

# Control of Shared Representations Relies on Key Processes Involved in Mental State Attribution

Stephanie Spengler,<sup>1\*</sup> D. Yves von Cramon,<sup>1,2</sup> and Marcel Brass<sup>3</sup>

<sup>1</sup>Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

<sup>2</sup>Max Planck Institute for Neurological Research, Cologne, Germany

<sup>3</sup>Department of Experimental Psychology and Ghent Institute of Functional and Metabolic Imaging, Ghent University, Ghent, Belgium

---

**Abstract:** Action observation leads to the automatic activation of the corresponding motor representation in the observer through “mirror-matching.” This constitutes a “shared representational system,” which is thought to subserve social understanding by motor simulation. However, it is unclear how these shared representations can be controlled and distinguished. Brain imaging suggests that controlling shared representations, indexed by the ability to control automatic imitative responses, activates anterior fronto-medial cortex (aFMC), and temporo-parietal junction (TPJ). Crucially, these regions are also consistently implicated in mental state attribution and have provided an alternative account for higher-level social cognition. Here, we directly tested whether social-cognitive processes involve similar key computational mechanisms as the control of shared representations by using functional brain imaging to reveal overlapping brain circuits. We show in a within-subject design that commonly activated regions occurred selectively in aFMC and TPJ. Mentalizing and self-referential thoughts recruited a region in aFMC, which was also activated when controlling imitation. In the TPJ, an area overlapped between mentalizing, agency processing, and imitative control. Behavioral and neural correlates of mentalizing were further related to the individual ability for controlling imitation. Our findings support the assumption of shared key processes and suggest a novel link between embodied and social cognition. *Hum Brain Mapp* 30:3704–3718, 2009. © 2009 Wiley-Liss, Inc.

**Key words:** agency; anterior fronto-medial cortex; fMRI; mentalizing; mirror system; social cognition; temporo-parietal junction

---

## INTRODUCTION

Recent research has focused on the “mirror system,” a set of cortical regions, which are commonly activated during both the execution and observation of actions, supported by the findings from primate neurophysiology [di Pellegrino et al., 1992; Rizzolatti and Craighero, 2004] and human neuroimaging [e.g. Gangitano et al., 2004]. Such “mirroring” reactions have been found in a broad range of domains and comprise responses on a behavioral and/or neurophysiological level [for overview, see Heyes and Bird, 2007; Brass and Spengler, 2008]. For example, observation of hand movements can lead to an automatic facilitation of this action [Brass et al., 2001a; Kilner et al., 2003], elicit motor-evoked potentials from the specific muscles responsible for this movement [Fadiga et al., 1995; Strafella and

---

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: NEST/EDICI (European Community's Sixth Framework Programme); Contract grant number: 012929.

\*Correspondence to: Stephanie Spengler, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr.1a, 04103 Leipzig, Germany. E-mail: spengler@cbs.mpg.de

Received for publication 25 February 2009; Accepted 17 March 2009

DOI: 10.1002/hbm.20800

Published online 10 June 2009 in Wiley InterScience (www.interscience.wiley.com).

Paus, 2001], and activate regions in the rostral ventral premotor and parietal cortex in the observer, which are also active during the actual execution of this action [Buccino et al., 2001]. These findings thus predict a tendency for “automatic,” unintentional imitation in the observer. This assumption is corroborated by findings of unintentional imitation in social interaction [“chameleon effect,” Chartrand and Bargh, 1999] and a heightened display of imitative behavior in patients with prefrontal lesions [Brass et al., 2003; Lhermitte, 1986]. Because of these “shared representations” for perception and action, a simulationist account has proposed that the mirror system is involved in action understanding, social cognition, and imitation by motor simulation [Gallese and Goldman, 1998].

Although action observation may usually lead to motor contagion, an automatic tendency to imitate the observed action, overt imitation is not the default behavior in healthy adults. In most everyday situations, nonimitative responses are preferred over imitative responses. However, clearly crucial for purposeful behavior and necessitated by the action-perception overlap, the question how we are able to exert control over automatic imitative response tendencies has so far received less interest. A recent model has suggested that the nonoverlapping parts of the perception-action system might support mechanisms that enable a distinction between self- or other-related actions [“who” system, Jeannerod, 1999, 2003], coding explicitly by whom, oneself or another agent, the corresponding motor representation was evoked. This system would allow to discriminate and inhibit the other-evoked motor program. Furthermore, it would enable us to maintain and reinforce intentionally generated motor representations. Discrimination of self or other agency thus seems to be a crucial aspect to control shared representations [Decety and Sommerville, 2003]. In turn, the underlying elementary functions of this system have been proposed to be key contributors of more complex social-cognitive abilities, such as theory of mind (ToM) and empathy [Decety and Grezes, 2006; Decety and Lamm, 2006].

Neuroimaging studies have consistently revealed activated brain regions in the anterior fronto-medial cortex (aFMC) and the temporo-parietal junction (TPJ) subserving higher-level social-cognitive processing [Amodio and Frith, 2006; Decety and Grezes, 2006; Decety and Lamm, 2007]. These brain regions are thought to support computational mechanisms underlying self-other discrimination and mental state attribution and have thus provided an alternative account for social understanding through inferential and reflective processes [Frith and Frith, 2003]. The aFMC and the TPJ are essential to distinguish self- and other-generated actions and motor representations [Decety et al., 2002; Farrer et al., 2003; Farrer and Frith, 2002; Leube et al., 2003], to shift one’s cognitive [Ruby and Decety, 2001, 2003] and visual perspective [Aichhorn et al., 2005], and to support mentalizing in ToM tasks [Gallagher et al., 2000]. The median wall has also consistently been associated with tasks requiring self-referential processing [Northoff

and Bermpohl, 2004]. The activation of corresponding motor representations may be unintentional and unaware, whereas consciously reflecting on these shared representations by becoming aware of own intentions and forming explicit meta-representations may be needed for action monitoring and self-regulation [Decety and Grezes, 2006].

However, while there is extensive research on the neural circuits involved in mental state attribution and also on mirror system functions, little is known about the neural correlates and cognitive processes involved in the *control and distinction* of shared representations of the mirror system. Two fMRI studies investigated the neural correlates of controlling automatic imitative response tendencies, which can be used as a prime example or an index of controlling shared representations. Suppressing automatic imitative response tendencies yielded activation in the aFMC [Brass et al., 2001b; Brass et al., 2005a] and the TPJ [Brass et al. 2005a]. Comparing the activations of these studies on imitative control with the loci of activations reported in recent meta-analyses on mentalizing, self-reflection, and agency, persuasive overlap of activations can be observed in aFMC and TPJ [see e.g. Decety and Grezes, 2006; Gilbert et al., 2006]. This overlap of activations between studies may suggest that shared basic abilities exist, such as the ability to attend to intentions of oneself or others and for self-other differentiation. Hence, common computational processes might subserve both the control of shared representations and also later-developing social-cognitive capacities [Brass et al., 2005a; Brass and Spengler, 2008]. This would provide a novel link between two hitherto opposing theoretical views on social cognition [e.g. Keysers and Gazzola, 2007], by assuming that the control of mirror system functions (“simulation account”) relies on key computations of mentalizing and self-other distinction (“ToM account”). However, evidence for this claim was so far primarily based on descriptive, between-experiment, and between-subject comparisons of activated regions, making interpretations more difficult due to existing differences, for example, in imaging and analysis procedures between laboratories and differences in neuroanatomy between subject groups. Therefore, it would be more conclusive to observe this overlap of common, activated regions within one study and not between studies.

To overcome these shortcomings of previous studies, we conducted an fMRI study with a within-subject, within-session design. We predicted that, if successfully producing nonimitative responses recruits the same functional mechanisms as involved in social cognition, then occurring activity in the crucial brain areas (aFMC/TPJ) should overlap between both sets of tasks. Activity of the same brain regions under two conditions or tasks implies a common function, if one assumes that there is a broad specialization of regions irrespective of the task context [Cabeza and Nyberg, 2002; Henson, 2005]. In the present study, all participants performed a task in which they had to produce nonimitative responses, while viewing imitation-inducing movements [Brass et al., 2005a, Supporting Information

Fig. 1] and three social-cognitive tasks. These included key abilities, which have been suggested to be significantly involved in social cognition [Amodio and Frith, 2006; Decety and Grezes, 2006]: attending to mental states of others (ToM) or oneself ("self-referential task") and the ability to perceive and assign those mental states or actions to their respective source or agent ("agency task"). Specifically, it was predicted that brain activation for controlling automatic imitation should overlap in aFMC with self-referential thoughts and mentalizing, but in the TPJ area with activations during agency processing and mentalizing [e.g. Leube et al., 2003; Vogeley et al., 2001; Zysset et al., 2002]. According to our hypothesis, overlap of activations was tested by means of a conjunction analysis and a region of interest (ROI) procedure described by Mitchell [2008] and Saxe and colleagues [e.g. Saxe and Powell, 2006]. Additionally, we reasoned that shared functions may not only be identified by overlap of activated regions, but also possibly by relating dependent measures of one function with the dependent variables of a second function. This is a novel approach, which has, to our knowledge, so far not been used in the neuroimaging literature. If shared computational processes exist, it should be expected that the variables (behavioral indicators and/or hemodynamic parameters) show "cross-task" relations (i.e. associations between tasks). Specifically, it was hypothesized that behavioral indicators (amount of interference in reaction times/errors) of the imitation-inhibition task should correlate with the neuroimaging data of the other tasks in the in the two key regions (TPJ and aFMC).

