

Cognitive and affective theory of mind share the same local patterns of activity in posterior temporal but not medial prefrontal cortex

Corrado Corradi-Dell'Acqua,^{1,2} Christoph Hofstetter,^{1,2} and Patrik Vuilleumier^{1,2}

¹Swiss Centre for Affective Sciences, University of Geneva, CH-1205 Geneva, Switzerland and ²Laboratory for Neurology and Imaging of Cognition, Department of Neurosciences and Clinic of Neurology, University Medical Center, CH-1211 Geneva, Switzerland

Understanding emotions in others engages specific brain regions in temporal and medial prefrontal cortices. These activations are often attributed to more general cognitive ‘mentalizing’ functions, associated with theory of mind and also necessary to represent people’s non-emotional mental states, such as beliefs or intentions. Here, we directly investigated whether understanding emotional feelings recruit similar or specific brain systems, relative to other non-emotional mental states. We used functional magnetic resonance imaging with multivoxel pattern analysis in 46 volunteers to compare activation patterns in theory-of-mind tasks for emotions, relative to beliefs or somatic states accompanied with pain. We found a striking dissociation between the temporoparietal cortex, that exhibited a remarkable voxel-by-voxel pattern overlap between emotions and beliefs (but not pain), and the dorsomedial prefrontal cortex, that exhibited distinct (and yet nearby) patterns of activity during the judgment of beliefs and emotions in others. Pain judgment was instead associated with activity in the supramarginal gyrus, middle cingulate cortex and middle insular cortex. Our data reveal for the first time a functional dissociation within brain networks sub-serving theory of mind for different mental contents, with a common recruitment for cognitive and affective states in temporal regions, and distinct recruitment in prefrontal areas.

Keywords: mentalizing; social cognition; multivariate pattern analysis; perspective taking; emotional coding

INTRODUCTION

How we understand emotions felt by other people is a central but still unresolved question in cognitive-affective neuroscience. Many studies used brain imaging techniques to explore the neural foundations of this ability, and suggested that they might rely on brain structures involved in ‘theory of mind’ (ToM), i.e. cognitive processes mediating the representations of others’ mental states, like beliefs or goals (Saxe *et al.*, 2004; Amodio and Frith, 2006). Accordingly, portions of the dorsomedial prefrontal cortex (DMPFC), and posterior temporal cortex—including the temporoparietal junction (TPJ) and the superior temporal sulcus (STS)—were found active not only when participants judged the emotional state of others (Zaki *et al.*, 2009; Peelen *et al.*, 2010) but also when they evaluated mental states unrelated to emotions, such as beliefs or desires (Gallagher and Frith, 2003; Mar, 2011). These observations are in favor of a mentalistic (or representational, Flavell, 1999; Saxe *et al.*, 2004) interpretation of affect attribution, according to which emotions of other people are not represented exclusively in terms on their bodily manifestations (smiles, tears, shivers, etc.) but also as particular states of mind.

However, although mentalistic interpretations of emotional processing are plausible and parsimonious, they remain open to several criticisms. First, TPJ and DMPFC exhibit functional properties that are not limited to mentalizing and emotional processing, but also encompass attention, agency, self-reflection, visuospatial perspective taking, decision-making, etc. (Amodio and Frith, 2006; Decety and Lamm, 2007; Corbetta *et al.*, 2008; Corradi-Dell'Acqua *et al.*, 2008, 2013). This

heterogeneity might reflect a common process shared across tasks, or result from the crude spatial resolution of functional magnetic resonance imaging (fMRI), which maps neural signals originating from independent neuronal populations within overlapping cortical areas. In addition, these regions have no well-defined boundaries, and the overlap of activations across conditions or studies is often partial. Therefore, previous work could not distinguish whether activations in DMPFC and TPJ when assessing others’ emotion originate from the same neuronal source responsible for mentalizing, or from independent nearby sources.

Furthermore, developmental research has documented a clear temporal dissociation between the emergence of the ability to infer others’ cognitive vs. emotional states. In these studies, ToM was usually tested by the *false beliefs* task, in which participants read stories describing a state of the world that changes without the protagonist knowing about the change (e.g. ‘Virginie put her chocolate away in the cupboard and went outside. Marc moved the chocolate from the cupboard into the fridge. Later, Virginie came back’). The task probes for the participants’ awareness that the protagonist’s representation of the world (‘Virginie believes that the chocolate is in the cupboard’) is different from the true state (‘the chocolate is in the fridge’) (Saxe *et al.*, 2004). Under explicit request, children younger than four are unable to make this distinction (but see Baillargeon *et al.*, 2010; Kovács *et al.*, 2010), although they can already recognize others’ emotions and pain (Flavell, 1999; Saxe *et al.*, 2004). This dissociation suggests that cognitive ToM abilities may not be necessary (at least in children) for assessing people’s affective states.

In our study, we aimed to unveil the role of TPJ and DMPFC in the representation of others’ emotions (Peelen *et al.*, 2010) and to elucidate the nature of their functional overlap with representations of non-emotional mental states such as beliefs. We measured brain activity in 46 female volunteers while they read brief stories and subsequently answered questions about the protagonists’ emotional state. Cortical activations observed in this emotional task [E] were compared with: a false belief task, requiring the judgment of the protagonists’

Received 20 August 2012; Revised 8 April 2013; Accepted 11 June 2013
Advance Access publication 14 June 2013

The authors thank Catherine Hynes and Rebecca Saxe for making their experimental materials available, as well as Francesco Turri and David Lawi for their assistance in translating all experimental materials into French. This research was supported by the National Center of Competence in Research (NCCR) for Affective Sciences financed by the Swiss National Science Foundation (no. 51NF40-104897) and hosted by the University of Geneva, plus an award of the Geneva Academic Society to P.V.

Correspondence should be addressed to Corrado Corradi-Dell'Acqua, Swiss Centre for Affective Sciences, University of Geneva, 7 rue des Batoirés, 1205 Geneva, Switzerland. E-mail: corrado.corradi@unige.ch

beliefs [B]; a pain judgment task [Pa], requiring the judgment of the protagonists' aching sensations; and a false photograph control task [Ph], involving the assessment of stories about physical events without human protagonists (high-level control) (Saxe and Kanwisher, 2003; Hynes *et al.*, 2006; Saxe and Powell, 2006).

Our key question was whether brain responses associated with emotional judgment behaved more similarly to beliefs than to other human non-representational states (pain). Critically, we examined the similarity of cortical activations by employing multivoxel pattern analysis (MVPA), which allows determining whether the spatial variability of neural responses evoked by one condition is explained by the variability of a second condition. This approach provides a more stringent test to assess whether two distinct domains (e.g. emotions and beliefs) recruit the same cortical processes (Peelen and Downing, 2007; Corradi-Dell'Acqua *et al.*, 2011). In particular, we hypothesized that a pattern overlap between emotions and beliefs (but not between emotions and pain) in DMPFC and TPJ would provide unprecedented evidence that similar representational processes are implicated in understanding emotions and other mental states in people. Conversely, an overlap between emotions and pain (but not between emotions and beliefs) would reveal patterns which are predominantly involved in coding affective states in a non-representational fashion.

METHODS

Participants

Forty-six female subjects (age: 18–31 years) took part in the experiment. None had any history of neurological or psychiatric illness. Written informed consent was obtained from all subjects, who were naive to the purpose of the experiment. The study was approved by the ethics committee of Geneva University Medical School.

