

Mirror Neuron and Theory of Mind Mechanisms Involved in Face-to-Face Interactions: A Functional Magnetic Resonance Imaging Approach to Empathy

Martin Schulte-Rüther^{1,2}, Hans J. Markowitsch³, Gereon R. Fink^{1,4},
and Martina Piefke^{1,2}

Abstract

■ Empathy allows emotional psychological inference about other person's mental states and feelings in social contexts. We aimed at specifying the common and differential neural mechanisms of "self"- and "other"-related attribution of emotional states using event-related functional magnetic resonance imaging. Subjects viewed faces expressing emotions with direct or averted gaze and either focused on their own emotional response to each face (self-task) or evaluated the emotional state expressed by the face (other-task). The common network activated by both tasks included the left lateral orbito-frontal and medial prefrontal cortices (MPFC), bilateral inferior frontal cortices, superior temporal sulci and temporal poles, as well as the right cerebellum. In a subset of these regions, neural activity was significantly correlated with empathic abilities. The self- (relative to the other-) task differentially activated the

MPFC, the posterior cingulate cortex (PCC)/precuneus, and the temporo-parietal junction bilaterally. Empathy-related processing of emotional facial expressions recruited brain areas involved in mirror neuron and theory-of-mind (ToM) mechanisms. The differential engagement of the MPFC, the PCC/precuneus, and temporo-parietal regions in the self-task indicates that these structures act as key players in the evaluation of one's own emotional state during empathic face-to-face interaction. Activation of mirror neurons in a task relying on empathic abilities without explicit task-related motor components supports the view that mirror neurons are not only involved in motor cognition but also in emotional interpersonal cognition. An interplay between ToM and mirror neuron mechanisms may hold for the maintenance of a self-other distinction during empathic interpersonal face-to-face interactions. ■

INTRODUCTION

A key aspect of social interaction is the inference of other persons' emotional states by evaluating their facial expressions (Ekman, 1982). These can help us to gain access to someone's feelings and may act as triggers of empathy. Empathy is based upon processes of psychological inferences about other persons' mental and emotional states occurring within a specific social context. This context may provide the frame of reference for the integration of emotion and cognition to yield insights into the intentions and feelings of others (Decety & Jackson, 2004). Three core aspects are likely to play a role in the emergence and maintenance of empathy: (i) an intuitive feeling of having something in common with the other person which relies on socially shared emotional experiences; (ii) cognitive mechanisms of perspective-taking; and (iii) the ability to maintain a self-

other distinction during interpersonal interaction. Empathic feelings allow for socially appropriate emotional responses, be it a shared or reactive emotional state. Therefore, empathy is different from pure emotional contagion as the resulting emotional state *can* be different from the one being observed (see, for example, Davis, 1996; Stotland, 1969).

It has been suggested that shared emotional feelings might be traced back to mirror neuron mechanisms (Gallese, 2003), which had initially been proposed in the field of motor representations. Mirror neurons have been discovered in the monkey brain in frontal area F5 and parietal area PF (see, e.g., Rizzolatti, Fogassi, & Gallese, 2001, for a review). There is increasing evidence from neuroimaging studies that core components of a *human* mirror neuron system (hMNS) comprise parts of the inferior frontal cortex and the posterior parietal cortex (e.g., Koski, Iacoboni, Dubeau, Woods, & Mazziotta, 2003; Iacoboni et al., 1999). The mirror neuron account has recently been adapted to processes operating during interpersonal cognition (Gallese, Keysers, & Rizzolatti, 2004; however, see Jacob & Jeannerod, 2005 for a critique). It has been suggested that the

¹Research Center Jülich, Germany, ²University Hospital Aachen, Rheinisch-Westfälische Technische Hochschule (RWTH) Aachen, Germany, ³University of Bielefeld, Germany, ⁴University Hospital Cologne, Germany

brain networks controlling motor behavior and social interaction share (at least in part) computational parallels (Wolpert, Doya, & Kawato, 2003), and might be implemented in similar or overlapping brain regions. Recent studies indeed indicate that not only motor plans but also more abstract intentions (Iacoboni et al., 2005) or emotional states (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003) recruit the same areas in the inferior frontal cortex which are supposed to contain mirror neurons.

Carr et al. (2003) demonstrated that components of the hMNS (including the inferior frontal cortex) are involved in the imitation *and* observation of facial expressions of emotions. Reports on hMNS involvement in the mere perception of facial expressions of emotions, however, are yet rare. This might be due to the fact that most previous studies focused on the implicit processing of emotional facial expressions using distractor tasks such as gender detection or conditioning paradigms (Winston, O'Doherty, & Dolan, 2003; Critchley et al., 2000), and/or explicit tasks such as emotion categorization or judgment of emotion intensity (e.g., Winston et al., 2003; Critchley et al., 2000; George et al., 1993). All these tasks can be accomplished by the use of perceptual strategies, for example, pattern matching. The above studies may thus have failed to detect hMNS activity because the perceptual processing of emotional facial expressions was not embedded in an interpersonal context.

For pain and disgust, functional neuroimaging studies showed evidence that the processing of one's own emotional experience and the observation of another person having a similar emotional experience may engage overlapping brain networks (Singer et al., 2004; Wicker et al., 2003). Interestingly, in the study by Singer et al. (2004), the relative engagement of these networks was correlated with individual differences in empathic abilities.

The cognitive aspects of empathic interpersonal perspective taking are closely related to theory of mind (ToM). The ToM concept depicts meta-cognitive abilities such as the inference of beliefs, intentions, and desires of other people (Premack & Woodruff, 1978). Importantly, ToM also requires the *distinction* between one's own thoughts and intentions and those of others, which is a key component of interpersonal interactions.

The neural network supporting cognitive aspects of ToM include temporo-parietal areas, the medial prefrontal cortex (MPFC), and the temporal poles (e.g., Vogeley et al., 2001; see Frith & Frith, 2003; Vogeley & Fink, 2003 for reviews). Functional neuroimaging data on *emotional* ToM, however, remain scarce (but see Völlm et al., 2006). Few studies directly addressed the functional neuroanatomy of taking the self- or other-perspective in the context of emotional experience (Ochsner et al., 2004; Ruby & Decety, 2004). For example, Ochsner et al. (2004) asked their subjects to evaluate emotional responses to affective photographs from their own (i.e., first person) perspective or from the perspective of a character depicted in the photos (i.e., third-person perspective). This approach,

however, did not take into account that the empathic switching between the self- and the other-perspective goes beyond the "classic" first-person/third-person concept adopted by studies on perspective-taking (e.g., Vogeley & Fink, 2003; Vogeley et al., 2001). Rather, empathy relies on a dyadic social situation where the inferred intentions and *feelings* of the other directly affect the self as the target of the other's emotion. This account is also supported by studies of cognitive aspects of interpersonal interactions, suggesting that "on-line" social reasoning is most efficiently triggered in a context where the observed behavior of others is highly relevant to one's own intentions and actions (Kampe, Frith, & Frith, 2003; Gallagher, Jack, Roepstorff, & Frith, 2002). To our knowledge, no functional neuroimaging data are currently available on the neural circuits supporting empathic self-other interactions unfolding in the dynamics of emotional interpersonal contexts.

We used functional magnetic resonance imaging (fMRI) to investigate empathy-related neural processes, applying a paradigm which required both the decoding of other persons' emotional states from facial cues ("other"-condition) and the evaluation of one's own emotional response to these faces ("self"-condition). Subjects evaluated either the emotional state of the observed person or their own emotional reaction when viewing that person. Responses were given by choosing appropriate emotional descriptions from a list of adjectives depicting emotional states. The experimental setup thus differed from previous approaches in that it did not focus on a perceptual decision on the emotion expressed by a face (e.g., a decision between basic emotion categories), but rather assessed the interactive switching between the self- and the other-perspective during the attribution of emotion. Using this paradigm, we were able to experimentally construct an interpersonal context in which empathic social cognition could emerge. In addition, we manipulated the head direction of the stimulus faces such that they were either directed to the observer or averted by $\sim 45^\circ$. An age/gender decision task on neutral faces was used as a high-level baseline. Standard empathy scales were administered to assess individual differences in empathic abilities. Individual scores on these scales were used to determine whether measures of empathic capabilities may predict the amplitude of empathy-related brain activity, especially in areas belonging to the hMNS.

We expected that brain regions supporting ToM (e.g., MPFC, superior temporal regions) and structures previously associated with the hMNS (e.g., inferior frontal gyrus, superior temporal sulcus [STS], inferior parietal cortex) would be activated during both the attribution of emotion to oneself and the other person. However, we also hypothesized that some of these areas would show differential responses depending upon the involvement of either the self- or the other-perspective. hMNS engagement was expected to predominate in the

self-condition, whereas we considered ToM mechanisms to be specifically implicated in the other-condition. Moreover, we hypothesized that empathy would be triggered more easily in the direct gaze condition (when the self is the *target* of the other's emotion) because direct eye contact may signal communicative intent and is typically cueing inferences on the emotional and mental state of others (Kleinke, 1986). Also, it has been shown that direct (compared to averted) gaze is able to drive ToM areas in the brain (Kampe et al., 2003).

METHODS

Participants

Twenty-six subjects (12 men, mean age \pm *SD* = 24.4 \pm 3.0 years; 14 women, mean age \pm *SD* = 24.8 \pm 3.7 years) participated in the study which was approved by the local ethics committee. All participants (mean age \pm *SD* = 24.6 \pm 3.4 years) were right-handed, as assessed by the Edinburgh inventory (Oldfield, 1971), and were native speakers of German language. They were screened medically to rule out the use of medication affecting the central nervous system, a history of neurological or psychiatric disorders, head trauma, substance abuse, and other serious medical conditions. To control for sufficient abilities of verbalizing emotional states, subjects were screened for alexithymia using the German version of the Toronto Alexithymia Scale (TAS-26; Kupfer, Brosig, & Brähler, 2001). Subjects underwent neuropsychological testing, including standard measures of IQ, working memory, and attention. Moreover, they completed questionnaires on empathic abilities (Balanced Emotional Empathy Scale [BEES], Mehrabian, 1997; Empathic Concern Scale [ECS], Davis, 1983) for later analysis of a possible relationship between individual degrees of empathic capabilities and regional brain activation. The 30-item BEES extends the Emotional Empathic Tendency Scale (Mehrabian & Epstein, 1972) in that it is designed to balance "acquiescence bias" (i.e., the tendency of the majority of people to generally agree or disagree with statements included in a questionnaire). The BEES allows for a measure of one's tendency to empathize with the emotional experiences of others ("emotional empathy"). The ECS, a subscale of the Interpersonal Reactivity Index developed by Davis (1983), taps emotional aspects of empathic abilities in a similar way. For example, items of both scales include questions concerning one's own emotional state when observing the misfortune of another person or one's tendency to get caught up in the emotions of a crowd. Both scales can be easily administered as a paper-and-pencil test and have previously been used in functional neuroimaging studies to predict empathy-related brain activation (e.g., Singer et al., 2004).