## METHODS

### Participants

Eighteen volunteers (9 female, mean age 25 years, range 22–31) participated in the two scanning sessions. All participants were right-handed, had normal or corrected-to-normal vision, and were native German speakers. No participant had a history of neurological, major medical, or psychiatric disorder. Participants completed all fMRI experiments, but reaction time data from one subject was lost in the imitation-inhibition task due to a computer failure. The data of this subject was used in the analyses without RT data ( $n = 18$ ) but was omitted in further analyses ( $n = 17$ ). Experiment was undertaken with the understanding and written consent of each subject, with the approval of the appropriate local ethics committee, and in compliance with national legislation and the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki).

### Procedure

Participants performed two of the four tasks in one scanning session, and the remaining two tasks in a scanning session on the following day (each session lasting

~ 45 min). Within the individual sessions, task order was fully counterbalanced across subjects. Before scanning participants were informed, the task that they had to perform on that day and training (in the order of the tasks in the scanner) was given. On the day before, the first scanning session participants took part in a training session for the agency task.

## Tasks

### Imitation-inhibition task

This task used a modified version of the imitation-inhibition paradigm used by Brass et al. [2005a]. In the imitation-inhibition task [Brass et al., 2000], participants had to lift their index or middle finger in response to a number (1 = lift index finger, 2 = lift middle finger), whilst viewing simultaneously congruent or incongruent finger movements on a computer screen (Supporting Information Fig. 1). As in congruent trials, the video-taped hand executed an identical finger movement to the instructed movement; the participants were not required to distinguish their intended action from the observed action. Their response was therefore a *quasi-imitative reaction* with regard to the observed movement, which is normally indicated by faster reaction times on congruent trials [Brass et al., 2000]. Contrary, in incongruent trials, the instructed movement differed from the observed movement, which introduces an imitative tendency to execute the observed movement. This necessitates subjects to enforce the intended movement against the observed action, reflected in longer reaction times on incongruent trials [Brass et al., 2000]. The required response on incongruent trials reflects therefore a *nonimitative reaction* with regard to the perceived movement. The experiment consisted of 72 incongruent trials, 72 congruent trials, and 36 null events, which were presented randomly. Each trial (duration 6 s) started with a variable jitter interval (0, 500, 1,000, and 1,500 ms). The video sequence began with a frame showing the hand, which mirrored the right hand of the subject, in resting position for 650 ms. Two frames lasting 34 ms then showed the finger movement (congruent and incongruent trials). The video sequence finished with the finger in the end position (650 ms). In half of the trials for each condition, the imperative stimulus was presented 150 ms before the onset of the finger movement (soa 1), and in the other half with the onset of the movement (soa 2). The screen remained blank between trials. In null events, a blank screen was shown for 6 s. Responses were recorded with custom-built keypad, which used light sensors to detect finger lifting movements of the index and middle finger of the subjects' right hand. Total task length was 18 min.

### ToM task

This task, testing mentalizing or "ToM," has been widely used in the literature; here, we used the German

translation of the stimuli from Vogeley et al. [2001]. Three different conditions (experimental condition: ToM, control conditions: physical stories and unlinked sentences), comprising of each eight stories, were presented in randomized order. The stories required either the comprehension or attribution of a mental state to the actor in the story (ToM condition) or comprehension of a physical state [Happe et al., 1999] to understand the story. To ensure that participants stayed alert and performed the task, a forced choice response was required after reading the question relating to the story. Each trial began with displaying a cue relating to the condition (800 ms), the story (19.2 s), followed by the question (8 s), and three multiple-choice responses (7.5 s), which could be selected by a right-hand button press (1 = index finger, 2 = middle finger, and 3 = ring finger). After selecting an answer, the responses disappeared and a blank screen was shown for the remainder of the trial. Participants were informed about the different story types (condition) before the presentation of the stimuli. As we adapted the stories from Vogeley et al. [2001], they had already been checked for word count and syntactical and semantic complexity. No significant differences occurred also for the word count of the responses between conditions. A pilot study further confirmed that the time to read the stories and to select an answer was sufficient. Eight null events (blank screen for 32 s) were included in the design, occurring after the presentation of three stories. Total task length was 20.2 min.

### Self-referential task

To tap processes related to self-knowledge (i.e. the ability to differentiate self from other objects and recognize attributes and preferences of oneself [Amodio and Frith, 2006]), we used a replication of an evaluative task developed by Zysset et al. [2002]. Evaluative judgments are a special type of judgment, in which the internal scale is related to the person's value system (preferences, norms, aesthetic values, etc.). In contrast to memory retrieval in general, many evaluative judgments are self-referential in that they refer to the subjects "narrative" self [Gallagher, 2000]. In the evaluative conditions, participants had to read sentences and to decide whether they agreed with this statement (e.g. "I like Leipzig"). To isolate processes specific for self-related, evaluative judgments, we used semantic memory retrieval trials (e.g. "Leipzig is the capital of Germany") as a control condition. For each experimental condition, 60 sentences were randomly presented and 60 trials of null events (blank screen). Sentences had been carefully controlled for an equal distribution of yes- and no-responses and for difficulty in the memory condition [for further information on the stimulus material, see Zysset et al., 2002]. In each trial (6 s), after a variable jitter interval (0, 400, 800, 1,200, or 1,600 ms), the sentence was presented for 2 s, independent of whether a response had been given. Subjects were given 4 s after stimulus onset to

respond by pressing a response button with the right index finger (yes) or right middle finger (no). The screen remained blank between trials. Total task length was 18 min.

### Agency task

The task consisted of an acquisition phase and an experimental phase on 2 consecutive days, similar to the study of Sato and Yasuda [2005]. In the acquisition phase, when a white outline of a square appeared on the screen (responses were accepted up to up 1000 ms), participants could freely choose if they performed an index or middle-finger button press of the right hand, but they were instructed to react as quickly as possible. After each button press, a colored square (either red or blue) immediately appeared in the white outline (for 200 ms). The mapping of the two colors to the index- or middle-finger button press was counterbalanced across subjects. Between trials, a blank screen was presented for 1,000 ms. Participants performed three blocks of 300 learning trials each. On the following day before the fMRI session took place, subjects participated in one training block of 300 trials. Afterward, they were instructed about the experiment in the scanner: participant were told they had to react as quickly as possible with an index- or middle-finger button press to the white square outline (as in the learning phase), but that this time the colored squares could be either produced by themselves (as in the learning phase) or by the experimenter. In reality, the experimenter did not produce any action-effects, but all responses were generated by the computer. The appearance of the colored squares (i.e. the action-effect) was either varied by a temporal delay after button press (delay of 0, 200, 300, or 600ms) or by the congruency with respect to the learned action-effect association, so that under the congruent condition the same color that had followed the respective button press in the learning phase appeared or in the incongruent condition the opposite color appeared. After each square, participants had to judge on a four-digit rating scale with a button press of the left hand, if they produced the action-effect or the experimenter ("I produced the square", 1 = totally disagree, 4 = totally agree). To familiarize subjects with the task and to make it more plausible that the experimenter participated in the experiment, a longer practice session was given. Subjects first could practice parts of the task and the rating alone, but then they also practiced a version of the fMRI task, in which they were told that the experimenter could produce action-effects, while they were sitting at a second computer in the same room behind a dividing wall. Afterward, participants performed a total of 160 experimental trials of this task in the scanner, 32 trials for each of the temporal delays (with a congruent mapping), 32 trials with an incongruent mapping (equally distributed over the four temporal delays) and 40 null events (blank screen). All incongruent trials were analyzed as a whole, so that, in total, five conditions were entered



into the analysis. Each trial lasted in total 6 s, started with a variable jitter interval (0, 250, 500, and 750 ms), the display of the white outline (reactions were allowed up to 1 s afterward), the colored square (200 ms), 1 s blank screen, the rating scale for up to 2,500 ms, and the remainder of the trials consisted of a blank screen. Total task length was 20 min. This paradigm and procedure has been reported in more detail in Spengler et al. [2009], and further findings on this paradigm are reported there.