Stimuli and task

Participants read 12 short French-written stories ('scenarios') describing a person engaged in various situations (presented for 12 s), each of which was followed by a question ('judgment'—5 s) probing for participant's awareness of the protagonist's beliefs (beliefs [B]). We also included 12 different stories which were followed by a question about a person's emotional states (emotions [E]), and 12 others followed by a question about a person's somatic aching states (pain [Pa]). As high-level control, we used 12 stories with no protagonist but referring to an outdated physical representation on a map or photograph (photos [Ph]). Beliefs and photo stories were based on the *theory-of-mind* localizer used in previous studies (Saxe and Kanwisher, 2003; Saxe and Powell, 2006). Emotion and pain stories were based from other studies (Hynes *et al.*, 2006; Saxe and Powell, 2006) and modified to match (word-length, complexity, etc.) beliefs and photos stories, or were created *ad hoc* for this experiment. Supplementary Materials report all the stories used in this study.

The stories were presented using E-Prime 2.0 (Psychology Software Tools, Inc.) and projected inside the scanner bore with a LCD projector (CP-SX1350, Hitachi, Japan) on a screen subtending about 14.25° (vertical) × 19° degrees of visual angle. Key-presses were recorded on an MRI-compatible bimanual response button box (HH-2 × 4-C, Current Designs Inc., USA). Participants were instructed to press one of two possible keys, placed at each hand's reach, to indicate their responses. They had to press the key corresponding to the same side as the answer they believed to be correct (i.e. press the right hand button when they felt that the correct response was on the right side of the screen). Half of all correct responses were located on the right side, and the other half on the left side.

Please note that, although in the 'judgment' stage participants were asked to evaluate only one dimension (belief, emotion and pain), the

'scenarios' contained more heterogeneous information and, despite the intended differences between experimental conditions, might have elicited in the reader undesirable considerations about multiple dimensions at the same time. The idiosyncratic likelihood of scenarios to elicit considerations about each distinct dimension was therefore evaluated by a separate behavioral rating task in an independent group of 40 participants (age: 19–54 years). Results are displayed in Figure 1 and confirmed that a given scenario could elicit attributions of more than one category. Please see Supplementary Materials for full details on how these ratings were acquired.

Imaging processing

Data acquisition

A Siemens Trio 3-T whole-body scanner was used to acquire both T1-weighted anatomical images and gradient-echo planar T2-weighted MRI images with blood oxygenation level dependent contrast. The scanning sequence was a trajectory-based reconstruction sequence with repetition time (TR) of 2100 ms, echo time (TE) of 30 ms, flip angle of 90 degrees, 32 slices, slice thickness of 3 mm and no gap between slices.

Preprocessing

Statistical analysis was performed using the SPM8 software (<http://www.fil.ion.ucl.ac.uk/spm/>). For each subject, all functional images were realigned, normalized to a template based on 152 brains from the Montreal Neurological Institute (MNI), and then smoothed by convolution with an 8 mm full-width at half-maximum Gaussian kernel.

First-level analysis

Data from each participant were analyzed using the General Linear Model (GLM) framework implemented in SPM. For each experimental session, we modeled with a boxcar function a 12 s long block during which the scenario was presented separately from another 5 s long block during which the judgment took place. We modeled four 'judgment' vectors, one for each of the four kinds of stories (B, E, Pa and Ph), whereas we modeled only one 'scenario' vector, in which all 48 stories were treated as a unique condition, but completed by further parametric regressors (see below). This yielded, for each session, to five main regressors, for each of which we accounted for putative habituation effects in neural responses by using the linear time-modulation option implemented in SPM, creating an additional regressor in which the block order is modulated parametrically. Furthermore, to account for idiosyncratic differences across scenarios, we included four additional regressors in which specific 'scenarios' dimensions were modulated parametrically: the first described the number of characters in the scenarios, the remaining three described the likelihood to elicit mental attributions beliefs, emotions, and pain, respectively, as measured in the rating task (Figure 1). The scenarios of the 12 photos stories, which had no human protagonists and therefore were not included in the rating task, were associated with an artificial value of 0. Text-length was also controlled for 'judgment' events with a parametric regressor in which, however, no distinction was made between the four kinds of stories. This resulted in a total of 15 regressors (1 scenario vector + 4 judgment vectors + 5 time modulations + 3 rating modulations + 1 scenario text-length modulations + 1 judgment text-length modulations) that were convolved with a canonical hemodynamic response function and associated with regressors describing their first order temporal derivative. To account for movement-related variance, we also included six differential movement parameters as covariates of no interest. Thus, 36 regressors represented each of the two sessions