Mean TAS-26, BEES, and ECS scores of participants involved in the present study and population norms are

summarized in Table 1. Individual scores on these scales and all individual neuropsychological data were within the range of reference population norms for all tests administered.

Stimuli

Facial stimuli were constructed using commercially available software for the generation of synthetic faces (FaceGen 3.0, Singular Inversions, Vancouver, Canada). With this software, faces can be rotated freely in 3-D space and morphed on diverse dimensions (e.g., emotion categories, gender, age) to create specific faces with differential degrees of distinct emotional expressions, thus allowing for a higher degree of control of stimulus materials than in studies where real faces are used. There were 318 faces preselected according to the naturalness of their emotional expressions by three independent judges from a pool of more than 1500 faces created randomly by the software. Validity of emotional expressions was confirmed in a behavioral study. Preselected faces were rated by volunteers ($n = 9$) according to the procedure devised by Ekman and Friesen (1976). Faces were presented with either an angry, fearful, sad, disgusted, happy, or neutral expression, and participants were asked to indicate which emotion category was displayed in a given face. Applying this procedure, 192 face stimuli were finally selected for the fMRI study (64 angry, 64 fearful, 64 neutral faces; half male and half female in each category). A minimum of 65% correct choices was required for the selection of a face. The average percentage (\pm *SD*) of correct choices was 86.4% (10.4%). This degree of agreement in categorical judgments of facial expressions of emotion is in accordance with behavioral studies in which the Ekman and Friesen face stimuli were used (Ekman, 1984). The mean percentage of correct choices (i.e., correct recognition of the intended emotional expression of a face) was counterbalanced across the stimulus categories (angry and fearful faces). A subset of the resulting stimulus set (30 stimulus faces) was rated by volunteers ($n = 19$) for emotional arousal (as described by Bradley

Table 1. Group Mean Scores and Respective Population Norms for TAS-26, BEES, and ECS

	BEES	ECS	TAS-26
Group mean \pm <i>SD</i>	41.58 \pm 30.23	18.31 \pm 4.00	2.38 \pm 0.47
Mean (norms) \pm <i>SD</i>	45 \pm 24 ^a	20.36 \pm 4.02 ^{a,b}	2.42 \pm 0.53

^aNorms for ECS and BEES refer to the original English version. In the present study, a German translation was used. For this German version, population norms are unavailable.

^bMean scores of norms for the ECS are estimated from separate scores for male and female samples provided by Davis (1980).

& Lang, 1999) to ensure that the synthetic stimulus faces evoked emotional responses sufficient for empathic reactions. A univariate analysis of variance (ANOVA) confirmed a significant effect of emotion category on the arousal ratings [$F(2, 36) = 98.33, p < .001$]. Post hoc *t* tests revealed that this was due to a significant stronger arousal in the fearful ($p < .001$) and the angry faces ($p < .001$) each compared to neutral faces.

The present study was restricted to (i) two emotion categories to retain statistical power sufficient to detect potential differences *between* emotion categories and (ii) negatively valenced emotions because the switching between negative and positive emotional valence might have confounded the experimental design. We did not use positively valenced emotional facial expressions for our experiment as it is difficult to find two positive basic emotions which can easily be distinguished. The selection of fear and anger for a study on empathy is also well in accordance with recent studies investigating face-based emotion recognition in brain-damaged patients. Paired deficits in recognition *and* production of the same emotion could be revealed for three emotions, namely, fear, disgust, and anger (see Goldman & Sripada, 2005 for a review).

Experimental Tasks

Subjects were asked to either concentrate on their own feelings that emerged when they were looking at each emotional facial expression (self-task), or on the emotional state expressed by the stimulus face (other-task). Besides the factor task (self/other-perspective), two additional factors were varied systematically. These were emotion (angry/fearful faces) and head direction (directed toward the observer/averted by $\sim 45^\circ$). This experimental setup resulted in a $2 \times 2 \times 2$ factorial design. In addition, a gender and age decision task on neutral faces (also with direct or averted head direction) was included as a high-level baseline condition. Neutral instead of emotional faces were used in the baseline condition because emotion processing is known to comprise unconscious, preconceptual stages such that empathic responses might implicitly emerge in the mere presence of an emotional face (Coricelli, 2005).

Each face stimulus was presented for 2 sec. Thereafter, a list of four 1- or 2-word descriptions appeared on the screen for 3 sec. The subject's task was to choose from this list of words the description which best depicted either the emotional state expressed by the preceding face (other-task) or the subject's own emotional response to the stimulus face (self-task). The emotional adjectives used for these judgments were randomly drawn from four sets of words, depending on the experimental condition (self-task with angry or fearful faces; other-task with angry or fearful faces). Choice words had been generated in a pilot study: subjects ($n = 6$) performed the same paradigm but

generated the words by themselves instead of choosing them from a given list. Four sets of 15 emotional adjectives were extracted from these responses corresponding to the dimensions self versus other, and fear versus anger (see below). Written word frequency (estimated by deriving frequencies from a collection of linguistic databases of German written language [COSMAS2, www.ids-mannheim.de/cosmas2, Institut für deutsche Sprache, Mannheim]), word length, arousal, and positivity were kept equivalent across tasks. In a separate pilot experiment, each of these words was rated on emotional scales (arousal and positivity) by 19 participants, as described by Bradley and Lang (1999). A two-factorial multivariate analysis of variance (MANOVA) including these measures as dependent variables yielded no significant effect of the factors emotion [i.e., anger versus fear; Rao's $R(5,52) = 0.95, p = .74$], task [i.e., self versus other; Rao's $R(5,52) = 0.97, p = .92$], or their interaction [Rao's $R(5,52) = 0.93, p = .53$]. Similarly, univariate two-factorial ANOVAs for each measure did not yield significant differences between the factors of interest.

For the high-level baseline, the following combinations of age and gender descriptions were employed: "older female," "older male," "younger female," "younger male," with the order of descriptions randomly varying from trial to trial.

During the fMRI measurement, subjects made their choices using four fingers of their right hand (excluding the thumb) and a four-button response device. The word lists were followed by an intertrial interval with a duration jittered between 0.2 and 1.2 sec. Random jitter was included to prevent correlation of event regressors. The two experimental tasks (attribution of emotion to oneself or the other) and the high-level baseline task (age/gender decision) alternated block wise in a pseudorandomized counterbalanced order. Each block consisted of a series of four trials and was cued by a preceding verbal instruction (duration: 2.5 sec) (see Figure 1). For stimulus presentation and response collection, the software Presentation 9.0 (Neurobehavioral Systems, Albany, CA, USA; www.neurobs.com) was used. During the fMRI experiment, the stimulus display was back-projected onto a screen mounted on the head coil.

Magnetic Resonance Technical Parameters

Subjects were scanned with a Siemens Magnetom SONATA whole-body system (Erlangen, Germany) at 1.5 T using a standard radio-frequency head coil. Thirty slices (0.4 mm interslice gap) were individually positioned to cover the whole brain with reference to a midsagittal scout image. During task performance, gradient-echo, echo-planar T2*-weighted images (EPI) were acquired using blood-oxygen-level-dependent (BOLD) contrast (TE = 66 msec, TR = 3020 msec, $\alpha = 90^\circ$, FOV =

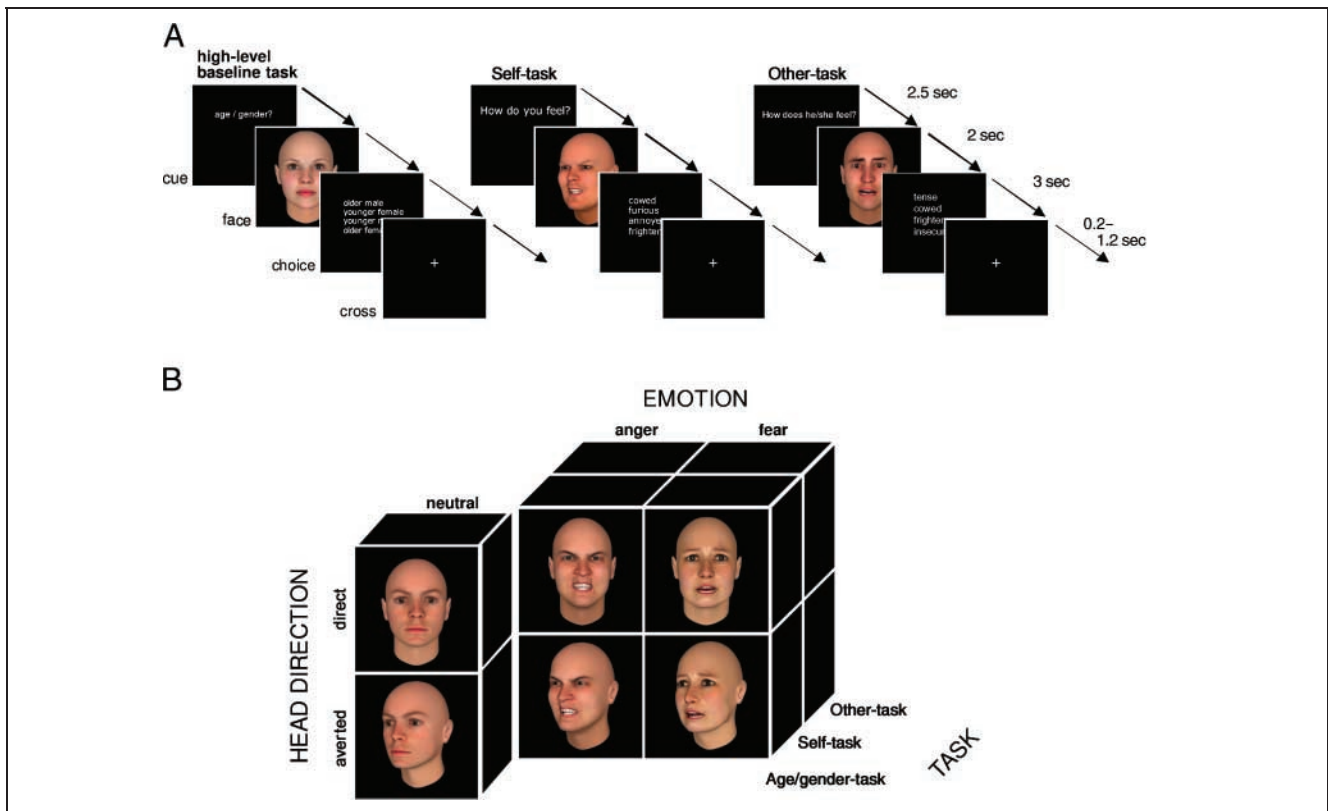


Figure 1. Experimental paradigm. (A) Time course of stimulus presentation. Trials were presented blockwise for the self-, other-, and high-level baseline tasks with each block consisting of four trials. The instruction cue was given only at the beginning of each block. The factors emotion category and gaze direction were pseudorandomized on a trial-by-trial basis. Choice words and task cues were originally presented in German language. (B) Experimental factors. We employed a $2 \times 2 \times 2$ factorial design with the factors task (self vs. other), head direction (direct vs. averted), emotion category (fear vs. anger), and a high-level baseline task with neutral faces and varying head direction (direct vs. averted).

