Puce A, Allison T, Bentin S, Gore JC, McCarthy G (1998) Temporal cortex activation in humans viewing eye and mouth movements. J Neurosci 18:2188–2199.

Quian Quiroga R, Kraskov A, Koch C, Fried I (2009) Explicit encoding of multimodal percepts by single neurons in the human brain. Curr Biol 19:1308–1313.

Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I (2005) Invariant visual representation by single neurons in the human brain. Nature 435: 1102–1107.

Reiman EM, Lane RD, Ahern GL, Schwartz GE, Davidson RJ, Friston KJ, Yun LS, Chen K (1997) Neuroanatomical correlates of externally and internally generated human emotion. Am J Psychiatry 154:918–925.

Russell JA (1980) A circumplex model of affect. J Pers Soc Psychol 39:1161–1178.

Samson D, Apperly IA, Chiavarino C, Humphreys GW (2004) Left temporoparietal junction is necessary for representing someone else's belief. Nat Neurosci 7:499–500.

Saxe R, Kanwisher N (2003) People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind." Neuroimage 19:1835–1842.

Saxe R, Moran JM, Scholz J, Gabrieli J (2006) Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. Soc Cogn Affect Neurosci 1:229–234.

Scherer KR (1984) On the nature and function of emotion: a component process approach. In: Approaches to emotion (Scherer KR, Ekman P, eds), pp 293–317. Hillsdale, NJ: Erlbaum.

Scherer KR, Ellgring H (2007) Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? Emotion 7:113–130.

Schulte-Rüther M, Markowitsch HJ, Fink GR, Piefke M (2007) Mirror neuron and theory of mind mechanisms involved in face-to-face interactions: a functional magnetic resonance imaging approach to empathy. J Cogn Neurosci 19:1354–1372.

Vaina LM, Solomon J, Chowdhury S, Sinha P, Belliveau JW (2001) Functional neuroanatomy of biological motion perception in humans. Proc Natl Acad Sci U S A 98:11656–11661.

Van Overwalle F, Baetens K (2009) Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. Neuroimage 48: 564–584.

Völlm BA, Taylor AN, Richardson P, Corcoran R, Stirling J, McKie S, Deakin JF, Elliott R (2006) Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. Neuroimage 29:90–98.

Vuilleumier P, Armony JL, Driver J, Dolan RJ (2001) Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. Neuron 30:829–841.

Winston JS, O'Doherty J, Dolan RJ (2003) Common and distinct neural responses during direct and incidental processing of multiple facial emotions. Neuroimage 20:84–97.

Zaki J, Weber J, Bolger N, Ochsner K (2009) The neural bases of empathic accuracy. Proc Natl Acad Sci U S A 106:11382–11387.