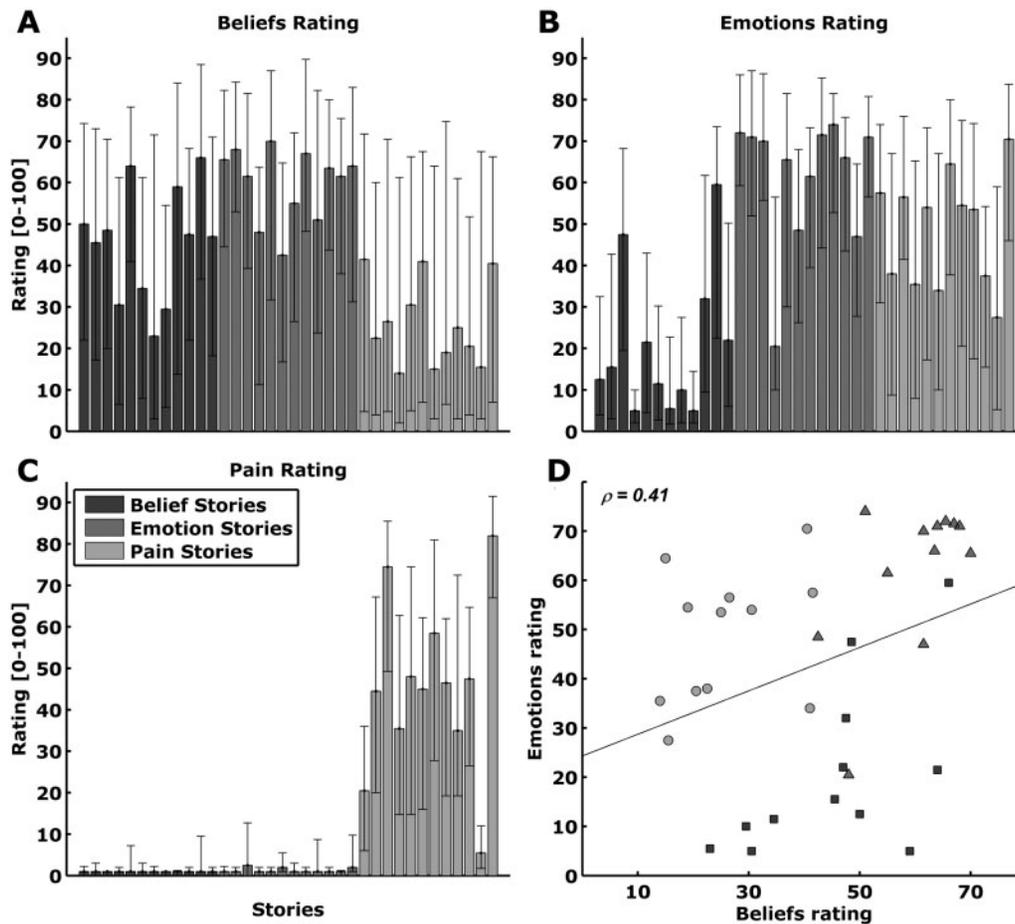


Fig. 1 Likelihood of each scenario to elicit attributions of (A) beliefs, (B) emotions and (C) pain to the story protagonist. Each bar reflects the median value of ratings made by 40 independent participants on a visual analogical scale ranging from 0 (the scenario never elicited considerations about beliefs/emotion/pain) to 100 (the scenario elicited a lot of considerations . . .). Inter-quartile differences are displayed as error bars. Dark gray bars refer to belief stories, light gray bars refer to pain stories, whereas intermediate gray bars refer to emotional stories. (D) Median ratings for attributions of beliefs plotted against median ratings for emotions, together with a regression line describing the relative dependency between these values throughout the 36 scenarios. The Spearman's rho correlation coefficient is also reported ($\rho = 0.41$, $P < 0.05$). As in the bar graphs, dark gray squares refer to belief stories, triangles refer to emotion stories, whereas light gray circles refer to pain stories.

modeled in the first level analysis. Low-frequency signal drifts were filtered using a cutoff period of 128 s.

Second-level analysis

Functional contrasts, describing effects of parameter estimate in both experimental sessions, were then fed in a second-level one-sample *t*-test. Clusters, isolated with a height threshold of $t_{(45)} = 3.28$ ($P < 0.001$ uncorrected), were considered significant if larger than the 95th percentile of the distribution of the largest clusters across the whole brain after 5000 replications of the analysis on permuted datasets. This second-level statistical analysis was carried out using the SnPM toolbox of SPM (<http://go.warwick.ac.uk/tenichols/snpm>).

Multivoxel pattern analysis

MVPA was carried out on preprocessed data which, at variance with the standard univariate approach, did not undergo normalization and smoothing. For each subject, we first analyzed our data with GLMs which were identical to those carried out for the univariate analysis. Following previous studies (Peelen *et al.*, 2010; Corradi-Dell'Acqua *et al.*, 2011), for each voxel of the individual native brain image, a cubic volume of interest surrounding the voxel was defined (5 voxels per side, 125 voxels total). For each of the two experimental sessions,

we extracted the parameter estimates associated with each of the four conditions used in the judgment epochs (B, E, Pa and Ph) in this volume of interest. Thus, each of the four regressors in each session was associated with a unique multivoxel pattern of β s in the volume-of-interest. These patterns were then correlated one with another, thus resulting in a symmetrical 8×8 matrix. The correlation coefficients r in this matrix reflected the similarity between the activation patterns evoked by different pairs of conditions. These coefficients were Fisher transformed $z = 0.5 \times \log_e[(1+r)/(1-r)]$ (Peelen *et al.*, 2010; Corradi-Dell'Acqua *et al.*, 2011) and then assigned to the center voxel of the cube. This resulted in z maps for each individual, which were then normalized to the MNI template.

We first identified activation patterns which were unique for each condition. This was achieved by testing whether the inter-session pattern similarity for identical conditions (e.g. z_{B-B} , correlation between belief patterns in session 1 and belief patterns in session 2) was, not only larger than 0, but also larger than the inter-session similarity between different conditions (e.g. z_{B-E} , described as the average of the correlation between belief patterns in session 1 and emotion patterns in session 2, and of the correlation between emotion patterns in session 1 and belief patterns in session 2). Then, we assessed whether two conditions of interest (e.g. belief and emotions) were associated with a correlation coefficient (z_{B-E}) larger than the coefficient obtained with other pairings (e.g. belief and pain). This yielded, at the group

level, differential statistical tests (i.e. $z_{B-B} > z_{B-P}$ or $z_{B-E} > z_{E-P}$) that were performed as in the standard univariate analysis (discussed earlier).

RESULTS

Participants performed accurately in all experimental conditions (mean 87.7% correct, see Supplementary Materials for behavioral results).

Univariate analysis of brain activations

Scenarios epochs

The scenarios text could elicit in participants spontaneous mental attributions concerning the protagonist's beliefs, emotions and pain to different degrees. We reasoned that the strength and content of such attributions across all scenarios would correspond to the scores obtained in the rating task. We therefore looked for brain regions significantly modulated by the parametrical regressor representing the scenarios' likelihood to elicit mental state attributions along each dimension, using the scores displayed in Figure 1 (see Methods). As the three likelihood vectors are not orthogonal (Figure 1D), and modeled together in the same GLM, their associated effects should be interpreted as reflecting exclusively the residual variance specific for each dimension (Andrade et al., 1999).

Table 1 and Figures 2A and 3A report the full results from this analysis of scenarios epochs. In line with previous observations (see Mar, 2011 as meta-analysis), belief-effects arose in the inferior parietal cortex bilaterally, extending ventrally to the location identified as TPJ in other studies (Figure 2A, red clusters). Further activations were also found bilaterally in the middle temporal gyrus (MTG), extending to STS in the right hemisphere only, in the hippocampus and inferior frontal gyrus on both sides. At the medial level, we observed activations in the medial prefrontal cortex (both dorsal and ventral aspects) and in the precuneus (PC) extending to the posterior cingulate cortex (PCC).

By contrast, emotion effects were found at the level of the supramarginal gyrus, bilaterally (Figure 2A, green clusters). At the medial level, we found activation of the middle cingulate cortex and of the ventral portion of medial prefrontal cortex (VMPFC), overlapping with the belief-related activation (Figure 2A, yellow cluster).

Finally, pain effects were found in the bilateral supramarginal gyrus and medial cingulate cortex (Figure 3A, blue clusters), overlapping with the emotion-related activation (cyan clusters). Further activity was found in the lateral orbital gyrus bilaterally, and in the insular cortex. No overlap was found between belief-related and pain-related activation.

Judgment epochs

Figure 2B (red clusters) and Table 2 display the regions activated during the judgment of beliefs as compared with photos ($B > Ph$). These increases were reminiscent of those identified in the scenario epochs, including bilateral TPJ and MTG/STS, PC, PCC and both the ventral and dorsal aspects of medial prefrontal cortex. Similar activations were found also when comparing beliefs against pain ($B > Pa$) (Table 3). Furthermore, the PCC was also activated, together with the left middle occipital gyrus and the anterior calcarine sulcus, when comparing the judgment of beliefs vs. emotions ($B > E$) (Figure 4A).