200 mm, voxel size = $3.1 \times 3.1 \times 4 \text{ mm}^3$, matrix size = 64×64). Scanning was performed continuously during one single session lasting ~ 25 min. After functional neuroimaging, high-resolution anatomical images were acquired using a T1-weighted 3-D magnetization-prepared, rapid acquisition gradient-echo (MP-RAGE) pulse sequence (TE = 3.93 msec, TR = 2200 msec, TI = 1200 msec, $\alpha = 15^\circ$, FOV = 240 mm, matrix size = 179×256 , 128 sagittal slices, slice thickness = 1.33 mm).

Image Processing and Data Analysis

Preprocessing of imaging data and statistical analysis were accomplished using SPM2 (Wellcome Department of Imaging Neuroscience, London, UK; www.fil.ion.ucl.ac.uk/spm) implemented in MATLAB 6.5 (The Mathworks, Natick, MA, USA). The first four images of each functional time-series were discarded to allow for T1 equilibration effects. The remaining volumes were slice-time corrected, realigned (Friston, Ashburner, et al., 1995), and corrected for movement-by-field-inhomogeneity interactions using the realign and unwarp module of SPM2 (Andersson, Hutton, Ashburner, Turner, & Friston, 2001). Individual differences in brain anatomy were

accounted for by normalizing each subject's dataset to an EPI template of the Montreal National Institute (MNI) series implemented in SPM2, transforming all images to "MNI-space," a standard stereotactic reference space (which is approximately equivalent to the atlas of Talairach & Tournoux, 1988). Subsequently, functional volumes were spatially smoothed with an $8 \times 8 \times 8$ mm Gaussian kernel (full width half maximum) to compensate for remaining variation in gyral and sulcal anatomy across subjects, and to meet the requirements of the Gaussian random fields theory for subsequent statistical analysis. Individual T1 anatomical scans were coregistered to the corresponding mean image of the EPI time-series. Thereafter they were also normalized into MNI space, applying the same linear transformation parameters obtained by the normalization procedure of the functional volumes described above. Boxcar functions of 2 sec duration (corresponding to the on-time of each face stimulus) were convolved with a model of the hemodynamic response function (Friston, Holmes, et al., 1995), which thus comprised the presentation of facial stimuli, but not the word-choice procedure. Instruction cues were modeled in analogous manner using boxcar functions of 2.5 sec length, respectively. Parameter

estimates of the resulting general linear model were obtained for each subject and each voxel, corresponding to a mass-univariate approach. For population inference, a second level analysis was performed, using the parameter estimates for each task-related regressor in a within-subject ANOVA model (with subjects treated as random effects). Departures from sphericity assumptions were accommodated by applying the nonsphericity correction implemented in SPM2 (Glaser & Friston, 2003). Specific effects at each voxel were tested by applying appropriate linear contrasts to the parameter estimates. Differential contrasts of interest were calculated according to the experimental factors task (self-task vs. other-task, and vice versa), emotion (fear vs. anger, and vice versa), and head direction (directed to the observer vs. averted, and vice versa) to assess differential modulation of the BOLD signal induced by each factor. In addition, a conjunction analysis was performed to detect for activations common to the self- and the other-task (relative to the high-level baseline).

We report results that survive a statistical height threshold of $p < .05$, family wise error (FWE) corrected for multiple comparisons, and an extent threshold of at least 15 voxels. This conservative threshold was adopted to reduce the risk of false-positive activations (type II error). Additionally, for those brain regions targeted by a priori hypotheses (see Introduction), a region-of-interest (ROI) approach was applied, using a height threshold of $p < .05$ (FWE corrected across this particular region). To assess correlations between brain activations and individual measures of emphatic abilities (as assessed by the BEES and the ECS; see above), individual parameter estimates of regressors coding for the self- and other-tasks were extracted for regions previously implicated in the hMNS at voxels showing peak activation in the group conjunction analysis of the “self”- and the “other”-tasks (each relative to the high-level baseline). Individual means of these parameter estimates were correlated with individual scores on the BEES and ECS scales.

Localization of Activations

SPM_T maps resulting from the group analysis were superimposed onto a group mean MR image which was calculated from the normalized anatomical T1-images of each subject (see above). Standard stereotactic coordinates of the local maxima within areas of significant relative changes in neural activity associated with different experimental conditions were determined and anatomically localized by comparing activation maps superimposed on the anatomical group mean brain with a standard atlas of brain anatomy (Duvernoy, 1999). In addition, an anatomical SPM2 toolbox developed by Eickhoff et al. (2005) was applied, which allows for the integration of probabilistic cytoarchitectonic maps of the brain and functional neuroimaging data.

Prescanning and Postscanning Procedures

To accommodate subjects to the experimental tasks, each subject accomplished six training blocks prior to the fMRI measurement outside the scanner. Training stimuli were not included into the stimulus set used for the fMRI experiment. None of the subjects reported problems with the task or the required reaction times of less than 3 sec. Immediately after scanning, subjects rated all stimulus faces for the intensity of feelings (i) expressed by the face stimulus and (ii) evoked in the participants. In addition, participants completed a debriefing questionnaire on the naturalness of stimulus faces, as well as the levels of difficulty of the self-task, other-task, and high-level baseline task for each emotion and head direction condition.

RESULTS

Behavioral Data

Response Times during Scanning

A 3×2 ANOVA with the factors head direction and task showed a significant main effect of task [$F(2, 50) = 33.94$; $p < .001$]. Paired post hoc t tests revealed that subjects responded significantly faster in the high-level baseline task than both experimental tasks ($p < .001$ for each comparison). A $2 \times 2 \times 2$ ANOVA with the factors task, head direction, and emotion showed that reaction times were significantly faster in the other- than in the self-task [$F(1, 25) = 16.930$; $p < .001$]. Differences in reaction times between directed and averted head direction as well as angry and fearful emotion were nonsignificant. The interaction between the factors task and head direction just failed to reach significance [$F(1, 25) = 4.012$; $p = .056$]. This interaction resulted from faster responses to faces directed toward the observer (relative to averted faces) in the other-task [$t(25) = 2.003$, $p = .056$, two-sided], but not in the self-task [$t(25) = -0.55$, 0.587 ; post hoc t tests for matched pairs].

Postscanning Debriefing Questionnaire

Comparisons between groups of different stimuli (direct/averted head direction; angry/fearful emotion) did not yield statistically significant differences for the ratings of task difficulty and naturalness of faces. The respective data are provided in Table 2. Ratings of task difficulty differed significantly between tasks [Friedman ANOVA, $n = 26$; $\chi^2(df = 2) = 15.31$, $p < .001$]. The self-task was considered more difficult than both the other-task (Wilcoxon Test, $n = 26$, $Z = -3.856$; $p < .001$) and the high-level baseline task (Wilcoxon Test, $n = 26$, $Z = -2.97$; $p < .01$), whereas the ratings of levels of difficulty did not differ significantly between the other- and the high-level baseline task (Wilcoxon Test, $n = 26$, $Z = -0.41$; $p = .678$). Also, there was no significant difference in

Table 2. Results of the Postscanning Questionnaire

	<i>n</i>	<i>Direct</i> <i>Mean ± SD</i>	<i>Averted</i> <i>Mean ± SD</i>	<i>p</i>	<i>Fear</i> <i>Mean ± SD</i>	<i>Anger</i> <i>Mean ± SD</i>	<i>p</i>	<i>All</i> <i>Mean ± SD</i>
Difficulty emotional expression	26	2.48 ± 0.57	2.90 ± 0.79	.032*	2.88 ± 0.91	2.65 ± 0.82	.234	2.73 ± 0.50
Difficulty own emotion	26	3.19 ± 0.71	3.33 ± 0.90	.565	3.35 ± 0.77	3.04 ± 0.90	.145	3.22 ± 0.60
Difficulty age/gender detection	26	2.67 ± 0.70	2.68 ± 0.69	.700				2.68 ± 0.61
Naturalness of emotional faces	26	2.92 ± 0.92	2.96 ± 0.79	.778	2.92 ± 0.77	3.15 ± 0.81	.04*	2.99 ± 0.73
Naturalness of neutral faces	22	2.93 ± 0.82	2.93 ± 0.84	.952				2.93 ± 0.76

Comparison of mean ratings ($\pm SD$) given by the subjects on items of the postscanning questionnaire for the different experimental conditions. Rating scales ranging from 1 to 5 were used for the naturalness (1 = very unnatural, 2 = unnatural, 3 = moderately natural, 4 = natural, 5 = very natural) and the difficulty score (1 = very easy, 2 = easy, 3 = moderate, 4 = difficult, 5 = very difficult). Mean ratings were compared by means of a Wilcoxon matched-pairs test; p values for each test are given. n differs in one instance due to missing data.

SD = standard deviation; * = trend of statistical significance ($p < .05$, uncorrected).

the mean ratings of neutral and emotional faces' naturalness (Wilcoxon Test, $n = 23$, $Z = -0.112$; $p = .911$).

Correlations between Ratings of Difficulty and Reaction Times

To clarify whether increased difficulty ratings were related to longer reaction times, respective correlation analyses (Spearman rank-order correlation) were performed. Individual postscanning ratings of difficulty for each of the three tasks (self, other, high-level baseline) were compared to the mean reactions times related to performance of each task during scanning. There were no positive correlations for self ($r = .136$, $p = .506$) or other conditions ($r = -.164$, $p = .424$), and an inverse correlation for the high-level baseline task ($r = -.432$, $p < .05$). Furthermore, the difference between individual ratings of difficulty for the self- and the other-task was compared to the difference between the mean reaction times associated with the same conditions. This correlation analysis did not yield a significant result as well ($r = .012$, $p = .955$).

Postscanning Rating of Stimulus Faces

The mean ratings of emotion intensity for each stimulus face were compared by two-way 2×2 ANOVAs with the factors emotion and head direction. The intensity of emotion induced by the face stimuli in the subjects (self-task) was rated significantly higher in direct relative to averted head direction trials [$F(1, 25) = 7.859$; $p < .01$]. This was not paralleled by respective differential ratings of the intensity of the emotions expressed in the face stimuli (i.e., in the other task) [$F(1, 25) = 2.333$; $p = .139$]. Concerning the factor emotion, angry faces were rated

significantly higher for emotion intensity than fearful faces in the other-task [$F(1, 25) = 11.695$; $p < .01$], but not in the self-task [$F(1, 25) = 3.458$; $p = .075$]. Interactions between emotion and head direction were non-significant for either experimental task.

fMRI Data

Common Effects of the Self- and the Other-task (Relative to the High-level Baseline)

Areas of significant differential activation revealed by a conjunction analysis of the self- and the other-task (each compared to the high-level baseline) were located in the left STS and the left lateral orbito-frontal cortex (BA 47, extending into the inferior frontal cortex, BA 44 and 45), left MPFC, the left middle frontal gyrus, the left presupplementary motor area (pre-SMA), and the right cerebellum (see Table 3 and Figure 2A). Using an ROI approach (see Methods), activations were also significant in the right STS, the right inferior frontal gyrus (BA 45), and the temporal pole ($p < .05$, FWE corrected across ROI; see Table 3 and Figures 3 and 5).