### fMRI Data Acquisition

Subjects lay supine on the scanner bed, with two fingers of the right and four fingers of the left hand (for the agency task) positioned over the buttons of a response box. Participants' hands were carefully stabilized, and form-fitting cushions were used to prevent arm, hand, and head motion. To attenuate scanner noise, participants were provided with earplugs.

Imaging was performed using 3T scanner (Bruker Medspec 30/100), equipped with the standard birdcage head coil. Twenty-four axial slices (19.2 cm FOV;  $64 \times 64$  pixel matrix; thickness 4 mm; spacing 1 mm) parallel to bicommissural line (AC-PC) were acquired using a single-shot gradient echo-planar imaging (EPI) sequence (TR = 2,000 ms, TE = 30 ms, flip angle =  $90^\circ$ ) sensitive to blood oxygenation level-dependent contrast. Prior to the functional runs, a set of 20 anatomical 2D MDEFT and EPI-T1 images were acquired for each subject. In a separate session, high-resolution 3D whole-brain images (160 slices and 1-mm slice thickness) were acquired from each subject to improve the localization of activation foci using a T1-weighted three-dimensional (3D) segmented MDEFT sequence covering the whole brain.

### fMRI Data Analysis

The fMRI data were analyzed using the software package LIPSIA [Lohmann et al., 2001]. First, the functional data was motion-corrected with a 3D motion correction using six degrees of freedom (three translational and three rotational). To correct for the temporal offset between the slices acquired in one scan, a cubic spline interpolation algorithm was applied. The functional slices were then coregistered with the high-resolution full-brain scan by a rigid, affine linear transformation using three translational and three rotational parameters. Subsequently, these rotational and translational parameters were transformed to standard size [Talairach and Tournoux, 1988] by linear scaling. The transformation parameters obtained from this step were then applied to the functional data, so that the functional slices were also registered into the stereotactic space by means of a trilinear interpolation. Voxel size was interpolated during the coregistration to a spatial resolution of  $3 \times 3 \times 3$  mm. Depending on the design of the task, a suitable high-pass filter was applied (cutoff

frequency: imitation-inhibition: 1/65 Hz, ToM task: 1/125 Hz, self-referential task: 1/72 Hz, agency task: 1/65 Hz) and the data was filtered with a spatial Gaussian filter with 7.2-mm full-width half maximum (FWHM). The statistical evaluation was based on a least-squares estimation using the general linear model for serially autocorrelated observations [random effects model; Friston et al., 1995; Worsley and Friston, 1995]. For all tasks, except for the ToM task, an event-related design was used; that is, the hemodynamic response function was modeled with respect to the experimental conditions (for each stimulus, event = onset of stimulus presentation). The design matrix was generated using a synthetic hemodynamic response function [gamma function, Glover, 1999] and its first derivative [Friston et al., 1998] and a response delay of 6 s. For the ToM task, the design matrix was generated using a box-car function. The model equation, including the observation data, the design matrix, and the error term, was convolved with a Gaussian kernel of dispersion of 4 s FWHM [Worsley and Friston, 1995]. Afterward, contrast images, that is, estimates of the raw score differences between specified conditions, were generated for each subject. The single-subject contrast images from the first level were entered into a second-level random effects analysis, using a one-sample *t*-test across the contrast images of all subjects to indicate whether observed differences between conditions were significantly different from zero. Subsequently, *t* values were transformed into Z scores. The relevant contrasts for the individual tasks included incongruent > congruent condition (imitation-inhibition task), ToM > physical stories (ToM task) and evaluative > semantic memory trials (self-referential task). For the agency task, we subjected a parametric regressor containing all five conditions and a parametric regressor coding the judgments of the subjects to a second-level analysis. In case of the imitation-inhibition task, we additionally correlated the individual, behavioral data for each participant (interference score reaction times, incongruent-congruent trials, *z*-transformed) with the contrast images from the incongruent > congruent contrast at the second level. This allowed us to further explore whether successfully executing nonimitative responses (indicated by a smaller interference score) was related to higher activity in regions involved in these processes (e.g. aMFC).

To protect against false positive activations, we used a double-threshold approach, that is, combining a voxel-based threshold with a minimum cluster size [Forman et al., 1995]. This nonarbitrary voxel cluster size was determined by using the program AlphaSim (afni.nimh.nih.gov/afni/doc/manual/AlphaSim; [Ward, 2000]). The multiple comparison correction used a combination of single-voxel probability thresholding on one hand, and cluster-size and cluster-*z*-value thresholding on the other hand. The probability threshold was set to  $Z > 2.56$  ( $P < 0.005$ , uncorrected), the cluster-size and cluster-*z*-value threshold were computed by the program using Monte-Carlo simulations. The program contains thresholds that

specify features that a cluster must have in order to qualify as being significant: the first feature is cluster size (i.e. number of voxels in a cluster). The second feature is the maximum  $z$ -value in a cluster. Thus, a cluster may qualify as being significant if it is either large enough, or if it has a large maximum  $z$ -value, or both. On the basis of a Monte-Carlo simulation (1,000 iterations), with our brain volume and an individual voxel height threshold of  $Z > 2.56$  ( $P < 0.005$ , uncorrected), areas were determined which were significant at an overall imagewise false-positive rate of 5% or 1%. Activations exceeding this cluster based threshold are therefore considered to be activated at an experiment-wise threshold of  $P < 0.05$ , corrected for multiple comparisons. Therefore, only activations exceeding this threshold are reported.

For the ROI approach, we used procedure specified by Mitchell [2008]. On the group level, activated voxels (at increasingly more conservative thresholds) in the imitation-inhibition task were used to extract brain activation values for the other tasks. Furthermore, individually defined rTPJ ROIs were defined separately for each participant from the activation in the imitation-inhibition task. To determine the individual rTPJ region an iterative procedure was used, so that first the main contrast incongruent > congruent trials (imitation-inhibition task) was first examined at the lowest statistical threshold ( $P < 0.05$ , uncorrected). Subsequently, this comparison was then visualized iteratively at progressively more conservative statistical thresholds (i.e., 0.05, 0.025,  $10^{-1}$ ,  $10^{-2}$ ,  $10^{-3}$ ,  $50^{-3}$ ,  $10^{-4}$ , and  $50^{-4}$ ) until the maximally circumscribed rTPJ ROI was determined. Furthermore, this region had to be anatomically localized in the TPJ region (ventral part of the inferior parietal lobe and/or the caudal parts of the posterior superior temporal gyrus) and had to include at least three contiguous voxels.

To identify regions commonly activated in the inhibition of imitative responses and social-cognitive measures, a conjunction analysis was performed. This method aims to finding commonly activated voxels, e.g., in two tasks or contrasts, by using the logical combination method [Joseph et al., 2002; Nichols et al., 2005]. The group  $Z$ -maps of the selected contrasts were thresholded at  $Z > 3.09$  ( $P < 0.001$ , uncorrected). These thresholded  $Z$ -maps were then overlaid, showing only voxels activated in both maps (at  $P < 10^{-6}$  for two tasks, at  $P < 10^{-9}$  for three tasks), which is the most conservative test for a conjunction [Nichols et al., 2005]. To identify peak activations and cluster sizes, the group  $Z$ -maps were averaged and masked with the overlap map (from the conjunction analysis), subsequently being referred to as masked-average  $Z$ -map.

Last, regions were explored which yielded a negative correlation with aFMC activation in the imitation-inhibition task. First, for each subject, parameter estimates in the aFMC for the correlation between behavioral performance and the contrast incongruent > congruent were extracted. These values were then used in a second-level, simple regression on a whole brain level to identify regions that

show negative correlations with these values. This analysis was performed on the unsmoothed contrast images from the comparison incongruent > congruent condition. In other words, we aimed to identify regions that show a decrease in activation with an increase in activation in the aFMC region.

## RESULTS

### Behavioral Results

All tasks showed behavioral effects which were highly similar to those reported by earlier studies [Brass et al., 2005; Sato and Yasuda, 2005; Vogeley et al., 2001; Zysset et al., 2002].

In the *imitation-inhibition task*, participants showed an interference effect, as reaction times on incongruent trials were significantly slower than reaction times on congruent trials [ $t(16) = 8.4$ ,  $P < 0.0001$ ]. Concerning the errors, subjects made more errors in the incongruent than in the congruent condition [ $t(16) = -4.1$ ,  $P < 0.001$ ].