We then inspected more directly whether the specific contrasts $B > Pa$ and $B > E$ implicated portions of the medial prefrontal cortex or TPJ bilaterally by using a ROI analysis with predefined anatomical masks. The medial prefrontal mask was based on a standard anatomical atlas (AAL database, Tzourio-Mazoyer et al., 2002). For TPJ, as there are no clear anatomical boundaries, we followed Mars et al. (2012) who defined this area, on the dorsal-to-ventral axis, between

Table 1 Univariate fMRI Analysis: scenarios epochs

	Side	Coordinates			$t_{(45)}$	Cluster size
		x	y	z		
Attribution of beliefs (critical cluster size = 305 consecutive voxels)						
Medial prefrontal cortex (dorsal aspect)	M	8	58	12	13.19	8762 [†]
Medial prefrontal cortex (ventral aspect)		6	54	-6	6.91	
Precuneus/posterior cingulate	M	4	-52	28	16.97	3008 [†]
TPJ	R	60	-58	28	14.29	5587 [†]
MTG/STS		54	8	-26	13.81	
Inferior frontal gyrus		48	30	-8	6.29	
Hippocampus	R	24	-10	-16	6.13	383 [§]
TPJ	L	-52	-58	26	12.15	1801 [‡]
MTG	L	-64	-16	-14	11.74	3577 [†]
Inferior frontal gyrus		-48	30	-8	7.37	
Hippocampus		-22	-10	-14	6.06	
Attributions of emotions (critical size = 344 voxels)						
Medial Prefrontal cortex (ventral aspect)	M	-4	44	-10	4.21	382 [§]
Middle cingulate gyrus	M	-2	-18	36	6.15	1349 [§]
Cuneus	M	8	-82	28	4.58	376 [§]
Supramarginal gyrus	R	66	-22	30	7.11	682 [§]
Supramarginal gyrus	L	-58	-32	32	7.76	787 [§]
Attributions of pain (critical size = 220 voxels)						
Middle cingulate gyrus	M	-4	-28	26	6.51	5018 [†]
Cuneus		-10	-70	34	7.31	
Supramarginal gyrus	R	64	-32	36	7.03	708 [§]
Lateral orbital gyrus	R	46	42	6	5.60	372 [§]
Supramarginal gyrus	L	-58	-36	34	8.60	1196 [‡]
Lateral orbital gyrus	L	-42	42	10	6.50	1371 [‡]
Posterior orbital gyrus		-28	34	-14	8.15	
Middle posterior insula	L	-40	-4	-6	4.70	248 [§]

Regions whose activity changed linearly with the likelihood of attributions concerning the protagonist's (i) beliefs, (ii) emotions, and (iii) pain, as measured by a rating task in an independent group of participants. All regions survive permutation-based correction for multiple comparisons at the cluster level (with an underlying height threshold of $t_{(45)} = 3.28$, corresponding to $P < 0.001$ uncorrected). Coordinates (in standard MNI space) refer to maximally activated foci: x = distance (mm) to the right (+) or the left (-) of the midsagittal line; y = distance anterior (+) or posterior (-) to the vertical plane through the anterior commissure (AC); z = distance above (+) or below (-) the inter-commissural (AC-PC) line. L and R refer to the left and right hemisphere, respectively. M refers to medial/bilateral clusters.

[†] $P < 0.001$, [‡] $P < 0.01$, [§] $P < 0.05$ corrected for the whole brain at the cluster level

the intraparietal and superior temporal sulci, and on the posterior-to-anterior axis, between the MNI coordinates $y = -62$ and -32 (this mask is larger than what usually designated as TPJ, thus providing a conservative small volume correction). When focusing on the medial prefrontal ROIs, both DMPFC and VMPFC showed significant increases in the contrast $B > Pa$ (see Table 2 for details). No effect was found for the contrast $B > E$, neither in the prefrontal nor in the TPJ ROIs. Taken together, these findings accord with previous studies (Hynes et al., 2006; Völlm et al., 2006; Sebastian et al., 2012) reporting an involvement of these prefrontal and temporo-parietal areas in both belief and emotion judgments.

Brain activations arising during the emotion judgment epochs were also specifically examined, in comparison with the photo condition ($E > Ph$). As shown in Figure 2B (green clusters), these effects implicated a network highly similar to that previously identified for the judgment of beliefs (Figure 2B, yellow clusters). This network included the DMPFC and VMPFC, bilateral MTG, and the right TPJ (as further verified via ROI analysis). This network was also found when testing the contrast $E > Pa$, whereas testing $E > B$ only revealed activation in the DMPFC (Figure 4A). No effect in TPJ was found in the contrast $E > B$, not even under small volume correction. These contrasts confirm that our paradigm evoked not only shared effects between emotion and beliefs but also dissociations. These effects are depicted in Figure 4A, involving relatively greater recruitment of DMPFC for emotion judgments and of PC for belief judgments.

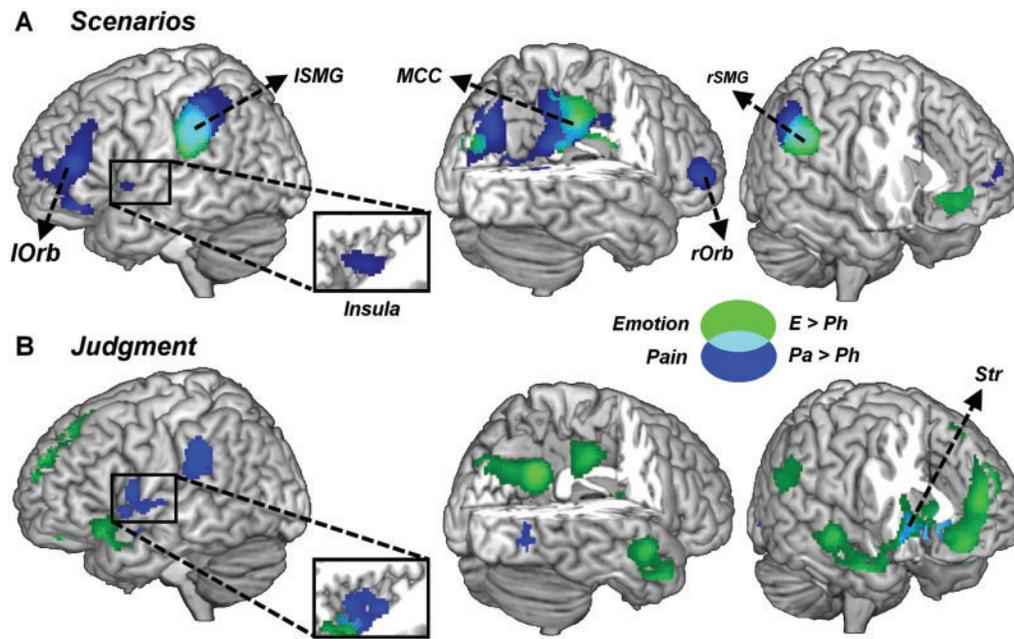


Fig. 3 Univariate analysis of fMRI data: common effects for emotions and pain. Whole-brain maps illustrate significant increases associated with (A) scenarios and (B) judgment epochs. The same parametric effects associated with attributions of emotions (scenarios) and differential increases in the contrast $E > Ph$ (judgment) shown in Figure 2 are also displayed here (always in green). Parametric effects associated with attributions of pain (scenarios) and increases in the contrast $Pa > Ph$ (judgment) are displayed in blue. Overlaps between green and blue clusters are displayed in cyan. SMG, supramarginal gyrus; MCC, middle cingulate cortex; Orb, lateral orbital cortex; Str, ventral striatum.

Finally, the judgment of pain compared with photos ($Pa > Ph$) led to activations in a distinct network (Figure 3B and Table 3), comprising, among others, the left supramarginal and insular structures already activated during the scenario epochs (discussed earlier). These effects were also confirmed in the contrasts $Pa > E$ and $Pa > B$ (Table 3).

Multivoxel pattern analysis

The existence of overlapping areas between conditions in univariate analyses does not necessarily imply that these conditions recruit the same cortical representations. We therefore used MVPA on the activation parameters associated with judgment epochs and probed for regions across the whole brain that would exhibit fine-grained patterns of activity differentially modulated across the three kinds of judgment.