Main Effect of the Self- versus the Other-task

The main effect of the self- (relative to the other-) task revealed significant differential bilateral activations of the temporo-parietal junction (TPJ), the MPFC, the posterior cingulate cortex (PCC)/precuneus, the middle frontal gyrus, the left inferior frontal gyrus, the pre-SMA, and the right cerebellum. The reverse contrast (other- vs. self-task) did not show significant differential activations, even at a lower statistical threshold ($p < .001$, uncorrected). Data are displayed in Table 4 and Figure 2B.

Table 3. Conjunction Analysis of the Self- and the Other-task (Each Compared to the High-level Baseline)

Anatomical Region	Side	Brodmann's Area	MNI coordinates			t
			x	y	z	
Inferior frontal gyrus pars orbitalis	L	47	-46	28	-8	12.70
Inferior frontal gyrus pars triangularis/opercularis [#]	L	44/45	-54	24	14	10.92
Inferior frontal cortex pars triangularis* ^a	R	45	56	32	0	4.82
Inferior frontal cortex pars triangularis* ^a	R	45	60	28	16	4.18
Cerebellum	R		28	-78	-40	10.54
Superior temporal sulcus	L	21	-62	-46	2	10.22
Superior temporal sulcus* ^b	R	21	52	-32	-2	5.04
Superior frontal gyrus pars medialis	L	9	-8	54	42	6.95
Superior frontal gyrus pars medialis [#]	L	9/10	-8	58	34	6.32
Middle frontal gyrus	L	6/9	-44	8	52	6.03
Supplementary motor area	L	6/8	-4	20	50	5.65
Temporal pole* ^{#,c}	R	38/20	46	16	-36	5.21
Temporal pole* ^{#,c}	L	38/20	-48	12	-36	4.30

Group activations for the conjunction of the self- and the other-task (each compared to the high-level baseline), $p < .05$, FWE corrected (except * $p < .05$, corrected for multiple comparisons across a small volume of interest).

[#]Local submaxima within a bigger cluster of activation.

ROIs: ^aright inferior frontal cortex pars triangularis/opercularis; ^bright superior temporal gyrus, right middle temporal gyrus; ^cbilateral temporal poles.

L = left hemisphere; R = right hemisphere; Coordinates are given in MNI space.

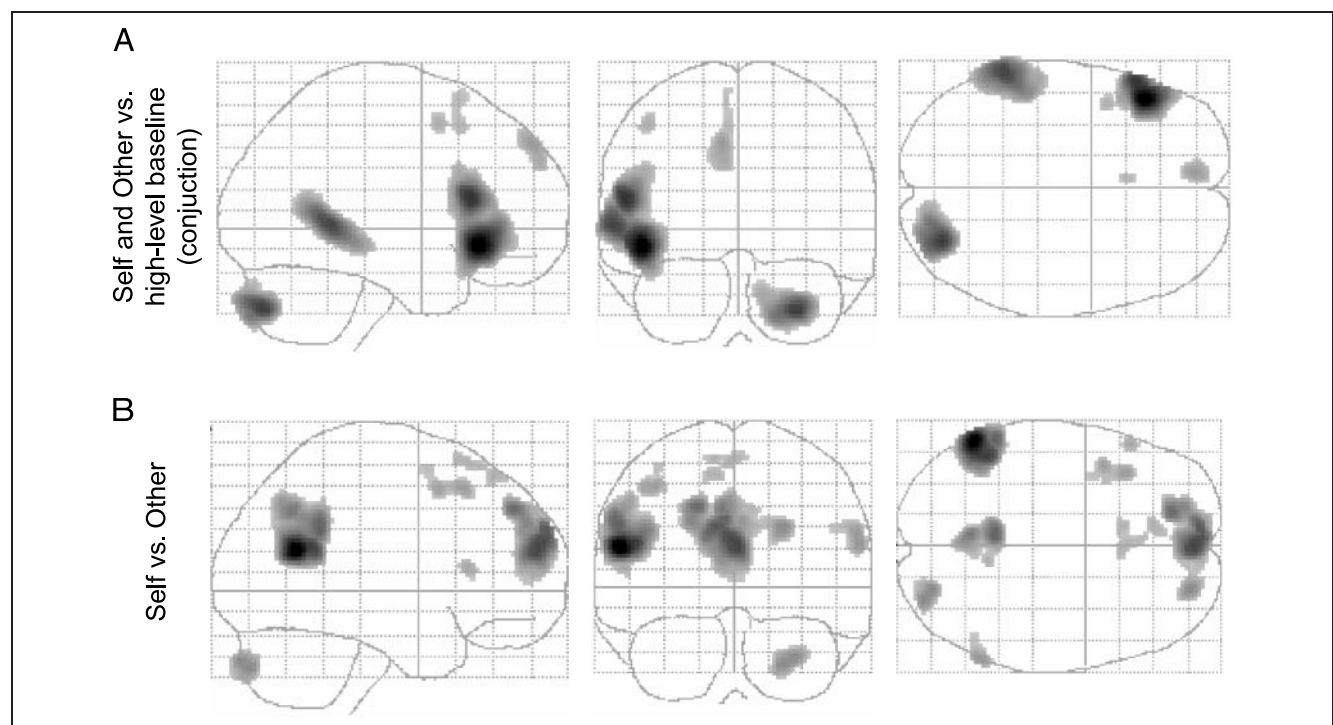


Figure 2. (A) Statistical parametric map (SPM) showing the group activations for the conjunction of the self- and the other-tasks (each vs. the high-level baseline). (B) SPM of group activations for the contrast between the self- and the other-task. SPMs depict statistically significant relative increases in neural activity at $p < .05$, FWE corrected for multiple comparisons across the whole brain volume (see Tables 3 and 4 for details).

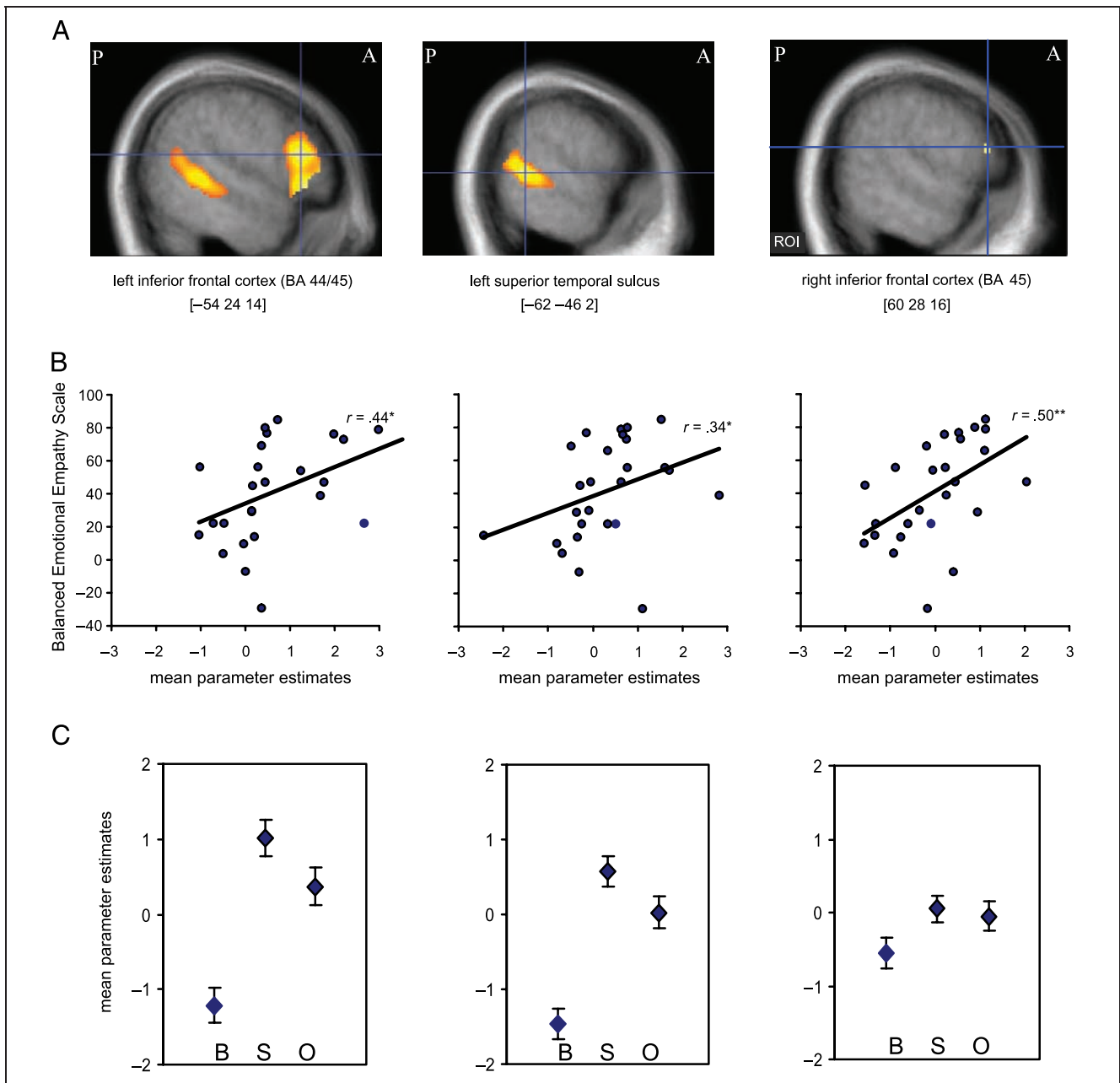


Figure 3. Covariation of mirror neuron activation and empathic ability. (A) Increase of brain activation in regions associated with the hMNS, as revealed in the conjunction analysis of self- and other-tasks (each vs. high-level baseline). Activations are superimposed on the anatomical group mean image (see Methods), depicting statistically significant relative increases in neural activity at $p < .05$, FWE corrected across the whole brain volume or (if indicated) at $p < .05$, FWE corrected for ROI. (B) Corresponding significant correlations ($*p < .05$, $**p < .01$) between individual empathy scores on the BEES (Mehrabian, 1997) and activation level (mean parameter estimates) in peak voxels of the depicted regions (see Table 5 for details). Coordinates are given in millimeters (mm). Solid lines represent the linear best fit. r = correlation coefficient. (C) Plots show the mean and standard errors of condition-specific parameter estimates, which reflect the relative contribution of each condition to the amplitude of the adjusted BOLD signal relative to the fitted mean in the respective brain region. Note, the scaling of mean parameter estimates is in arbitrary units. As the experimental design did not include a low-level baseline, absolute values cannot be interpreted, for example, as deviations from a physiological baseline. In particular, negative values *do not* indicate a deactivation in comparison to a resting state baseline (as described, for example, by Gusnard et al., 2001). S = self-task; O = other-task; B = high-level baseline task.