For the *ToM task*, participants displayed similar performance across all three conditions [ToM stories: 7.39 (SD 0.7), physical stories: 6.94 (SD 1.21), unlinked sentences: 6.72 (1.45)], and no significant differences could be found between conditions (repeated measurement ANOVA:  $F(2,34) = 1.58$ ,  $P > 0.2$ ).

In the self-referential task, we analyzed the amount of yes and no decisions, as it was not possible to differentiate between correct and false responses for the evaluative condition due to the subjective nature of this task. In the semantic and the evaluative, 50.9% and 46.7% were yes responses [ $t(17) = -1.86$ ,  $P > 0.05$ ], showing that the conditions were well-matched and no response tendencies occurred. The two task conditions showed no significant differences [ $t(17) = 1.39$ ,  $P > 0.1$ ] in reaction times [evaluative condition: 2,464 ms (SD 1,215), semantic condition: 2,062 ms (432)].

For the *agency task*, in accordance with Sato and Yasuda [2005], a main effect of condition [repeated measurement ANOVA:  $F(4,68) = 94.6$ ,  $P = 0.000$ ] could be found for the rating data of the subjects. This indicated that participants showed a decreased sense of agency when discrepancies between predicted and actual action-effects increased. This reduction of the experience of agency was either due to longer delays between action and effect or was even more pronounced when the action-effect was incongruent with the preceding action. Reaction times for the ratings showed no significant differences between conditions [ $F(4,68) = 1.4$ ,  $P > 0.2$ ].

### Neuroimaging Results

#### **Conjunction analysis: Common regions activated in imitation-inhibition task and the social-cognitive tasks**

Our primary aim was to identify regions commonly activated by the paradigms tapping social-cognitive processes

and the imitation–inhibition task. Therefore, the Z-maps of the key comparisons (listed in the Methods section, see also Supporting Information Material and Supporting Information Table I) were subjected to a *conjunction analysis*. First, we tested for combinations of pairs of tasks (e.g. overlap imitation–inhibition task and ToM task) and second, for regions activated in three or more tasks (e.g., overlap imitation–inhibition task, ToM task, and agency task). As the main aim of the analyses was to focus on the overlap between imitative control and social-cognitive sub-processes, we only report results of conjunctions including the imitation–inhibition task. Regions, which showed overlap between a pair of tasks ( $P < 10^{-6}$ ), are listed in Supporting Information Table II.

As predicted by our hypotheses, significant overlap of activations could only be found very selectively in the two key regions, aFMC and TPJ. We identified a significant overlap in the right TPJ between the imitation–inhibition task and the ToM task, as well as between the imitation–inhibition task and the agency task. No commonly activated region in the TPJ could be found for the self-referential task and the imitation–inhibition task. The results of the conjunction moreover revealed that the region of the aFMC, activated in the imitation–inhibition task (correlation of aFMC activation with the behavioral performance, Fig. 2a), fully overlapped with involved regions in the ToM and the self-referential task. To test further, which brain regions showed an overlap for combinations of three tasks, we performed an additional conjunction analysis, including only the tasks, which already yielded an overlap in the analysis of the task pairs ( $P < 10^{-9}$ , Fig. 1). For the conjunction of the imitation–inhibition task, ToM, and agency task, we found again a region in the right TPJ, which showed a reliable overlap between all three tasks ( $x: 49, y: -46, z: 21, Z\text{-max}: 3.75$ , nine voxel). The conjunction imitation–inhibition task, ToM, and self-referential task again showed that the aFMC activation of the imitation–inhibition task overlapped completely with the ToM task and the self-referential task ( $x: -5, y: 50, z: 12, Z\text{-max}: 4.60$ , 27 voxels).

Additionally, to protect against the possibility that smoothing of the data during preprocessing resulted in imprecise spatial localization, the conjunction analysis was repeated with unsmoothed data ( $P < 50^{-9}$  for conjunction of three tasks). Replicating our above-reported results, we found an overlap for the three tasks in the aFMC ( $x: -5, y: 50, z: 12, Z\text{-max}: 4.57$ , 39 voxels) and the TPJ ( $x: 46, y: -43, z: 27, Z\text{-max}: 3.33$ , three voxels).

### ROI analysis: Common regions activated in imitation–inhibition task and the social-cognitive tasks

Furthermore, we conducted a second analysis to confirm the findings from the conjunction. This approach followed a ROI analysis recently used by several researchers [Mitch-

ell, 2008; Saxe and Powell, 2006] to reveal if the regions activated in the imitation–inhibition task subserve also the social-cognitive tasks. Therefore, from the activated regions in the imitation–inhibition task, the aFMC (27 voxels) and the TPJ (212 voxels) ( $P < 0.001$ ), we extracted the parameter estimates for the other two tasks and compared the responses for the main conditions. In the TPJ region, the ToM condition showed greater activation than the physical condition [ $t(17) = 1.9, P < 0.04$ ] and the agency task also yielded a significant parametric increase of activations over the conditions [ $t(17) = 2.1, P < 0.03$ ]. As predicted, in the aFMC, we also found significant activations for the ToM task [Tom > physical stories,  $t(17) = 6.0, P < 0.0001$ ] and the self-referential task [evaluative > semantic memory condition,  $t(17) = 8.14, P < 0.0001$ ].

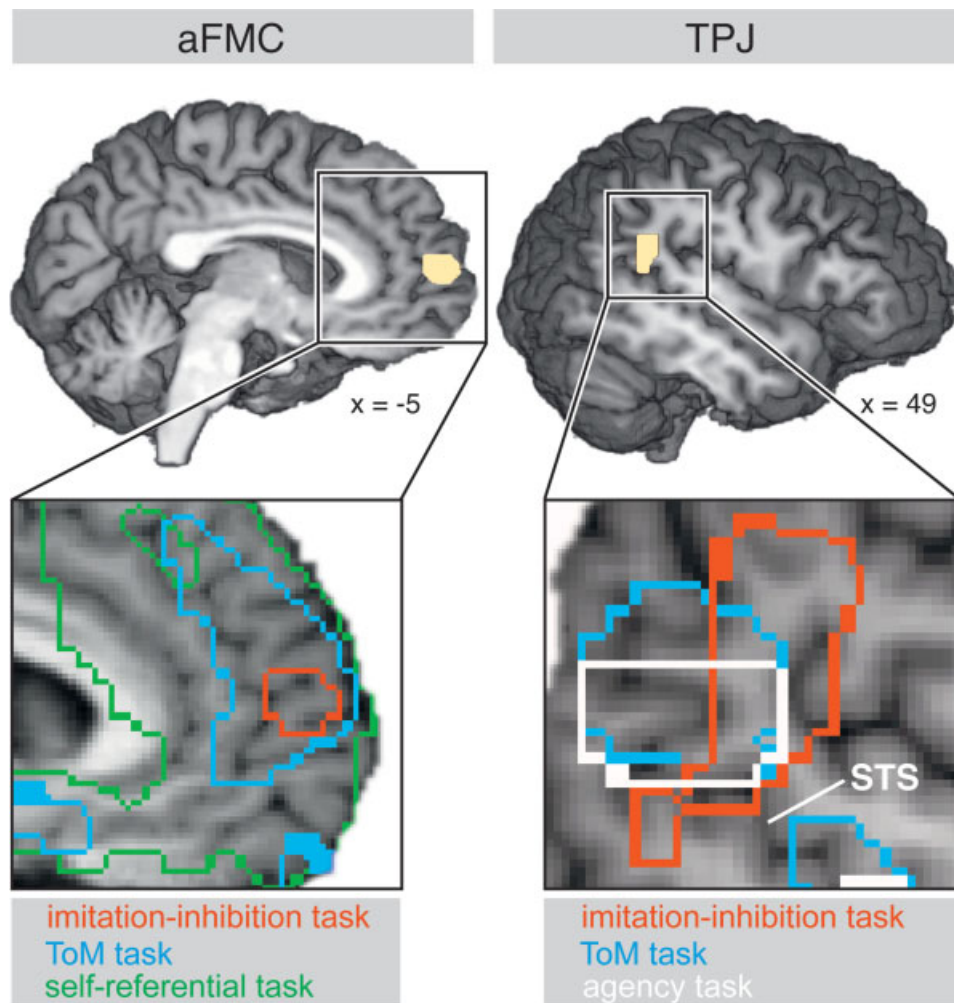
In addition, we used a more conservative defined ROI in the imitation–inhibition task ( $P < 0.0001$ ) to exclude the possibility that the selected region encompassed other adjacent functional areas. At this threshold, the difference in the ToM task [ $t(17) = 2.2, P < 0.02$ ] and the agency task [ $t(17) = 2.1, P < 0.03$ ] in the TPJ (57 voxels) was significant. Similarly, in the aFMC (eight voxels), activity was greater for ToM than physical stories [ $t(17) = 5.6, P < 0.0001$ ] and evaluative than semantic memory trials [ $t(17) = 7.4, P < 0.0001$ ].