We first distinguished beliefs and emotion patterns by searching for regions in which the inter-session correlation of voxel-by-voxel activity between similar conditions (e.g. $z_{B,B}$ or $z_{E,E}$) was larger than the inter-session correlation between different conditions (belief vs. emotion, $z_{B,E}$). Evidence for belief-patterns independent from emotions ($z_{B,B} > z_{B,E}$) was found in the ventral striatum and DMPFC (Table 4 and Figure 4B, red clusters). Instead, evidence for emotion-specific patterns independent from beliefs ($z_{E,E} > z_{B,E}$) was found in a more extensive portion of DMPFC, approximately 10 mm ventral to the previous analysis (Figure 4B, green clusters). None of these tests showed any specific effect in TPJ, not even with liberal statistical thresholds.

We use a similar approach to dissociate beliefs and emotion patterns from those evoked by pain and photos judgments (i.e. $z_{B,B} > z_{B,Pa}$, $z_{E,E} > z_{E,Ph}$). These analyses confirmed a differential role for the DMPFC in belief and emotion judgments (only $z_{B,B} > z_{B,Ph}$ showed no difference under the chosen threshold—see Supplementary Materials). No region was instead consistently associated with pain-specific patterns [$z_{Pa,Pa} > z_{B,Pa}$, $z_{Pa,Pa} > z_{E,Pa}$, etc.].

Next, we tested for local cortical patterns that were common for beliefs and emotions. We reasoned that these regions would exhibit similar correlations, not only when each condition is paired with its

homologous but also when the two conditions are paired together (i.e. $z_{B,B} \approx z_{E,E} \approx z_{B,E}$). Furthermore, these correlation coefficients (not only $z_{B,B}$ and $z_{E,E}$ but, critically, also $z_{B,E}$) should be greater than those obtained when pairing either emotions or beliefs with the other conditions, pain or photos. Figure 4C shows the regions identified in the contrast $z_{B,E} > \text{mean}(z_{B,Pa}, z_{E,Pa}, z_{B,Ph}, z_{E,Ph})$, predominantly involving bilateral areas in TPJ and MTG. A more conservative analysis, testing the conjoint effect of four separate tests [$(z_{B,E} > z_{B,Pa}) \cap (z_{B,E} > z_{E,Pa}) \cap (z_{B,E} > z_{B,Ph}) \cap (z_{B,E} > z_{E,Ph})$], each at a height threshold of $P < 0.001$ uncorrected, confirmed a robust common effect in the left TPJ, in particular the posterior portion of the superior temporal gyrus. No involvement of the medial prefrontal cortex was found.

Figure 4C also plots, for visualization purposes, the average parameter estimates from both the ‘shared’ (TPJ and MTG) and ‘independent’ regions (DMPFC), showing how all these exhibit high correlation for both belief ($z_{B,B}$) and emotion ($z_{E,E}$). The critical difference between these regions is demonstrated by the coefficient $z_{B,E}$ which, in TPJ and MTG, was as high as $z_{B,B}$ and $z_{E,E}$, whereas in DMPFC it was lower than the condition-specific values and comparable with other control pairings ($z_{B,Pa}$, $z_{E,Pa}$, etc.). Using a similar approach, we also searched for common patterns between beliefs and pain, and between emotion and pain, but these analyses revealed no effects.

DISCUSSION

Our results reveal a striking dissociation in the information represented within the human brain networks mediating theory of mind. Using fMRI with MVPA, we provide the first evidence for shared patterns of cortical activity in temporo-parietal areas (TPJ/MTG) between the judgment of other people’s emotions and beliefs (but not pain). In contrast, although beliefs and emotions both produce reliable and coherent patterns in DMPFC, these were independent from one another, thus suggesting the presence of distinct neuronal populations involved in these different mentalizing abilities.

Table 2 Univariate fMRI analysis: judgment epochs

	Side	Coordinates			$t_{(45)}$	Cluster size
		x	y	z		
Belief > Photos (critical size = 244 voxels)						
Medial prefrontal cortex (dorsal aspect)	M	8	56	12	6.62	1806 [‡]
Medial prefrontal cortex (ventral aspect)		-4	44	-14	6.13	
Precuneus/posterior cingulate	M	6	-54	30	9.29	6072 [†]
TPJ	R	58	-46	22	6.43	1391 [*]
MTG/STS	R	66	-24	-4	5.90	672 [§]
Hippocampus	R	26	-22	-16	5.13	402 [§]
TPJ	L	-48	-58	24	4.50	541 [§]
MTG	L	-64	-20	-12	6.52	325 [§]
Hippocampus	L	-20	-28	-12	5.28	258 [§]
Emotions > photos (critical size = 404 voxels)						
Medial prefrontal cortex (dorsal aspect)	M	8	56	10	9.43	2957 [†]
Medial prefrontal cortex (ventral aspect)		-4	44	-14	6.66	
Middle cingulate gyrus	M	-0	-14	38	6.14	728 [§]
Precuneus/posterior cingulate	M	8	-52	28	9.02	1795 [‡]
Ventral striatum	R	16	10	-8	5.37	3113 [‡]
Temporal pole	R	62	-8	-12	6.10	
Ventral striatum	L	-12	12	-8	6.21	
Temporal pole	L	-30	10	-22	5.95	
TPJ	R	66	-40	22	4.27	220 [¶]
Pain > photos (critical size = 284 voxels)						
Lingual gyrus	R	18	-72	2	4.53	486 [§]
Parieto-occipital fissure	R	16	-72	46	4.34	631 [§]
Ventral striatum	R	12	4	-8	5.46	812 [§]
Ventral striatum	L	-14	8	-12	4.35	
Supramarginal gyrus	L	-56	-34	36	5.39	389 [§]
Middle-posterior insula	L	-38	-8	-6	4.61	674 [§]

Regions showing significant activation for explicit judgment of the protagonist's (i) beliefs, (ii) emotions and (iii) pain, relative to the control 'photos' condition.

[†] $P < 0.001$, [‡] $P < 0.01$, [§] $P < 0.05$ corrected for the whole brain at the cluster level.

[¶] $P < 0.05$ corrected for bilateral TPJ.

Shared neural systems for beliefs and emotions

Previous studies consistently implicated posterior temporal regions in mentalizing abilities. Indeed, not only these regions are active during ToM tasks (Mar, 2011), but understanding the mental states of others is selectively affected after TPJ impairment (Samson *et al.*, 2004; Young *et al.*, 2010). Some studies tested whether these regions are also recruited when assessing people's emotions, with mixed results: whereas some reported a similar magnitude of activation in TPJ for the judgment of characters' beliefs and emotions (Hynes *et al.*, 2006; Völlm *et al.*, 2006; Sebastian *et al.*, 2012), others found TPJ and MTG preferentially active for beliefs (Zaitchik *et al.*, 2010; Bruneau *et al.*, 2012). By using a powerful dataset ($N = 46$) and a fine-grained MVPA approach, we were able to confirm the existence of activity patterns equally associated with the judgment of beliefs and emotions. Critically, the internal similarity of beliefs and emotion patterns (as determined by the inter-session correlation of activity between homologous conditions) was comparable with the similarity between emotions and beliefs (inter-session correlation of between different conditions). By contrast, emotion and belief patterns showed no similarity with other conditions. These results suggest that emotion and belief patterns in TPJ and MTG reflect the same neuronal activity, equally recruited in these two independent conditions.