To extract neurofunctional components which were differentially implicated in the self-related processing of faces, but not the other-related task, the self versus other comparison was exclusively masked with the other

versus high-level baseline contrast (at $p < .05$, uncorrected). Masking corroborated activations in the regions indicated above, except in the left inferior frontal cortex and the right cerebellum.

Table 4. Contrast between the Self- and the Other-Task

<i>Anatomical Region</i>	<i>Side</i>	<i>Brodmann's Area</i>	<i>MNI coordinates</i>			<i>t</i>
			<i>x</i>	<i>y</i>	<i>z</i>	
Temporo-parietal junction	L	21/22/39	-56	-60	20	9.52
Supramarginal gyrus [#]	L	40	-56	-48	30	7.51
Temporo-parietal junction	R	21/22	62	-54	20	6.15
Inferior frontal cortex pars triangularis/opercularis	L	44/45	-54	24	10	5.73
Superior frontal gyrus pars medialis	R	10	2	58	20	8.00
Superior frontal gyrus pars medialis [#]	L	10	-8	62	30	7.66
Superior frontal gyrus [#]	L	9	-18	46	40	7.28
Superior frontal gyrus pars medialis	L	8/9	-10	36	56	5.49
Gyrus rectus ^{*a}	L	11	-6	56	-18	5.06
Middle orbital gyrus ^{*a}	R	11	2	40	-8	4.41
Anterior cingulate cortex ^{*a}	M	24	0	28	18	4.34
Supplementary motor area	L	6/8	-10	20	64	5.45
Supplementary motor area [#]	R	6/8	2	20	64	5.44
Superior frontal gyrus pars medialis [#]	L	8	-4	30	60	5.34
Middle frontal gyrus	L	9	-38	22	50	6.06
Posterior cingulate cortex	L	23	-6	-48	32	7.41
Precuneus [#]	L	7	-6	-62	38	6.88
Precuneus/posterior cingulate cortex [#]	R	7	8	-52	30	5.52
Cerebellum	R		30	-86	-34	6.48

Group activations for the contrast between the self- and the other-task (the main effect of task), $p < .05$, FWE corrected (except $*p < .05$, corrected for multiple comparisons across a small volume of interest).

[#]Local submaxima within a bigger cluster of activation.

ROIs: ^amedial prefrontal cortex.

L = left hemisphere; R = right hemisphere; Coordinates are given in MNI space.

Direct versus Averted Head Direction

The main effect of emotional facial expressions with direct (relative to averted) head direction revealed bilateral activations of the calcarine sulcus. The reverse contrast (averted vs. direct head direction) did not show any significant differential activation.

Fearful versus Angry Emotion Expressed by the Faces

At $p < .05$ (FWE corrected), no significant differential activations were associated with the contrasts between fearful and angry emotional facial expressions.

Interactions

At $p < .05$ (FWE corrected), there were no significant interactions between the experimental factors.

Covariance of Individual Empathy Scores and Local Brain Activation

Table 5 summarizes the covariance between local brain activity and measures of individual differences in empathic abilities for regions implicated in the hMNS which were activated in the conjunction analysis of the self- and the other-task (each relative to the high-level baseline). The inferior frontal gyrus (bilaterally) and the left STS showed a significant correlation with empathy scores on the BEES (Mehrabian, 1997). Correlations between the ECS (Davis, 1983) and brain activation were less pronounced, albeit still significant for the left STS and right inferior frontal gyrus. Correlations are illustrated in Figure 3.

DISCUSSION

The present study aimed at identifying the neural systems mediating key components of empathy. Using

Table 5. Correlations between Mirror Neuron Activation and Individual Empathy Scores

<i>Anatomical Region</i>	<i>MNI coordinates</i>			<i>BEES</i>		<i>ECS</i>	
	<i>x</i>	<i>y</i>	<i>z</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Left inferior frontal gyrus pars triangularis	-54	24	14	.44*	<.05	.29 [#]	.078
Right inferior frontal gyrus pars triangularis ¹	60	28	16	.50**	<.01	.37*	<.05
Left superior temporal sulcus	-62	-46	2	.34*	<.05	.41*	<.05
Right superior temporal sulcus	52	-32	-2	.13	.54	.16	.42

Covariation of measures of empathic abilities and brain activation for regions activated in the conjunction analysis of the self- and the other-task (each compared to the high-level baseline).

r = correlation coefficient, [#] = trend toward significance ($p < .1$); BEES = Balanced Emotional Empathy Scale; ECS = Empathic Concern Scale (see text for references). All regions were significantly activated in the conjunction analysis ($p < .05$, FWE corrected; except ¹ $p < .05$, FWE corrected across a small volume of interest).

fMRI, we demonstrated that the attribution of feelings either to an emotional facial expression or to oneself (in response to a face) recruits brain areas involved in emotion processing, mirror neuron mechanisms, and ToM. More importantly, we were able to disentangle brain activations related to the self- or the other-perspective in dyadic face-to-face interactions by directly contrasting the self- and other-related attribution of emotions.

An important point of the current experiment was the attempt to create an interactive and interpersonal context in which empathy could emerge. Although interpersonal interaction is experimentally limited (e.g., the subject is lying in the MR scanner, perceives emotional facial expressions, and is instructed to evaluate these social cues with respect to the “agents” coded by the stimulus faces or with respect to themselves), the current paradigm is well designed to study components of emotional interpersonal interaction. It accounts for a key aspect of empathy, namely, that it relies on a dyadic social situation where the inferred emotional state of another person directly affects oneself as the target of the other’s emotion. Previous studies have successfully used such virtual social settings to study “on-line” social interaction (e.g., Schilbach et al., 2005).

Self- and Other-perspectives and Empathy

Mirror Neuron Mechanisms Involved in Empathy

Both the self- and the other-task (compared to the high-level baseline) activated the inferior frontal gyrus bilaterally, including Broca’s area (BA 44/45) and its homologue in the right hemisphere, which have previously been implicated in the hMNS. Furthermore, this area was differentially activated in the self- relative to the other-task, indicating that mirror neuron mechanisms may play a key role in the attribution of emotions to oneself during processes of empathic interpersonal cognition. This assumption is in accordance with the adaptation of the mirror neuron approach (see Introduction) to mental processes occurring in social interactions, for

example, the detection of intentions (Iacoboni et al., 2005), emotional states (Carr et al., 2003), empathy, or mind-reading (Gallese, 2003). The operation of mirror neuron mechanisms in emotional social interactions may constitute the neural basis of the observer’s resonance of the emotional state of another individual, for example, by means of some kind of inverse mapping of the underlying neural representations. Further evidence for such a mechanism comes from electromyographic studies of covert facial muscle activity during the passive viewing of emotional faces (Dimberg, Thunberg, & Elmehed, 2000; Dimberg & Thunberg, 1998). In these studies, muscles, which are involved in the creation of emotional facial expressions, were also activated during the mere observation of the faces.

The mirror neuron account of imitation postulates that visual representations of observed movements are formed in the STS and relayed via the parietal cortex to inferior frontal areas, thereby converting visual information into a motor plan (Iacoboni et al., 1999). When an action is executed, activity in frontal areas is relayed back to the STS, which then transforms the motor plan back into some predicted visual representation (Iacoboni et al., 2001). The inferior frontal gyrus may thus constitute a “mirror” interface for the ability to imitate other people’s actions. It is conceivable that such mechanisms of interpersonal motor behavior are involved in social communication in a more general sense, especially in the context of face-to-face interaction. A facial expression of emotion can be referred to as a goal-directed social action used to intentionally affect the mental state of others (Wolpert et al., 2003). Similar to motor actions, the observation of emotions expressed by a face might be “mirrored” in the brain to help understand the underlying social meaning by facial mimicry. Functional neuroimaging data demonstrating inferior frontal activation during both imitation and passive viewing of emotional faces (although less pronounced during passive observation) are in line with this view (Dapretto et al., 2006; Carr et al., 2003). Our data corroborate and extend these previous findings in that they show mirror mecha-

nisms acting as key players in the decoding of socially meaningful emotional expressions and the interpersonal attribution of emotions. This hypothesis is also supported by a study of Lawrence et al. (2006), who reported activation of Broca's area and motor/premotor regions during a task which required the decoding of mental and emotional states from video clips showing gestures, body postures, and facial expressions. Moreover, a study by Decety and Chaminade (2003) showed mirror neuron activation (in both emotion processing areas and frontoparietal networks including the inferior frontal gyrus) during the passive viewing of actors telling a story in cases where an actor's emotional facial expression corresponded to the emotional content of a story.

A number of neuroimaging experiments on the decoding of emotion from facial displays failed to reveal inferior frontal or other mirror neuron activation (cp. Adolphs, 2002 for a recent review; but see, for example, Gorno-Tempini et al., 2001; George et al., 1993). We suggest that this may depend on the fact that perceptual tasks not embedded in an interpersonal context were used in these studies. As a result, emotional facial expressions might not have been perceived as (target-directed) interpersonal actions. In contrast, by introducing a task which required the evaluation of both other persons' emotional facial expressions *and* one's own reaction to the observed emotional faces, we explicitly focused on the dyadic social nature of emotional face-to-face communication.

Although the subjects' choice options for the evaluation of faces were more varied in the emotional task than the high-level baseline, it is unlikely that the observed differential brain activations (e.g., in the inferior frontal cortex) could be attributed to increased task difficulty resulting from a higher "effort" of reading the choice words in the emotional face conditions. The training session prior to the fMRI measurement can be assumed to sufficiently have familiarized participants with the possible response options for all conditions. Condition-specific differences in the variability of choice words are thus not likely to have caused differences in reading effort as could have been the case, for example, with new words in every trial. Moreover, postscanning ratings of task difficulty differed significantly between the self- and the other-task (where choice word variability was comparable), but *not* the other- and the high-level baseline task. This agrees well with the finding of significantly longer reaction times in the self- relative to the other-task. Postscanning data and reaction times both converge on the view that something inherent in the self-condition rendered the task more difficult for the volunteers than the other-task, but not a confound. The demands of self-reflection and the self-other distinction which were differentially associated with the self-condition may well have caused an increased level of difficulty, as reflected by the related longer reaction times and higher ratings of difficulty. Reaction time differences between the emo-

tional tasks and the high-level baseline are also likely to reflect the operation of additional cognitive processes required by the evaluation of emotional facial expressions which we intended to measure. Moreover, reaction times and postscanning ratings of task difficulty were not positively correlated for any of the conditions included in the experiment, also suggesting that the condition-specific differences in behavioral performance depended on the controlled experimental manipulation, rather than on confounding variables.