Furthermore, we reanalyzed these ROIs (at  $P < 0.001$ ) with unsmoothed data to account for the loss in spatial localization during the smoothing procedure. Again, we found that the TPJ (50 voxels) yielded significant activations in the ToM task [ $t(17) = 2.0, P < 0.03$ ] and the agency paradigm [ $t(17) = 2.2, P < 0.02$ ]. In the aFMC (14 voxels), the contrasts ToM > physical stories [ $t(17) = 6.6, P < 0.0001$ ] and evaluative > semantic memory condition [ $t(17) = 8.8, P < 0.0001$ ] were still very significant.

However, it might be argued that group-based analyses produce blurred activations and could mix the specific activations of one functional region with that of an adjacent area [e.g. Saxe and Powell, 2006]. We therefore investigated individual ROIs of each participant using an iterative procedure (see fMRI analysis) described by Mitchell [2008]. For each participant, we tried to identify a region in the rTPJ, being more active for incongruent than congruent trials in the imitation–inhibition task (Supporting Information Table III). In line with our results from the group analysis, these individually defined rTPJ regions also showed significant activations in the ToM task [ $t(11) = 2.7, P < 0.02$ ] and the agency task [ $t(11) = 2.8, P < 0.02$ ].

### Relationship between tasks: Brain-behavior associations and interindividual differences

To gain further support for our findings from the conjunction analysis on the group level, we sought to undermine those results by adding an *interindividual differences perspective*. As the behavioral performance in the



**Figure 1.**

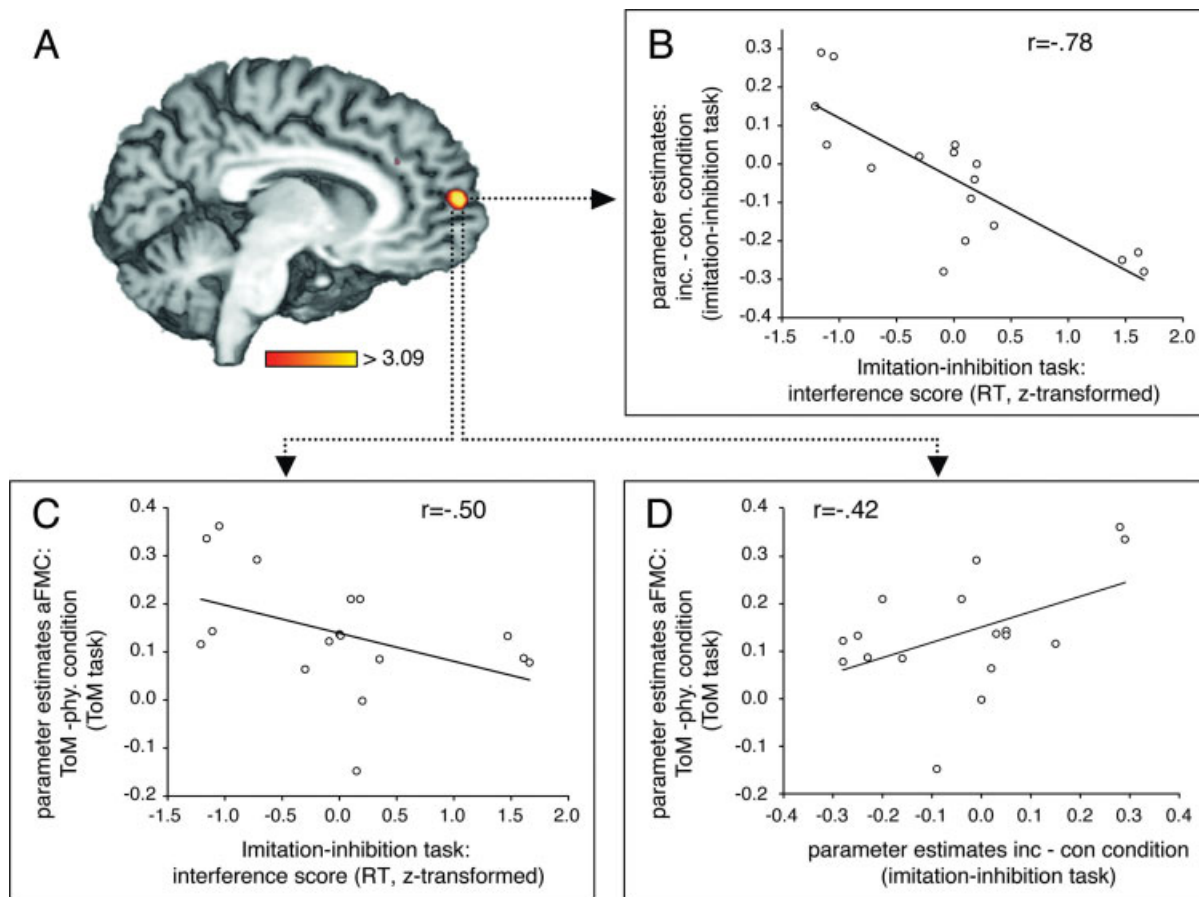
Overlap of actions of the individual tasks revealed by a conjunction analysis. For each individual task Z maps were thresholded at 3.09 (corresponding to  $P < 0.001$ ) and then subjected to a conjunction analysis (resulting in an overall threshold of  $P < 10^{-9}$  for the conjunction), identifying only regions that survived this threshold in all comparisons.

imitation-inhibition task (i.e., the individual ability to successfully execute nonimitative responses) was strongly associated with brain activity in aFMC (Fig. 2a), it might be a reliable indicator of the individual differences in recruiting neural regions and associated cognitive functions. If this would be the case, then the differential ability for nonimitative responses should predict activity in aFMC in other tasks activating this region, if it is assumed (a) that a basic cognitive process, implemented in this region, is involved in these tasks, and (b) that those tasks produce enough interindividual variability. Our analysis only targeted the aFMC region, as this region showed differences in activity related to performance. The interference score of the imitation-inhibition task was therefore correlated with aFMC activation in the ToM and self-refer-

ential task. For the ToM task only, we found that activity of the aFMC was predicted by the interference score ( $r = -0.50$ ,  $P < 0.05$ ), with subjects effectively performing nonimitative actions also showed more activity in aFMC in the ToM task and vice versa (Fig. 2b). Furthermore, individual aFMC activity in the imitation-inhibition task, extracted from the aFMC region, was also positively correlated with activity in the ToM task in this region ( $r = 0.42$ ,  $P < 0.05$ , one-tailed, Fig. 2c).

We further applied this strategy to the behavioral data of the ToM task, by median-splitting our subjects again in a “high interference” and “low-interference” group and analyzing the errors in the ToM task between those groups. One subject showing exactly the median value was omitted, leaving evenly distributed groups ( $n = 8$  for





**Figure 2.**

Brain-behavior relationships in the imitation-inhibition task and the theory of mind task. **A:** Correlation of the interference score (reaction times, z-transformed) with the activity in the imitation-inhibition task (incongruent > congruent condition) in aFMC ( $x: -8, y: 56, z: 12$ ). **B:** Correlation of the interference score (reaction times, z-transformed) with the activity in the

theory of mind task (ToM > physical stories) in aFMC ( $x: -8, y: 56, z: 12$ ). **C:** Correlation of activity in the imitation-inhibition task (incongruent > congruent condition) and activity in the theory of mind task (ToM > physical stories) in aFMC ( $x: -8, y: 56, z: 12$ ). [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

each group). Mirroring the brain imaging results, the “low interference” group showed significantly less errors in the ToM task than the “high interference” group [ $t(14) = 2.6, P < 0.02$ , Fig. 3].

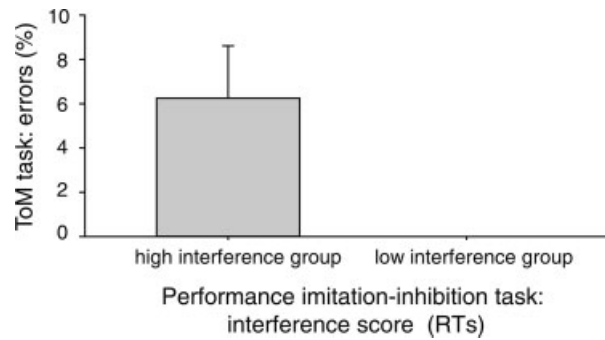
### Relationship between aFMC and “Mirror System” regions

In an additional analysis, we aimed to identify regions showing a negative relationship with aFMC activations, as this region was associated with the behavioral performance of the subjects in the imitation-inhibition task. The rationale behind this idea was that if aFMC activation is a reliable indicator of the individual ability to control shared representations of the mirror system, then activation in this region possibly should also be related to activations in the mirror system. Correlation of the

parameter estimates in aFMC revealed only one region, which was negatively associated with aFMC, in ventral premotor cortex (vPMC) ( $x: -50, y: 0, z: 24, P < 0.001$ , uncorrected).