Recent models (Shamay-Tsoory *et al.*, 2010; Sebastian *et al.*, 2012) proposed that understanding others' emotions is mediated (at least in adults) by those representational abilities that are also necessary for understanding non-emotional mental states. Although a literal interpretation of this view cannot explain how impairments in ToM abilities can be associated with spared affective judgments (Kalbe *et al.*, 2010; Sebastian *et al.*, 2012), our data favor an alternative

Table 3 Univariate fMRI Analysis: judgment epochs

	Side	Coordinates			$t_{(45)}$	Cluster size
		x	y	z		
Belief > emotions (critical size = 380 voxels)						
Precuneus	M	-6	-60	44	5.26	507 [§]
Angular gyrus/middle occipital gyrus	L	-38	-74	34	6.48	466 [§]
Anterior calcarine sulcus	L	-14	-56	16	6.78	489 [§]
Emotions > belief (critical size = 422 voxels)						
Medial prefrontal cortex (dorsal aspect)	M	-6	-60	18	4.69	511 [§]
Ventral striatum	R	12	2	0	5.40	1591 [§]
Ventral striatum	L	-14	8	4	5.22	
Belief > pain (critical size = 197 voxels)						
Precuneus/posterior cingulate	M	-2	-56	36	7.93	2333 [‡]
TPJ	R	62	-54	26	8.29	1354 [‡]
MTG/STS	R	52	2	-28	7.74	790 [‡]
Superior frontal sulcus	R	24	32	50	5.33	238 [§]
TPJ	L	-52	-56	30	7.74	1447 [‡]
MTG	L	-64	-14	-12	7.36	825 [‡]
Superior frontal sulcus	L	-28	30	48	5.04	319 [§]
Parahippocampal gyrus	L	-26	-34	-16	5.06	226 [§]
Medial prefrontal cortex (dorsal aspect)	M	6	60	16	3.84	86 [*]
Medial prefrontal cortex (ventral aspect)	M	2	54	-16	3.65	18 [*]
Pain > belief (critical size = 204 voxels)						
Supramarginal gyrus	L	-53	-36	36	4.82	210 [§]
Inferior Frontal Gyrus	L	-50	8	18	6.19	394 [§]
Emotion > pain (critical size = 237 voxels)						
Medial prefrontal cortex (dorsal aspect)	M	-6	-62	18	7.93	2998 [‡]
Medial prefrontal cortex (ventral aspect)		-2	48	-18	6.66	
Precuneus/posterior cingulate	M	-4	-52	30	8.90	1632 [‡]
TPJ	R	60	-56	28	7.87	699 [§]
MTG	R	58	-2	-26	7.93	1213 [‡]
Precentral gyrus	R	40	-28	64	4.57	305 [§]
TPJ	L	-54	-60	32	8.23	1139 [‡]
MTG	L	-60	-10	-14	9.83	1760 [‡]
Inferior frontal gyrus	L	-48	28	8	5.48	
Pain > emotion (critical size = 250 voxels)						
Supramarginal gyrus	L	-58	-34	36	5.30	451 [§]
Inferior frontal gyrus	L	-46	6	16	6.97	913 [§]
Middle posterior insula	L	-40	-12	4	4.64	

Regions showing significant activation when comparing explicit judgment of beliefs, emotions, and pain, one against the other.

[‡] $P < 0.01$, [§] $P < 0.05$ corrected for the whole brain at the cluster level.

* $P < 0.05$ corrected for the medial prefrontal cortex.

interpretation according to which, due to flexible engagement of domain-specific functions, mentalistic representations are one of many possible components or strategies underlying our comprehension of others' emotions. In this perspective, ToM-patterns in TPJ/MTG may constitute the neural signature of such mentalizing strategy, recruited for the judgment of both affective and non-affective mental states.

Finally, our data extend previous accounts by offering new insights on what information might be shared between the two kinds of inferences. Indeed, belief-related activity in TPJ/MTG was found both during spontaneous evaluations triggered by the text (scenario epochs) and during the cued judgments of mental states (judgment epochs); instead, emotion-related activity in these regions was found exclusively during the judgment epochs (Figure 2). Furthermore, emotion judgments in our task (and presumably in earlier studies with similar materials, Hynes *et al.*, 2006) followed scenarios associated with spontaneous belief attributions (Figure 1A). It is therefore plausible that cued judgments about others' beliefs and emotions share some key process related to the appraisal of those mental states (character's thoughts, goals, etc.) that are crucial for both belief and emotion inferences. Indeed, appraisal theories of emotions (Scherer, 1984, 2009) propose that affective experience is critically determined