Although the portion of the inferior frontal gyrus and the pre-SMA, which were activated by our experimental task, are also involved in language processing (Grezes & Decety, 2001; Poldrack et al., 1999), it is unlikely that activations in this area were related to overt language production in the present study: Participants were silently looking at the stimulus display and reading was involved in all conditions. Moreover, the significant covariation of activation in the inferior frontal cortex and individual empathic abilities renders a language-dependent involvement of this region unlikely (see Table 5), in particular, as there was a corresponding correlation between empathic abilities and neural activity also in BA 44/45 in the right hemisphere. This is also in good accordance with the notion of a bilateral representation of the hMNS (Aziz-Zadeh, Koski, Zaidel, Mazziotta, & Iacoboni, 2006).

Motor-related regions activated in the present study include the SMA and the cerebellum. The maximum of SMA activation was located in the pre-SMA, a subregion which has specifically been implicated in more cognitive functions than the SMA proper. In particular, pre-SMA activation has been related to motor imagery and the observation of movements with the aim to imitate (Grezes & Decety, 2001). There is also evidence that neural activity in the pre-SMA can be differentially modulated by emotionally triggered movements (Oliveri et al., 2003), suggesting a role of this area in the transformation of emotional experiences into motor actions. These findings converge with our data showing differential pre-SMA engagement during the self-referential decoding of emotion from facial expressions, that is, during a task which does not rely only on emotional "contagion" (possibly mediated by mirror neuron mechanisms), but also on higher cognitive functions such as ToM and mental imagery.

Taken together, activations revealed by the conjunction analysis of the self- and the other-task are in accordance with our suggestion that mirror neuron mechanisms operate during the self- and other-related decoding of emotional facial expressions. Moreover, the stronger mirror neuron activation during the self-compared to the other-condition and the accompanying specific involvement of areas involved in higher cognition suggest specific roles of mirror mechanisms and cognitive control in the self-attribution of emotion.

As expected, subjects did not always attribute the same emotion to themselves which was displayed by a

stimulus face in the self-condition. Rather, reactive responses were highly frequent (see Appendix). This agrees with previous studies demonstrating that facial mimicry to observed emotional expressions also occurs in the *absence* of a contagious emotional response (e.g., Hess & Blairy, 2001) and may be correlated with individual differences in empathic ability (Sonnby-Borgstrom, 2002). The exact neural mechanisms of contagious and reactive conscious emotional reactions and their relationship to unconscious mimicry of emotional facial expressions during face-to-face interaction are unknown, to date. These questions, however, go beyond the concerns of the present study.

Individual Empathy Scores and Brain Activation

For several regions of the hMNS which were activated in the conjunction analysis of the self- and the other-task, a significant covariation of the intensity of local activation with individual empathy scores was observed (see Figure 3). Subjects who scored higher on the BEES (Mehrabian, 1997) showed stronger activations in the left and right inferior frontal cortices, as well as in the left STS. These activations may reflect unconscious mimicry of observed facial expressions supporting the ability to infer feelings from other persons' faces. Findings from behavioral studies which demonstrate that individuals with high scores on empathy scales are more likely to exhibit such facial mimicry corroborate this assumption (Sonnby-Borgstrom, 2002).

To our knowledge, the present study is the first to report activation of the inferior frontal cortex (including BA 44/45) in a task where subjects explicitly reflected either on their own or someone else's emotional state (represented by our stimulus materials showing emotional facial expressions), in the *absence of any instruction to consciously imitate the observed face*. Previous studies which reported inferior frontal activation (Dapretto et al., 2006; Leslie, Johnson-Frey, & Grafton, 2004; Carr et al., 2003) used experimental settings with alternating blocks of imitation and observation. It is conceivable that brain activation during observation was influenced by the subjects' "preparedness" to imitate (or, alternatively, the suppression thereof), thus introducing motor components into the observation task. Our finding of hMNS activation in a paradigm, which required the "on-line" self- and other-related decoding of emotions from faces without any imitation condition, strongly supports the importance of mirror neuron mechanisms in emotional interpersonal communication.

The "Self" in Shared Emotional Experience

If it is true that similar neural mechanisms are recruited during the execution and perception of diverse types of behavior (e.g., body movements, emotional behavior, social interaction), the question then arises how the brain

distinguishes between self- and other-related processing. In this respect, empathizing cannot be considered only as the sharing of an emotional state. Rather, it also requires the maintenance of a distinction between the self and the other (Decety & Jackson, 2004), especially when an emotional response occurs that is not similar to the observed emotion. In the present study, we observed differential self-related activation of the MPFC, that is, in a brain area involved in ToM (Gallagher & Frith, 2003), social cognition (cp. Ochsner et al., 2004 for a review), and diverse types of self-related processing (Kampe et al., 2003; Ruby & Decety, 2003; Vogeley & Fink, 2003). Evidence suggests that social cognitive judgments primarily rely on the dorsal MPFC (dMPFC) regions, whereas self-referential cognitive and emotional processing additionally involve the more ventral aspects (vMPFC; roughly defined by $z < 10$) of the MPFC. It is consistent with this hypothesis that in our study, self-related MPFC activations were not restricted to the dMPFC, but rather extended into the ventral aspect of the MPFC (see Figure 4). Interestingly, a similar pattern of results was recently reported in the work of Mitchell, Macrae, and Banaji (2006) (see also Mitchell, Banaji, & Macrae, 2005 for a related study). The authors report a double dissociation of ventral and dorsal aspects of the MPFC when perceivers made "mental state inferences" (in the sense of preference statements) about themselves or another person who either resembles themselves, or not. The vMPFC was significantly stronger activated in that study when subjects were mentalizing about similar than dissimilar others, whereas the dMPFC showed the opposite pattern. In the vMPFC, activations were comparable for judgments referring to oneself or a similar other (compared to a dissimilar other), whereas the dMPFC area did not show significant differences in activations associated with self- relative to other-conditions. Our study corroborates and extends these findings by showing evidence that the dissociation of self- and other-related mentalizing also occurs during the "on-line" interpersonal attribution of emotions cued by angry and fearful facial expressions. Anatomical differences in the MPFC also support a ventral-dorsal functional differentiation: The vMPFC is strongly interconnected with the amygdala, ventral striatum, and orbito-frontal cortex (thereby relating the ventral region to emotion processing), whereas the dMPFC lacks such interconnections (indicating more cognitive functions of the dorsal area) (Ongur, Ferry, & Price, 2003). Our data thus contribute to the current knowledge of vMPFC involvement in self-referential and emotional processing, in that they demonstrate differential activations in this region when one introspectively assesses his or her feelings in response to emotional facial expressions.

It is conceivable that the ability to attribute emotions to oneself and to others is at least, to some degree, based on memories, in particular, on episodic emotional autobiographical memory. This assumption could explain activation of memory-related areas in our paradigm. The basal

Figure 4. MPFC activations for the conjunction of self and other versus the high-level baseline and for the contrast self versus other. Activations are superimposed on the anatomical group mean image (see Methods), depicting statistically significant relative increases in neural activity at $p < .05$, FWE corrected for ROI (see Table 4 for the exact MNI coordinates, and Figure 3 for an explanation of BOLD signal change plots).

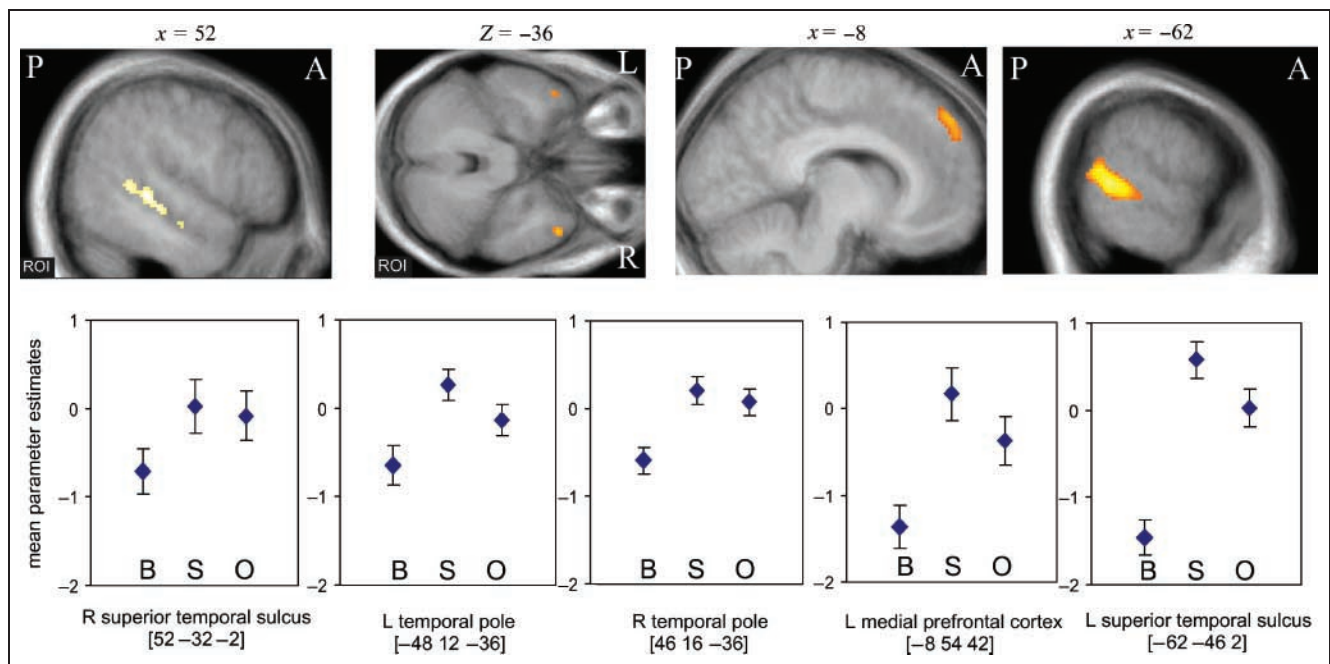
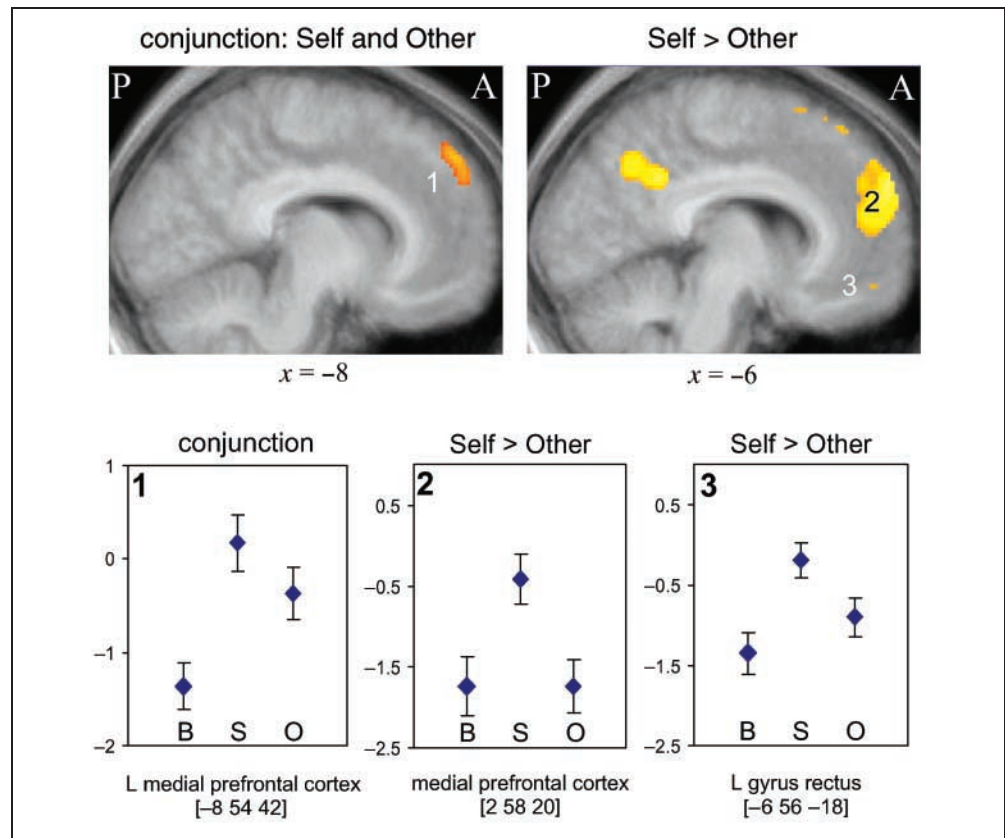