## DISCUSSION

The aim of the present study was to test whether key social-cognitive processes, such as mentalizing, involve brain regions that overlap with cerebral circuits involved the control of shared representations, assessed by imitative control. In particular, we predicted that the inhibition of imitative response tendencies leads to activation in the TPJ that overlaps with agency processing and ToM, reflecting the need to distinguish self- and other-related actions or perspectives. In the aFMC, it was expected that controlling imitative responses should show common activations with



**Figure 3.**

Behavioral performance in the theory of mind task (% errors) according to performance in the imitation–inhibition task (median split according to interference score RTs).

mental state attribution and self-referential processing, indicating the formation of intentional states. The present within-subject fMRI study confirmed these hypotheses derived from between-study comparisons. By overlaying activations from the individual tasks, we identified common brain activation in the TPJ and aFMC for social-cognitive key abilities and the control of the shared representational system. Furthermore, we found that the individual ability for imitative control was also functionally related to neural and behavioral correlates of mentalizing. Our data strongly support the hypothesis that not the shared representational system per se shows a functional overlap with social-cognitive abilities, in particular, mentalizing processes, but rather the control of shared representations.

### Overlap of activations in aFMC and TPJ

According to Cabeza and Nyberg [2002], one potential interpretation of overlap of activated regions in different tasks is the sharing of cognitive subprocess (“sharing view”). Following the abstract version of this interpretation, we assume that the overlap of activation for the control of shared representations and social-cognitive tasks is due to common underlying computational mechanisms. Recent research on the role of the TPJ and the aFMC supports the idea that rather domain-general cognitive operations are represented in these brain areas. In a series of studies on perspective taking, Ruby and Decety [2001, 2003] found consistent activations of the temporal-parietal region, irrespective of the investigated domain (motor or conceptual), reflecting the need to distinguish the self- and other-person perspective. In a similar vein, Decety and Lamm [2007] showed in a recent meta-analysis overlapping activations in the TPJ between agentive processing and ToM, but also during the reorienting of attention. Several studies have shown the consistent activation of the TPJ during reorienting of spatial attention [Corbetta et al., 2008; Mitchell, 2008]. This may be taken as evidence that the TPJ is involved in low-level, domain-general computa-

tional processes allowing a comparison of internal predictions with external events, which can be applied to different domains ranging from attentional to higher-level social-cognitive processes. Similarly, aFMC activations have been found for a variety of different social-cognitive functions and other types of cognitive processing [e.g. Amodio and Frith, 2006; Ochsner et al., 2004; Ramnani and Owen, 2004]. It is currently a matter of debate whether this region is dedicated to selective processing of social information or provides more general cognitive functions, such as a common form of meta-representation applicable to different sorts of input, as suggested for the TPJ region [Frith, 2007; Ochsner et al., 2004; Olsson and Ochsner, 2008]. Contributing to this discussion, the current study provides strong evidence for the idea that the aFMC and the TPJ are not only involved in complex social-cognitive tasks, but also in simple visuo-motor tasks, if a common set of computational processes is required in both tasks. Our findings thus provide the missing link between these different classes of tasks by assuming that social-cognitive abilities have evolved from the control of shared representations [Brass et al., 2005a; Brass and Spengler, 2008].

### Nonoverlapping activations between tasks

One potential alternative interpretation of our data is that not the same region is activated in all tasks but rather different subregions of this area [“subdivision view,” Cabeza and Nyberg, 2002]. To address this concern, we additionally defined the ROIs from the imitation–inhibition task at very conservative statistical thresholds. Even in these strictly defined regions in aFMC and TPJ, significant differences were found in the social-cognitive tasks, excluding the possibility of functionally distinct but neighboring regions, under the current spatial resolution of fMRI. However, this does not preclude the possibility that tasks differ in the extend of activation in a given area. The conjunction analysis revealed that the activations of the social-cognitive tasks extended into adjacent regions. In line with this observation, several recent studies have suggested that subdivisions within aFMC and TPJ might exist. In a meta-analysis of neuroimaging studies, Perner and Leekam [2008] proposed a distinction between dorsal parts (corresponding to the TPJ region) and more ventral aspects (around the posterior temporal sulcus, pSTS). Studies, which involved a perspective difference, including visual perspective-taking tasks, mentalizing tasks and deception, activated the TPJ area, whereas tasks with no contrast in perspectives were located around the pSTS. Concordantly, the imitation–inhibition task activated the dorsal aspect, as this task also requires a distinction of perspectives by assigning motor representation to either oneself or someone else.

Gilbert et al. [2006] showed that in anterior PFC an anterior–posterior gradient could be found, with rostral regions being sensitive to multiple-task co-ordination and

caudal regions in aFMC subserving mentalizing and self-reflection functions. Others have proposed that additionally a dorsal–ventral distinction within aFMC might exist. Based on the previous meta-analyses [Koski and Paus, 2000; Steele and Lawrie, 2004], it has been suggested that the anterior cingulate cortex (ACC) might be distinguished in a posterior cognitive zone and an anterior emotional region. Furthermore, a subcallosal region is related to emotional responses (including autonomic and visceral aspects) [Amodio and Frith, 2006; Koski and Paus, 2000; Ochsner et al., 2004]. Similarly, emotional in contrast to cognitive perspective-taking activated ventral versus dorsal aFMC, respectively [Hynes et al., 2006]. In all studies on imitative control, the activated region falls within the region associated with mentalizing [Amodio and Frith, 2006; Gilbert et al., 2006], but is located anterior and superior to regions associated with emotional processing. Conversely, these regions, sensitive to emotional processing, are activated by the self-referential task and partly by the ToM task, reflecting additional affective evaluation aroused by the task [Frith and Frith, 2003]. However, this emotional evaluation is not elicited no necessary for imitative control, thus showing no activations in these regions. Furthermore, a ventral–dorsal dissociation in regions anterior to the ACC has also been observed when reflecting about a similar versus dissimilar person, yielding activations in ventral compared to dorsal regions of aFMC, respectively [Mitchell et al., 2006]. Interestingly, the activation elicited by the control of imitation was localized in the ventral part closely to the part activated by mentalizing for similar others, whereas the ToM task also extended into more dorsal regions linked to mental state inferences about dissimilar others.

Finally, another important aspect that has to be considered, are the specific cortical networks in which the aFMC and the TPJ are embedded during the execution of different tasks [“network view,” Cabeza and Nyberg, 2002]. Consistent with the literature, mentalizing recruited additionally the temporal pole and the precuneus [Gallagher and Frith, 2003]. Activation of the precuneus is typically also found in self-referential tasks [Northoff and Berm-pohl, 2004]. Both areas have been associated with retrieval of autobiographical memory [Fink et al., 1996], to create a wider context for the material currently being processed [Frith and Frith, 2003]. These processes and associated regions are therefore crucial in the ToM and the self-referential task. However, in the imitation–inhibition task, these functional mechanisms are not of relevance to solve the task at hand, thus showing partly different networks active during imitative control.

### Functional associations between imitative control and social cognition

Additionally to the observed functional neuroanatomical overlap of tasks, we analyzed brain-behavior relationships to find further support for shared subfunctions between

tasks. Activity in aFMC was related to behavioral performance during imitative control, and furthermore this index of the individual ability to control shared representations was related to mentalizing functions. This strongly confirms the idea of a common computational process underlying both tasks. Studies focusing on common functions across tasks typically have concentrated on anatomical overlap [e.g. Derrfuss et al., 2004; Mitchell, 2008]. In addition, the current study extends this method by using a novel approach aiming at functional associations *between* tasks. Given the number of potential alternative interpretations for activation overlap in a given brain region, it is highly desirable to obtain converging evidence for a shared process view, from different dependent variables such as correlations of brain activity and behavioral data between tasks. In line with this idea, behavioral performance in the imitation–inhibition task significantly predicted performance in the ToM task. Further expanding this approach, a positive correlation was found between activity in aFMC during imitative control and mentalizing. This suggests that across different tasks there might be an individual level of regional activity, indexing the interindividual “functioning” in a certain brain region.