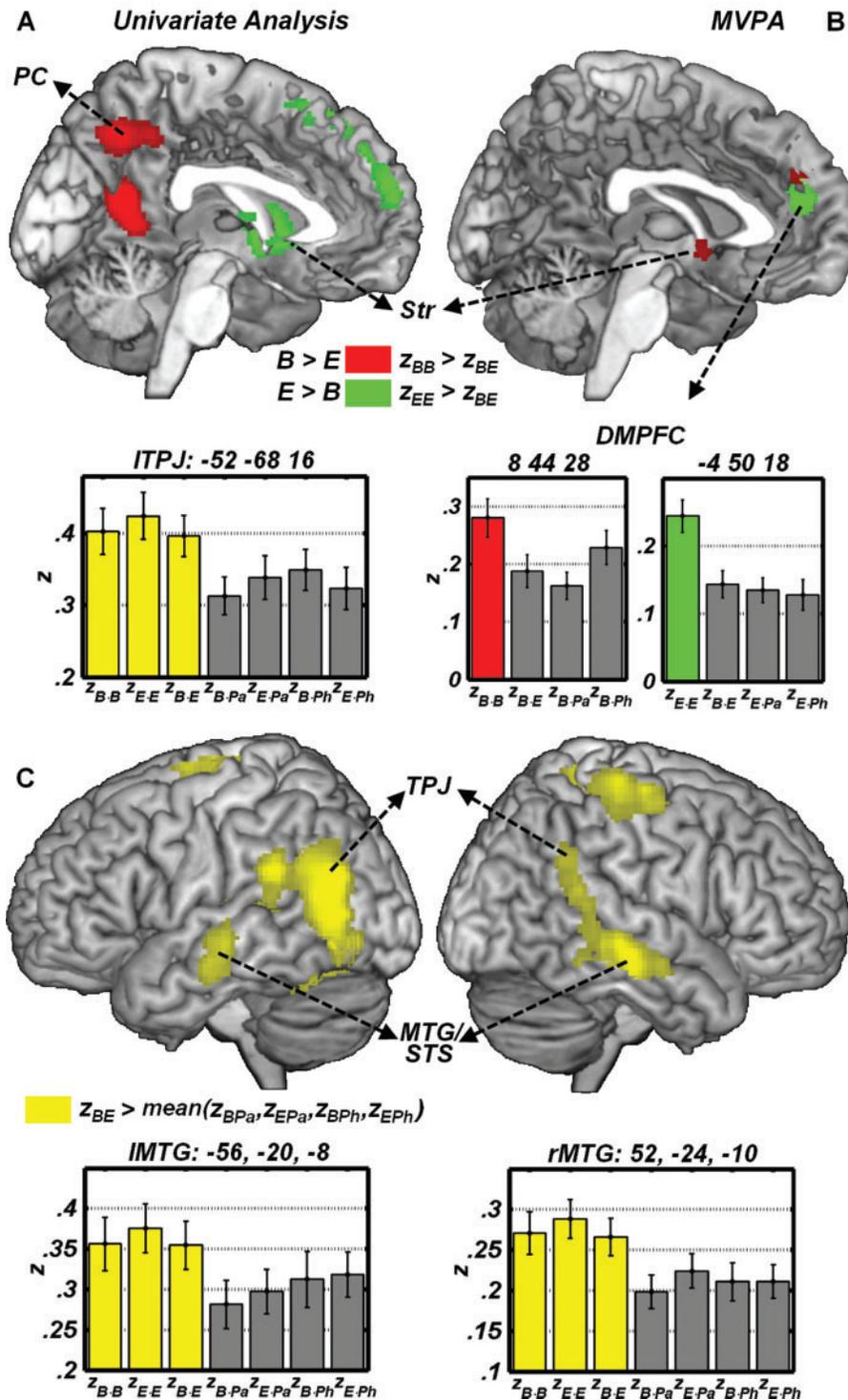


Fig. 4 Whole-brain maps showing regions with dissociated activity judgments of beliefs and emotions in (A) univariate analysis and (B) MVPA. Red clusters refer to specific effects of beliefs; green clusters refer to specific effects of emotions. (C) MVPA results showing regions with shared voxel-by-voxel patterns between judgments of emotions and beliefs (yellow clusters). The Fisher-transformed correlation coefficients (averaged across all voxels in each region) are plotted with standard error bars for visualization purposes. Color codes for the graph bars correspond to the statistical test used to identify each region.

by the contextual evaluation of events, current actions and goals, others' beliefs and norms, as well as by the interaction thereof (e.g. pride or shame are based on a representation of how the social environment considers one's actions). It stands to reason that an accurate evaluation of emotions in other people should also be built on a contextual appraisal of their mental states reflected at the brain level by the engagement of TPJ and MTG.

Specific neural systems for beliefs and emotions

Like the TPJ and MTG, the medial prefrontal cortex also exhibited increased activity for both beliefs and emotions. However, unlike TPJ and MTG, the cortical signals evoked in medial prefrontal areas were independent (although overlapping) for each type of judgment. Our results converge with, but also extend, previous investigations reporting that, in both healthy and brain-damaged individuals, medial

Table 4 MVPA with small-volume searchlight approach

	Side	Coordinates			$t_{(45)}$	Cluster size
		x	y	z		
Beliefs-patterns independent from emotions: $z_{B,B} > z_{B,E}$ (critical size = 302 voxels)						
Ventral striatum	R	20	2	-8	6.67	614 [§]
Medial prefrontal cortex (dorsal aspect)	M	8	44	28	3.96	86 [†]
Emotions-patterns independent from beliefs: $z_{E,E} > z_{B,E}$ (critical size = 289 voxels)						
Medial prefrontal cortex (dorsal aspect)	M	-4	50	18	4.31	390 [§]
Patterns common for belief and emotion: $z_{B,E} > \text{mean}(z_{B-Pa}, z_{E-Pa}, z_{B-Pa}, z_{E-Pa})$ (critical size = 295 voxels)						
Superior frontal gyrus	M	-14	-16	72	4.85	967 [‡]
Precentral gyrus	R	38	-28	66	4.94	790 [‡]
TPJ	R	58	-46	24	3.83	1233 [‡]
MTG/STS	R	52	-24	-10	5.84	
TPJ	L	-52	-68	16	6.45	4525 [†]
MTG	L	-56	-20	-8	4.49	
Conjunction analysis: $(z_{B,B} > z_{B-Pa}) \cap (z_{E,E} > z_{E-Pa}) \cap (z_{B,B} > z_{B-Pa}) \cap (z_{B,E} > z_{E-Pa})$						
TPJ	L	-46	-64	14	3.46	30

[†] $P < 0.001$, [‡] $P < 0.01$, [§] $P < 0.05$ corrected for the whole brain at the cluster level.

[†] $P < 0.05$ corrected for the medial prefrontal cortex.

prefrontal regions appear critically involved in the evaluation of others' beliefs (Bird *et al.*, 2004; Mar, 2011) or emotions (Hynes *et al.*, 2006; Shamay-Tsoory *et al.*, 2006, 2010; Peelen *et al.*, 2010; Leopold *et al.*, 2012).

When testing for activity during scenario epochs, belief-related effects were observed in both the DMPFC and VMPFC, whereas emotion-related effects were associated with VMPFC only. In our analysis we fed the emotion and belief values from the rating task (Figure 1) in the same GLM. This ensured that, in regions where both effects coexist (e.g. VMPFC), the emotion parameters were estimated on a portion of variance that was not explained by belief ratings. Furthermore, during the scenarios epochs, participants were not aware of which specific dimension (beliefs, emotion, pain) would be the object of the subsequent judgment (although they knew that there would be a judgment), so that the effects found in VMPFC and DMPFC (Figure 2A) can be considered as exclusively text-driven. In this perspective, the functional role of VMPFC in emotional coding diverges quite sharply from the TPJ and MTG: whereas the latter regions were not sensitive to emotion-relevant features in the scenarios text, but were engaged by emotion attribution in the subsequent judgment epoch, the prefrontal regions instead encoded emotional information already early in the scenario epoch.

The analysis of brain activity during the judgment epochs also implicated VMPFC and DMPFC in both emotion and belief judgments (like for TPJ/MTG). However, unlike for TPJ and MTG, our fine-grained MVPA approach suggests that in prefrontal areas no pattern of activity was shared between these two dimensions. On the contrary, we found strong evidence that DMPFC exhibited coherent and specific patterns for emotions and beliefs, which were clearly dissociable from each other despite their spatial proximity. This result would not have been obtained by relying only on a standard univariate approach. These functional properties of DMPFC suggest the existence of independent neuronal populations in this region, each responding selectively to one kind of mental state attribution. An alternative (less likely) interpretation might instead surmise that the same neurons could change their firing rate according to the different kinds of inference.

In either case, it is plausible that, whereas the TPJ and MTG hold broad representational processes common to both beliefs and emotions, DMPFC maintains more specific information unique to each condition. Interestingly, a study employing a repetition suppression

paradigm (Jenkins *et al.*, 2008) reported that activity associated with mentalizing and self-reflection relied on the same medial prefrontal structures. Therefore, at variance with TPJ/MTG, the DMPFC might be specifically involved in *simulation-based* mentalizing, i.e. when using one's own mind as a model for others (Jenkins *et al.*, 2008). This interpretation fits nicely with both current and previous data, as *simulation-based* accounts do not predict that patterns associated with different cognitive and affective states should be common; instead, each state inferred in other people should trigger a specific activity pattern similar to the one produced when this same state is experienced by oneself. Further studies are needed to verify this prediction.

Specific and shared effects of pain

We found that text-based evaluation of emotion and pain in others similarly enhanced activity in supramarginal and cingulate regions, while additional increases arose in the insula for pain only. In addition, for pain, the same regions were also active during the judgment epochs. The supramarginal, cingulate, and insular activations reported in this condition are remarkably similar to those previously implicated in perceiving physical pain in others via photographs (Corradi-Dell'Acqua *et al.*, 2011; Lamm *et al.*, 2011), or purely verbal description (Gu and Han, 2007; Bruneau *et al.*, 2012). The same regions are also often found to exhibit 'shared' sensitivity to pain, as they respond to both direct and vicarious nociception, in studies using both univariate and MVPA approaches (Corradi-Dell'Acqua *et al.*, 2011; Lamm *et al.*, 2011).