Figure 5. Group activations for the conjunction of the self- and the other-tasks (each vs. the high-level baseline) in regions associated with ToM. Activations are superimposed on the anatomical group mean image (see Methods), depicting statistically significant relative increases in neural activity at $p < .05$, FWE corrected across the whole brain volume or (if indicated) at $p < .05$, FWE corrected for ROI (see Table 3 for the exact MNI coordinates and Figure 3 for an explanation of BOLD signal change plots).

temporal cortices including the temporal poles (which were activated in both the self- and the other-task) have been implicated in emotional autobiographical memory retrieval (Piefke, Weiss, Zilles, Markowitsch, & Fink, 2003; Fink et al., 1996), and the recollection of familiar faces and scenes (Sugiura, Shah, Zilles, & Fink, 2005; Nakamura et al., 2000). One might thus speculate that our subjects used their own personal past experiences to gain access to other persons' thoughts and emotional states.

The medial parietal and adjacent posterior cingulate cortices have been implicated in emotional and non-emotional types of social cognition, for example, in making forgivability judgments in social scenarios (Farrow et al., 2001) and watching social interactions (Iacoboni et al., 2004). Furthermore, an important role of the anterior precuneus in self-related information processing has been suggested (e.g., Shah et al., 2001; see Cavanna & Trimble, 2006 for a review). These data are consistent with our finding of precuneus/PCC involvement in emotional aspects of self-related cognition, specifically, in the evaluation of one's own emotional state during interpersonal face-to-face interaction.

ToM Mechanisms and Perspective Taking

In the present study, the MPFC, the TPJ, and basal temporal cortices, including the temporal poles, were activated in the conjunction of the self- and the other-tasks (each compared to the high-level baseline; see Figure 5). This is consistent with the idea that ToM abilities are crucial for empathic understanding, in particular, for the more cognitive aspects of empathy. However, circumscribed TPJ and MPFC areas were also exclusively involved in the self-task. This pattern of activation suggests the differential operation of distinct sub-modules in the medial prefrontal and temporo-parietal cortices in self- or other-related processing within a broader network supporting mentalizing abilities.

There is a current debate on the question whether areas at the TPJ are involved in the "preprocessing" of social signals that "aid" ToM reasoning (Gallagher & Frith, 2003) or in ToM itself (Saxe & Kanwisher, 2003). Our data suggest that the two functions may be implemented in separate TPJ regions. The conjunction analysis of the self- and the other-task showed STS/TPJ activations in a region which has been implicated in the processing of socially relevant cues (Frith & Frith, 2003). Such processing of social cues occurred during the decoding of facial features signaling an emotional state in both of our experimental tasks. A more posterior TPJ area (extending into the inferior parietal cortex) was activated differentially during the self-related attribution of emotion. This region has been implicated in agency (Jeannerod, 2001), self-related visuospatial perspective-taking (Vogeley & Fink, 2003; Ruby & Decety, 2001), but also in the other-perspective (e.g., Ruby & Decety, 2003). The posterior TPJ and the inferior parietal cortex could thus play a role

in the *self-other distinction* (Decety & Sommerville, 2003). Following this hypothesis, we suggest that these regions act as mediators between the self- and the other perspective, thereby allowing us to mentally separate the observer's view from the perspective of ourselves. Although both perspectives may sometimes overlap, we are usually able to disentangle the two points of view. It is in good accordance with such an account, that the misattribution of self-generated actions to an external source in schizophrenic patients has been related to the abnormal function of the inferior parietal cortex (Brunelin et al., 2006; Spence et al., 1997).

The hypothesis of a temporo-parietal mechanism for the differentiation between the self- and the other-perspective may also help explain TPJ and inferior parietal activations in ToM studies using false-belief tasks (Saxe & Kanwisher, 2003). To successfully answer questions on typical false-belief stories (as used, for example, by Gallagher et al., 2000), simultaneous access to one's own and someone else's knowledge about a specific situation is required, but both perspectives also need to be *distinguished* correctly.

Direct versus Averted Head Direction

Although of minor importance with regard to the aim of the present study, the effect of direct (versus averted) head direction showed a significant differential effect in the postscanning rating of emotion intensity: As expected, faces directed toward the observer (relative to faces with averted head direction) enhanced one's own emotional response (self-task), but did not affect the intensity of emotion perceived in the facial expressions (other-task). These results suggest that direct gaze affects empathic reactions on the behavioral level. However, this behavioral effect was not paralleled by differential activations beyond early visual processing areas. We suggest that this discrepancy may result from the fact that the effect might have been too small for causing significant changes in activation patterns. Instead, we observed activation related to direct (relative to averted) head direction in the striate cortex only. This difference in the early perceptual domain is likely to reflect visual differences in the stimuli.

Fearful versus Angry Facial Expressions

We did not observe differential activations related to the processing of angry versus fearful stimulus faces. This finding suggests that our paradigm tapped a general mechanism involved in the interpersonal decoding of at least negatively valenced emotional faces rather than specific responses to distinct emotion categories. Further studies are needed to investigate brain mechanisms underlying the interpersonal evaluation of faces expressing other negatively and positively valenced emotions to corroborate this assumption.

Conclusions

Our present data support the notion that mirror neuron and ToM mechanisms are involved in empathy. Importantly, we show evidence for an hMNS engagement in an empathy-related experimental paradigm without explicit motor component (e.g., imitation), corroborating the view that mirror neuron mechanisms are not only involved in motor behavior but also act as key players in emotional

interpersonal cognition. The differential engagement of the MPFC, the PCC/precuneus, and temporo-parietal subregions in the self-related attribution of emotion indicates a neurofunctional segregation within the brain circuits subserving emotional face-to-face interactions (which include prefrontal, temporal, and parietal areas). Furthermore, our data suggest that posterior TPJ and inferior parietal regions mediate the distinction between the self- and the other-perspective in emotional interpersonal cognition.

APPENDIX

I. List of Words Used as Response Alternatives in the Emotional Conditions and the Associated Frequency of Choices (German)

<i>Self-task</i>				<i>Other-task</i>			
<i>Fearful Faces</i>		<i>Angry Faces</i>		<i>Fearful Faces</i>		<i>Angry Faces</i>	
besorgt	90	angegriffen	102	erschüttert	79	angriffslustig	98
beunruhigt	90	unbehaglich	98	hilflos	79	aggressiv	98
betroffen	87	unangenehm	85	hilfesuchend	73	hasserfüllt	74
mitleidig	81	gereizt	71	erschrocken	66	verärgert	67
mitleidend	78	verunsichert	67	geschockt	65	wütend	66
gefordert	73	unsicher	64	angstvoll	61	zornig	65
hilflos	67	ängstlich	49	besorgt	56	sauer	57
verunsichert	49	verärgert	49	ratlos	55	rachsüchtig	56
verstört	42	verängstigt	48	schockiert	53	erbost	48
angespannt	41	verachtet	43	ängstlich	52	grimmig	45
traurig	31	aggressiv	38	unsicher	46	gehässig	43
ängstlich	26	genervt	32	verwirrt	46	misslaunig	41
schuldig	25	verletzt	28	verletzt	40	angegriffen	33
nervös	22	wütend	25	überrumpelt	37	hämisch	19
überfordert	22	zornig	23	angespannt	19	genervt	13

II. List of Words Used as Response Alternatives in the Emotional Conditions and the Associated Frequency of Choices (English Translation)

<i>Self-task</i>				<i>Other-task</i>			
<i>Fearful Faces</i>		<i>Angry Faces</i>		<i>Fearful Faces</i>		<i>Angry Faces</i>	
afraid	90	attacked	102	appalled	79	belligerent	98
alarmed	90	uneasy	98	helpless	79	aggressive	98
concerned	87	uncomfortable	85	defenseless	73	hating	74
pitiful	81	testy	71	scared	66	annoyed	67
compassionate	78	uncertain	67	startled	65	angry	66
required	73	insecure	64	anxious	61	irate	65
helpless	67	fearful	49	afraid	56	sulky	57

APPENDIX II. (continued)

<i>Self-task</i>				<i>Other-task</i>			
<i>Fearful Faces</i>		<i>Angry Faces</i>		<i>Fearful Faces</i>		<i>Angry Faces</i>	
insecure	49	annoyed	49	perplexed	55	revengeful	56
deranged	42	scared	48	shocked	53	furious	48
tense	41	disdained	43	fearful	52	fierce	45
sad	31	aggressive	38	insecure	46	nasty	43
fearful	26	irritated	32	confused	46	grumpy	41
guilty	25	harmed	28	harmed	40	attacked	33
nervous	22	angry	25	upset	37	malicious	19
overstrained	22	irate	23	tense	19	irritated	13

In the tables, numbers given after each word indicate how often the respective word was chosen by the subjects for responses in each experimental condition.

Acknowledgments

M. P. was supported by the Hans-Lungwitz-Foundation (Berlin) and the START program of the RWTH Aachen. H. J. M. is currently a fellow at the Hanse Institute of Advanced Study. G. R. F. is supported by the Deutsche Forschungsgemeinschaft (DFG; KFO 112) and the Bundesministerium für Bildung und Forschung (BMBF; Brain Imaging Center West; Förderkennzeichen 01GO0509). We thank our colleagues in the MR and Cognitive Neurology groups at the Institute of Medicine (Research Center Jülich) for their support and helpful advice.

Reprint requests should be sent to Martin Schulte-Rüther, Institute of Neurosciences and Biophysics—Medicine, Research Center Jülich, Leo-Brand Str. 5, 52425 Jülich, Germany, or via e-mail: m.schulte@fz-juelich.de.