### Shared representations and social cognition: Towards an empirical and conceptual clarification

The interpretation put forward in the current study is based on the assumption that a shared representational system exists for perceived and internally planned actions [“common coding theory,” Prinz, 1997]. However, this system does not code explicitly by whom the corresponding motor representation was caused [Jeannerod, 1999]. This view suggests that the control of shared representations requires mechanisms that enable us to distinguish between self- or other-related actions. Furthermore, mechanisms that are involved in the control of shared representations functionally overlap with social-cognitive abilities, such as mentalizing. These findings complement the proposal that shared representations form a basis for action understanding and social cognition [Gallese and Goldman, 1998; Rizzolatti and Craighero, 2004]. Although this “mirroring” mechanism may be an essential part of mature mentalizing functions, it might only provide a basis for more advanced mind-reading abilities [Frith and Frith, 2006b], as an initial estimation of the mental states of the “mirrored” person. Mirroring the responses of others might be ideal to constantly track and monitor the changing actions and emotions of interaction partners, but this motor or emotional contagion does not always and unequivocally convey the cause for this action or emotion [Mitchell et al., 2006]. Conversely, the formation of more higher-level mentalizing capacities may be based on the ability to form representations of mental states of others and distinguish the other-perspective from the self-perspective [Decety and Grezes, 2006]. This might be necessary, for example, when attributing mental states that are different from own beliefs in the

false-belief task or distinguishing between being imitated from imitating another person [Decety et al., 2002; Hurley, 2005]. These specific mechanisms are thus rather related to the control of shared representations than to sharing representations and have reliably activated a set of regions, including the aFMC and TPJ [see also Brass and Spengler, 2008; Decety and Grezes, 2006; Frith and Frith, 2006b]. Hence, this description of cognitive processes implicated in mentalizing overlaps with our conceptualization of imitative control, as a prime example for assigning agency and exerting control over the shared representational system. This suggests that a pivotal precondition for mental state attribution is not the sharing of representations, but rather the building and distinction of representations related to self and other. It further implies that the functional system, which developed to distinguish self and other in the motor domain, subsequently generalized to more abstract representations, such as mental states and higher-level social cognition [Brass et al., 2005a; Brass and Spengler, 2008]. Hence, both processes, the management and control of shared representations and the understanding of other people's mental states, require the ability to distinguish one's own mental representation from the other person's representation. The present study strongly corroborates this theoretical view by showing shared neural resources in two key regions between the control of shared representations and functions of social cognition, such as mental state attribution.

Interestingly, the aFMC was the region in which brain activity was most reliably related to behavioral performance, suggesting that this region primarily drives the control of shared representations. One intriguing, but still speculative idea is that this region of medial prefrontal cortex is concerned with top-down modulation of posterior brain regions and possibly regions of the shared representational system [Frith and Frith, 2006a]. In general, the idea that prefrontal regions exert top-down control over posterior brain regions is very common in the domain of cognitive control [e.g. Brass et al., 2005b; Desimone and Duncan, 1995; Miller and D'Esposito, 2005]. Such controlled or top-down processes might also have a relevant role in social cognition [Satpute and Lieberman, 2006], with a network including the aFMC as the source of top-down signals that modulate signal processing in regions analyzing social signals, such as the pSTS [Frith and Frith, 2006a]. Similarly, it has been proposed that the aFMC is crucially involved in top-down control during emotional processing and empathic responses [Cheng et al., 2007; Decety and Lamm, 2007; Lamm et al., 2007]. Automatic activation of shared representations for emotional experiences of self and other might be modulated by the aFMC network, which supports meta-cognitive, reflective awareness of these emotional states [Olsson and Ochsner, 2008]. This view suggests that shared representations of the mirror system could also be subject to top-down modulation. In a neuroimaging study, it was shown that activation in the left inferior frontal gyrus was modu-

lated by selective attention and thus is the site at which descending inhibitory processes filter out or suppress task-irrelevant actions [Chong et al., 2008]. The inferior frontal gyrus has been proposed to be part of the core mirror system [Rizzolatti and Craighero, 2004], but it should be noted that this region has also been found to be activated in other domains, such as inhibitory control [Chikazoe et al., 2007; Koechlin and Jubault, 2006; Swick et al., 2008]. However, possible top-down processes on the mirror system may further depend on a network of regions implicated in ToM processing, such as the aFMC and the TPJ [Hamilton, 2008; Nishitani et al., 2004; Oberman et al., 2005]. This idea is corroborated by our data showing a negative association between aFMC and left vPMC, a region which has been assigned also to the mirror system.

In this sense, our findings allow to integrate two hitherto mutually exclusive positions on the origin of mind reading, by linking shared representations and mentalizing through the mechanisms involved in the control of shared representations. A recent account has been put forward which proposes a first integration of those views [Keysers and Gazzola, 2006, 2007], assuming that social cognition comprises both a prereflective, experiential representation of bodily states (obtained through the shared representational system), which can be passed on to the aFMC to form reflective, meta-cognitive representation of those states. However, our findings suggest that the flow of information from the shared representational system is not unidimensional, but rather that the aFMC may control shared representations, by enhancing or inhibiting the automatically evoked activations of this system [see also Decety and Lamm, 2006; Olsson and Ochsner, 2008]. Thus, it would allow to introduce deliberation from external stimuli and flexibility of responses, enabling the individual to escape automatic responses and follow current intentions [Frith and Frith, 2006a]. This extended model would also assign an important role to the TPJ. Self-other distinction is a necessary condition to successfully manage and separate shared representations [Decety and Grezes, 2006], possibly via signaling discrepancies between predicted and perceived events [Decety and Lamm, 2007]. Furthermore, it might be crucial for promoting self-awareness functions of the aFMC needed for the building of meta-cognitive, reflective intentional states of either self or similar others [Decety and Lamm, 2006]. Recognizing the different perspectives, emotions or actions of others may be an essential building block for explicit awareness of oneself, knowing that others' minds can be different from one's own. This view also implies different functional roles of the two crucial brain regions, which is supported by our data, as the overlap with the agency task occurred only in the TPJ and the self-referential task solely activated a common region in the aFMC. A recent fMRI study showed a dissociation of these cortical regions during the control of shared representations, with the TPJ being related to agentic processing and the aFMC responding only to a conflict between intended and externally



triggered motor representations through enforcing one's own intention (Brass et al., in press).

Hence, an important topic for future research may be to determine further the specific functions of these "social" brain regions and their mode of interaction with each other. For now, this study offers considerable evidence for shared functioning between two different strings of research and provides evidence for the promising future application of a "cross-functional," within-subject approach [Cabeza and Nyberg, 2002; Decety and Lamm, 2007; Mitchell, 2008] to develop new accounts that integrate findings across different functions.