Insula and cingulate cortex have been suggested to represent others' affective states through the same bodily and interoceptive sensations that are associated with direct emotional experience (Singer *et al.*, 2009; Lamm and Singer, 2010). In this perspective, the sensitivity of these regions to both emotion and pain considerations might reflect the ability to code for affective sensations of other people in a non-mentalistic fashion, i.e. by imagining bodily manifestations (swellings, shivers, palpitations, etc.) of affective relevance from the story texts. Future studies need to further investigate this conjecture which, if true, would imply the presence of two strategies/pathways for inferring people's affect: one based on the appraisal of their mental states (as manifested in explicit judgments of emotion) and another based on the representation of somatic and interoceptive sensations (as in the case of pain here).

SUPPLEMENTARY DATA

Supplementary data are available at SCAN online.

REFERENCES

- Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7, 268–77.
- Andrade, A., Paradis, A.L., Rouquette, S., Poline, J.B. (1999). Ambiguous results in functional neuroimaging data analysis due to covariate correlation. *Neuroimage*, 10, 483–6.
- Baillargeon, R., Scott, R.M., He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, 14, 110–8.
- Bird, C.M., Castelli, F., Malik, O., Frith, U., Husain, M. (2004). The impact of extensive medial frontal lobe damage on "theory of mind" and cognition. *Brain*, 127, 914–28.
- Bruneau, E.G., Pluta, A., Saxe, R. (2012). Distinct roles of the "shared pain" and "theory of mind" networks in processing others' emotional suffering. *Neuropsychologia*, 50, 219–31.
- Corbetta, M., Patel, G., Shulman, G.L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, 58, 306–24.
- Corradi-Dell'Acqua, C., Civai, C., Rumiati, R.I., Fink, G.R. (2013). Disentangling self- and fairness-related neural mechanisms involved in the ultimatum game: an fMRI study. *Social Cognitive and Affective Neuroscience*, 8, 424–31.
- Corradi-Dell'Acqua, C., Hofstetter, C., Vuilleumier, P. (2011). Felt and seen pain evoke the same local patterns of cortical activity in insular and cingulate cortex. *Journal of Neuroscience*, 31, 17996–8006.

- Corradi-Dell'Acqua, C., Ueno, K., Ogawa, A., Cheng, K., Rumiati, R.I., Iriki, A. (2008). Effects of shifting perspective of the self: an fMRI study. *Neuroimage*, *40*, 1902–11.
- Decety, J., Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *Neuroscientist*, *13*, 580–93.
- Flavell, J.H. (1999). Cognitive development: children's knowledge about the mind. *Annual Review of Psychology*, *50*, 21–45.
- Gallagher, H.L., Frith, C.D. (2003). Functional imaging of "theory of mind". *Trends in Cognitive Sciences*, *7*, 77–83.
- Gu, X., Han, S. (2007). Neural substrates underlying evaluation of pain in actions depicted in words. *Behavioural Brain Research*, *181*, 218–23.
- Hynes, C.A., Baird, A.A., Grafton, S.T. (2006). Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. *Neuropsychologia*, *44*, 374–83.
- Jenkins, A.C., Macrae, C.N., Mitchell, J.P. (2008). Repetition suppression of ventromedial prefrontal activity during judgments of self and others. *Proceedings of National Academy of Science U S A*, *105*, 4507–12.
- Kalbe, E., Schlegel, M., Sack, A.T., et al. (2010). Dissociating cognitive from affective theory of mind: a TMS study. *Cortex*, *46*, 769–80.
- Kovács, Á.M., Téglás, E., Endress, A.D. (2010). The social sense: susceptibility to others' beliefs in human infants and adults. *Science*, *330*, 1830–4.
- Lamm, C., Decety, J., Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage*, *54*, 2492–502.
- Lamm, C., Singer, T. (2010). The role of anterior insular cortex in social emotions. *Brain Structure and Function*, *214*, 579–91.
- Leopold, A., Krueger, F., dal Monte, O., Pardini, M., Pulaski, S.J., Solomon, J., Grafman, J. (2012). Damage to the left ventromedial prefrontal cortex impacts affective theory of mind. *Social Cognitive and Affective Neuroscience*, *7*, 871–80.
- Mar, R.A. (2011). The neural bases of social cognition and story comprehension. *Annual Review in Psychology*, *62*, 103–34.
- Mars, R.B., Sallet, J., Schüffelen, U., Jbabdi, S., Toni, I., Rushworth, M.F.S. (2012). Connectivity-based subdivisions of the human right "temporoparietal junction area": evidence for different areas participating in different cortical networks. *Cerebral Cortex*, *22*, 1894–903.
- Peelen, M.V., Atkinson, A.P., Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *Journal of Neuroscience*, *30*, 10127–34.
- Peelen, M.V., Downing, P.E. (2007). Using multi-voxel pattern analysis of fMRI data to interpret overlapping functional activations. *Trends in Cognitive Science*, *11*, 4–5.
- Samson, D., Apperly, I.A., Chiavarino, C., Humphreys, G.W. (2004). Left temporoparietal junction is necessary for representing someone else's belief. *Nature Neuroscience*, *7*, 499–500.
- Saxe, R., Carey, S., Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annual Review in Psychology*, *55*, 87–124.
- Saxe, R., Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage*, *19*, 1835–42.
- Saxe, R., Powell, L.J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychological Science*, *17*, 692–9.
- Scherer, K.R. (1984). On the nature and function of emotion: a component process approach. In: Scherer, K.R., Ekman, P., editors. *Approaches to Emotion*. Hillsdale, NJ: Erlbaum, pp. 293–317.
- Scherer, K.R. (2009). The dynamic architecture of emotion: evidence for the component process model. *Cognition & Emotion*, *23*, 1307–51.
- Sebastian, C.L., Fontaine, N.M.G., Bird, G., et al. (2012). Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. *Social and Cognitive Affective Neuroscience*, *7*, 53–63.
- Shamay-Tsoory, S.G., Harari, H., Aharon-Peretz, J., Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex*, *46*, 668–77.
- Shamay-Tsoory, S.G., Tibi-Elhanany, Y., Aharon-Peretz, J. (2006). The ventromedial prefrontal cortex is involved in understanding affective but not cognitive theory of mind stories. *Social Neuroscience*, *1*, 149–66.
- Singer, T., Critchley, H.D., Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Sciences*, *13*, 334–40.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*, 273–89.
- Völlm, B.A., Taylor, A.N.W., Richardson, P., et al. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *Neuroimage*, *29*, 90–8.
- Young, L., Camprodon, J.A., Hauser, M., Pascual-Leone, A., Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of National Academy of Science USA*, *107*, 6753–8.
- Zaitchik, D., Walker, C., Miller, S., LaViolette, P., Feczko, E., Dickerson, B.C. (2010). Mental state attribution and the temporoparietal junction: an fMRI study comparing belief, emotion, and perception. *Neuropsychologia*, *48*, 2528–36.
- Zaki, J., Weber, J., Bolger, N., Ochsner, K. (2009). The neural bases of empathic accuracy. *Proceedings of National Academy of Science USA*, *106*, 11382–7.