REFERENCES

Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, 1, 21–62.

Andersson, J. L., Hutton, C., Ashburner, J., Turner, R., & Friston, K. (2001). Modeling geometric deformations in EPI time series. *Neuroimage*, 13, 903–919.

Aziz-Zadeh, L., Koski, L., Zaidel, E., Mazziotta, J., & Iacoboni, M. (2006). Lateralization of the human mirror neuron system. *Journal of Neuroscience*, 26, 2964–2970.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings* (Rep. No. Technical Report C-1). The Center for Research on Psychophysiology, University of Florida.

Brunelin, J., Poulet, E., Bediou, B., Kallel, L., Dalery, J., D'amato, T., et al. (2006). Low frequency repetitive transcranial magnetic stimulation improves source monitoring deficit in hallucinating patients with schizophrenia. *Schizophrenia Research*, 81, 41–45.

Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences, U.S.A.*, 100, 5497–5502.

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus:

A review of its functional anatomy and behavioural correlates. *Brain*, 129, 564–583.

Coricelli, G. (2005). Two-levels of mental states attribution: From automaticity to voluntariness. *Neuropsychologia*, 43, 294–300.

Critchley, H., Daly, E., Phillips, M., Brammer, M., Bullmore, E., Williams, S., et al. (2000). Explicit and implicit neural mechanisms for processing of social information from facial expressions: A functional magnetic resonance imaging study. *Human Brain Mapping*, 9, 93–105.

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., et al. (2006). Understanding emotions in others: Mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, 9, 28–30.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85.

Davis, M. H. (1983). Measuring individual-differences in empathy—Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44, 113–126.

Davis, M. H. (1996). *Empathy—A social psychological approach*. Boulder, CO: Westview.

Decety, J., & Chaminade, T. (2003). Neural correlates of feeling sympathy. *Neuropsychologia*, 41, 127–138.

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, 3, 71–100.

Decety, J., & Sommerville, J. A. (2003). Shared representations between self and other: A social cognitive neuroscience view. *Trends in Cognitive Sciences*, 7, 527–533.

Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology*, 39, 39–45.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, 11, 86–89.

Duvernoy, H. M. (1999). *The human brain: Surface, three-dimensional sectional anatomy with MRI, and blood supply* (2nd ed.). New York: Springer.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic

- maps and functional imaging data. *Neuroimage*, *25*, 1325–1335.
- Ekman, P. (1982). *Emotion in the human face* (2nd ed.). New York: Cambridge University Press.
- Ekman, P. (1984). Expression and the nature of emotion. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 319–344). Hillsdale, NJ: Erlbaum.
- Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.
- Farrow, T. F. D., Zheng, Y., Wilkinson, I. D., Spence, S. A., Deakin, J. F. W., Tarriner, N., et al. (2001). Investigating the functional anatomy of empathy and forgiveness. *NeuroReport*, *12*, 2433–2438.
- Fink, G. R., Markowitsch, H. J., Reinkemeier, M., Bruckbauer, T., Kessler, J., & Heiss, W. D. (1996). Cerebral representation of one's own past: Neural networks involved in autobiographical memory. *Journal of Neuroscience*, *16*, 4275–4282.
- Friston, K. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., & Frackowiak, R. (1995). Spatial registration and normalization of images. *Human Brain Mapping*, *2*, 165–189.
- Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., Frackowiak, R. S., et al. (1995). Analysis of fMRI time-series revisited. *Neuroimage*, *2*, 45–53.
- Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *358*, 459–473.
- Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of “theory of mind”. *Trends in Cognitive Sciences*, *7*, 77–83.
- Gallagher, H. L., Happe, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of “theory of mind” in verbal and nonverbal tasks. *Neuropsychologia*, *38*, 11–21.
- Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage*, *16*, 814–821.
- Gallese, V. (2003). The manifold nature of interpersonal relations: The quest for a common mechanism. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *358*, 517–528.
- Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8*, 396–403.
- George, M. S., Ketter, T. A., Gill, D. S., Haxby, J. V., Ungerleider, L. G., Herscovitch, P., et al. (1993). Brain regions involved in recognizing facial emotion or identity: An oxygen-15 PET study. *Journal of Neuropsychiatry and Clinical Neurosciences*, *5*, 384–394.
- Glaser, D. E., & Friston, K. J. (2003). Variance components. In R. S. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan, C. J. Price, S. Zeki, et al. (Eds.), *Human brain function* (2nd ed.). San Diego, CA: Academic Press.
- Goldman, A. I., & Sripada, C. S. (2005). Simulationist models of face-based emotion recognition. *Cognition*, *94*, 193–213.
- Gorno-Tempini, M. L., Pradelli, S., Serafini, M., Pagnoni, G., Baraldi, P., Porro, C., et al. (2001). Explicit and incidental facial expression processing: An fMRI study. *Neuroimage*, *14*, 465–473.
- Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping*, *12*, 1–19.
- Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences, U.S.A.*, *98*, 4259–4264.
- Hess, U., & Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology*, *40*, 129–141.
- Iacoboni, M., Koski, L. M., Brass, M., Bekkering, H., Woods, R. P., Dubeau, M. C., et al. (2001). Reafferent copies of imitated actions in the right superior temporal cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *98*, 13995–13999.
- Iacoboni, M., Lieberman, M. D., Knowlton, B. J., Molnar-Szakacs, I., Moritz, M., Throop, C. J., et al. (2004). Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *Neuroimage*, *21*, 1167–1173.
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, *3*, e79.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, *286*, 2526–2528.
- Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: A critique. *Trends in Cognitive Sciences*, *9*, 21–25.
- Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage*, *14*, S103–S109.
- Kampe, K. K., Frith, C. D., & Frith, U. (2003). “Hey John”: Signals conveying communicative intention toward the self activate brain regions associated with “mentalizing,” regardless of modality. *Journal of Neuroscience*, *23*, 5258–5263.
- Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin*, *100*, 78–100.
- Koski, L., Iacoboni, M., Dubeau, M. C., Woods, R. P., & Mazziotta, J. C. (2003). Modulation of cortical activity during different imitative behaviors. *Journal of Neurophysiology*, *89*, 460–471.
- Kupfer, J., Brosig, B., & Brähler, E. (2001). *Toronto-Alexithymie-Skala-26 (TAS-26) Deutsche Version*. Göttingen: Hogrefe.
- Lawrence, E. J., Shaw, P., Giampietro, V. P., Surguladze, S., Brammer, M. J., & David, A. S. (2006). The role of “shared representations” in social perception and empathy: An fMRI study. *Neuroimage*, *29*, 1173–1184.
- Leslie, K. R., Johnson-Frey, S. H., & Grafton, S. T. (2004). Functional imaging of face and hand imitation—Towards a motor theory of empathy. *Neuroimage*, *21*, 601–607.
- Mehrabian, A. (1997). Relations among personality scales of aggression, violence, and empathy: Validation evidence bearing on the risk of eruptive violence scale. *Aggressive Behavior*, *23*, 433–445.
- Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*, *40*, 525–543.
- Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *17*, 1306–1315.
- Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, *50*, 655–663.
- Nakamura, K., Kawashima, R., Sato, N., Nakamura, A., Sugiura, M., Kato, T., et al. (2000). Functional delineation of the human occipito-temporal areas related to face and scene processing. A PET study. *Brain*, *123*, 1903–1912.

- Ochsner, K. N., Knierim, K., Ludlow, D. H., Hanelin, J., Ramachandran, T., Glover, G., et al. (2004). Reflecting upon feelings: An fMRI study of neural systems supporting the attribution of emotion to self and other. *Journal of Cognitive Neuroscience*, *16*, 1746–1772.
- Oldfield, R. C. (1971). Assessment and analysis of handedness—Edinburgh Inventory. *Neuropsychologia*, *9*, 97–113.
- Oliveri, M., Babiloni, C., Filippi, M. M., Caltagirone, C., Babiloni, F., Cicinelli, P., et al. (2003). Influence of the supplementary motor area on primary motor cortex excitability during movements triggered by neutral or emotionally unpleasant visual cues. *Experimental Brain Research*, *149*, 214–221.
- Ongur, D., Ferry, A. T., & Price, J. L. (2003). Architectonic subdivision of the human orbital and medial prefrontal cortex. *Journal of Comparative Neurology*, *460*, 425–449.
- Piefke, M., Weiss, P. H., Zilles, K., Markowitsch, H. J., & Fink, G. R. (2003). Differential remoteness and emotional tone modulate the neural correlates of autobiographical memory. *Brain*, *126*, 650–668.
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage*, *10*, 15–35.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind. *Behavioral and Brain Sciences*, *1*, 515–526.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, *2*, 661–670.
- Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: A PET investigation of agency. *Nature Neuroscience*, *4*, 546–550.
- Ruby, P., & Decety, J. (2003). What you believe versus what you think they believe: A neuroimaging study of conceptual perspective-taking. *European Journal of Neuroscience*, *17*, 2475–2480.
- Ruby, P., & Decety, J. (2004). How would you feel versus how do you think she would feel? A neuroimaging study of perspective-taking with social emotions. *Journal of Cognitive Neuroscience*, *16*, 988–999.
- 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.
- Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., et al. (2005). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, *44*, 718–730.
- Shah, N. J., Marshall, J. C., Zafiris, O., Schwab, A., Zilles, K., Markowitsch, H. J., et al. (2001). The neural correlates of person familiarity. A functional magnetic resonance imaging study with clinical implications. *Brain*, *124*, 804–815.
- Singer, T., Seymour, B., O’Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, *303*, 1157–1162.
- Sonnby-Borgstrom, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy. *Scandinavian Journal of Psychology*, *43*, 433–443.
- Spence, S. A., Brooks, D. J., Hirsch, S. R., Liddle, P. F., Meehan, J., & Grasby, P. M. (1997). A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control). *Brain*, *120*, 1997–2011.
- Stotland, E. (1969). Exploratory investigations of empathy. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 271–314). New York: Academic Press.
- Sugiura, M., Shah, N. J., Zilles, K., & Fink, G. R. (2005). Cortical representations of personally familiar objects and places: Functional organization of the human posterior cingulate cortex. *Journal of Cognitive Neuroscience*, *17*, 183–198.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme Medical Publishers.
- Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happe, F., Falkai, P., et al. (2001). Mind reading: Neural mechanisms of theory of mind and self-perspective. *Neuroimage*, *14*, 170–181.
- Vogeley, K., & Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends in Cognitive Sciences*, *7*, 38–42.
- Völlm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., et al. (2006). Neuronal correlates of theory of mind and empathy: A functional magnetic resonance imaging study in a nonverbal task. *Neuroimage*, *29*, 90–98.
- Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in My Insula: The common neural basis of seeing and feeling disgust. *Neuron*, *40*, 655–664.
- Winston, J. S., O’Doherty, J., & Dolan, R. J. (2003). Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *Neuroimage*, *20*, 84–97.
- Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *358*, 593–602.