## REFERENCES

- Aichhorn M, Perner J, Kronbichler M, Staffen W, Ladurner G (2006): Do visual perspective tasks need theory of mind? *NeuroImage* 30:1059–1068.
- Amodio DM, Frith CD (2006): Meeting of minds: The medial frontal cortex and social cognition. *Nat Rev Neurosci* 7:268–277.
- Brass M, Spengler S (2008): The inhibition of imitative behavior and attribution of mental states. In: Striano T, Reid V, editors. *Social Cognition: Development, Neuroscience and Autism*. Oxford: Blackwell.
- Brass M, Bekkering H, Wohlschlaeger A, Prinz W (2000): Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain Cogn* 44: 124–143.
- Brass M, Bekkering H, Prinz W (2001a): Movement observation affects movement execution in a simple response task. *Acta Psychol Amst* 106:3–22.
- Brass M, Zysset S, von Cramon DY (2001b): The inhibition of imitative response tendencies. *NeuroImage* 14:1416–1423.
- Brass M, Derrfuss J, Matthes-von Cramon G, von Cramon DY (2003): Imitative response tendencies in patients with frontal brain lesions. *Neuropsychology* 17:265–271.
- Brass M, Derrfuss J, von Cramon DY (2005a): The inhibition of imitative and overlearned responses: A functional double dissociation. *Neuropsychologia* 43:89–98.
- Brass M, Ullsperger M, Knoesche TR, von Cramon DY, Phillips NA (2005b): Who comes first? The role of the prefrontal and parietal cortex in cognitive control. *J Cogn Neurosci* 17:1367–1375.
- Brass M, Ruby P, Spengler S (in press): Inhibition of imitative behavior and social cognition. *Philos Trans R Soc B*.
- Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V, Seitz RJ, Zilles K, Rizzolatti G, Freund HJ (2001): Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *Eur J Neurosci* 13:400–404.
- Cabeza R, Nyberg L (2002): Seeing the forest through the trees: The cross-functional approach to imaging cognition. In: Zani A, Proverbio AM, editors. *The Cognitive Electrophysiology of Mind and Brain*. San Diego: Academic Press. pp 41–68.
- Chartrand TL, Bargh JA (1999): The chameleon effect: The perception-behavior link and social interaction. *J Pers Soc Psychol* 76:893–910.
- Cheng Y, Lin CP, Liu HL, Hsu YY, Lim KE, Hung D, Decety J (2007): Expertise modulates the perception of pain in others. *Curr Biol* 17:1708–1713.
- Chikazoe J, Konishi S, Asari T, Jimura K, Miyashita Y (2007): Activation of right inferior frontal gyrus during response inhibition across response modalities. *J Cogn Neurosci* 19: 69–80.
- Chong TT, Williams MA, Cunnington R, Mattingley JB (2008): Selective attention modulates inferior frontal gyrus activity during action observation. *NeuroImage* 40:298–307.
- Corbetta M, Patel G, Shulman GL (2008): The reorienting system of the human brain: From environment to theory of mind. *Neuron* 58:306–324.
- Decety J, Sommerville JA (2003): Shared representations between self and other: A social cognitive neuroscience view. *Trends Cogn Sci* 7:527–533.
- Decety J, Grezes J (2006): The power of simulation: Imagining one's own and other's behavior. *Brain Res* 1079:4–14.
- Decety J, Lamm C (2006): Human empathy through the lens of social neuroscience. *Sci World J* 6:1146–1163.
- 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* 13:580–593.
- Decety J, Chaminade T, Grezes J, Meltzoff AN (2002): A PET exploration of the neural mechanisms involved in reciprocal imitation. *NeuroImage* 15:265–272.
- Derrfuss J, Brass M, von Cramon DY (2004): Cognitive control in the posterior frontolateral cortex: Evidence from common activations in task coordination, interference control, and working memory. *NeuroImage* 23:604–612.
- Desimone R, Duncan J (1995): Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18:193–222.
- di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G (1992): Understanding motor events: A neurophysiological study. *Exp Brain Res* 91:176–180.
- Fadiga L, Fogassi L, Pavesi G, Rizzolatti G (1995): Motor facilitation during action observation: A magnetic stimulation study. *J Neurophysiol* 73:2608–2611.
- Farrer C, Frith CD (2002): Experiencing oneself vs another person as being the cause of an action: The neural correlates of the experience of agency. *NeuroImage* 15:596–603.
- Farrer C, Franck N, Georgieff N, Frith CD, Decety J, Jeannerod M (2003): Modulating the experience of agency: A positron emission tomography study. *NeuroImage* 18:324–333.
- Fink GR, Markowitsch HJ, Reinkemeier M, Bruckbauer T, Kessler J, Heiss WD (1996): Cerebral representation of one's own past: Neural networks involved in autobiographical memory. *J Neurosci* 16:4275–4282.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995): Improved assessment of significant activation in functional magnetic resonance imaging fMRI: Use of a cluster-size threshold. *Magn Reson Med* 33:636–647.
- Friston KJ, Frith CD, Frackowiak RS, Turner R (1995): Characterizing dynamic brain responses with fMRI: A multivariate approach. *NeuroImage* 2:166–172.
- Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R (1998): Event-related fMRI: Characterizing differential responses. *NeuroImage* 7:30–40.
- Frith CD (2007): The social brain? *Philos Trans R Soc Lond B Biol Sci* 362:671–678.
- Frith U, Frith CD (2003): Development and neurophysiology of mentalizing. *Philos Trans R Soc Lond B Biol Sci* 358:459–473.
- Frith CD, Frith U (2006a): How we predict what other people are going to do. *Brain Res* 1079:36–46.

- Frith CD, Frith U (2006b): The neural basis of mentalizing. *Neuron* 50:531–534.
- Gallagher II (2000): Philosophical conceptions of the self: Implications for cognitive science. *Trends Cogn Sci* 4:14–21.
- Gallagher HL, Frith CD (2003): Functional imaging of ‘theory of mind’. *Trends Cogn Sci* 7:77–83.
- Gallagher HL, Happe F, Brunswick N, Fletcher PC, Frith U, Frith CD (2000): Reading the mind in cartoons and stories: An fMRI study of ‘theory of mind’ in verbal and nonverbal tasks. *Neuropsychologia* 38:11–21.
- Gallese V, Goldman A (1998): Mirror neurons and the simulation theory of mind-reading. *Trends in Cogn Sci* 2:493–501.
- Gangitano M, Mottaghy FM, Pascual-Leone A (2004): Modulation of premotor mirror neuron activity during observation of unpredictable grasping movements. *Eur J Neurosci* 20:2193–2202.
- Gilbert SJ, Spengler S, Simons JS, Steele JD, Lawrie SM, Frith CD, Burgess PW (2006): Functional specialization within rostral prefrontal cortex (area 10): A meta-analysis. *J Cogn Neurosci* 18:932–948.
- Glover GH (1999): Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage* 9:416–429.
- Hamilton, AF (2008): Emulation and mimicry for social interaction: A theoretical approach to imitation in autism. *Q J Exp Psychol Colchester* 61:101–115.
- Happe F, Brownell H, Winner E (1999): Acquired ‘theory of mind’ impairments following stroke. *Cognition* 70:211–240.
- Henson R (2005): What can functional neuroimaging tell the experimental psychologist? *Q J Exp Psychol A* 58:193–233.
- Heyes CM, Bird G (2007): Mirroring, association and the correspondence problem. In: Haggard P, Rossetti Y, Kawato M, editors. *Sensorimotor Foundations of Higher Cognition, Attention and Performance XX*. Oxford: University Press. pp 461–479.
- Hurley S (2005): The shared circuits hypothesis: A unified functional architecture for control, imitation, and simulation. In: Hurley S, Chater N, editors. *Perspectives on Imitation*. Cambridge, MA: MIT Press. pp 177–194.
- Hynes CA, Baird AA, Grafton ST (2006): Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. *Neuropsychologia* 44:374–383.
- Jeannerod M (1999): The 25th Bartlett Lecture. To act or not to act: Perspectives on the representation of actions. *Q J Exp Psychol A* 52:1–29.
- Jeannerod M (2003): The mechanism of self-recognition in humans. *Behav Brain Res* 142:1–15.
- Joseph JE, Partin DJ, Jones KM (2002): Hypothesis testing for selective, differential, and conjoined brain activation. *J Neurosci Methods* 118:129–140.
- Keysers C, Gazzola V (2006): Towards a unifying neural theory of social cognition. *Prog Brain Res* 156:379–401.
- Keysers C, Gazzola V (2007): Integrating simulation and theory of mind: From self to social cognition. *Trends Cogn Sci* 11:194–196.
- Kilner JM, Paulignan Y, Blakemore SJ (2003): An interference effect of observed biological movement on action. *Curr Biol* 13:522–525.
- Koechlin E, Jubault T (2006): Broca’s area and the hierarchical organization of human behavior. *Neuron* 50:963–974.
- Koski L, Paus T (2000): Functional connectivity of the anterior cingulate cortex within the human frontal lobe: A brain-mapping meta-analysis. *Exp Brain Res* 133:55–65.
- Lamm C, Nusbaum HC, Meltzoff AN, Decety J (2007): What are you feeling? Using functional magnetic resonance imaging to assess the modulation of sensory and affective responses during empathy for pain. *PLoS ONE* 2:e1292.
- Leube DT, Knoblich G, Erb M, Grodd W, Bartels M, Kircher TT (2003): The neural correlates of perceiving one’s own movements. *NeuroImage* 20:2084–2090.
- Lhermitte F (1986): Human autonomy and the frontal lobes, Part II: Patient behavior in complex and social situations: The “environmental dependency syndrome”. *Ann Neurol* 19:335–343.
- Lohmann G, Muller K, Bosch V, Mentzel H, Hessler S, Chen L, Zysset S, von Cramon DY (2001): LIPSIA—A new software system for the evaluation of functional magnetic resonance images of the human brain. *Comput Med Imaging Graph* 25:449–457.
- Miller BT, D’Esposito M (2005): Searching for “the top” in top-down control. *Neuron* 48:535–538.
- Mitchell JP (2008): Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cereb Cortex* 18:262–271.
- Mitchell JP, Macrae CN, Banaji MR (2006): Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron* 50:655–663.
- Nichols T, Brett M, Andersson J, Wager T, Poline JB (2005): Valid conjunction inference with the minimum statistic. *NeuroImage* 25:653–660.
- Nishitani N, Avikainen S, Hari R (2004): Abnormal imitation-related cortical activation sequences in Asperger’s syndrome. *Ann Neurol* 55:558–562.
- Northoff G, Bermpohl F (2004): Cortical midline structures and the self. *Trends Cogn Sci* 8:102–107.
- Oberman LM, Hubbard EM, McCleery JP, Altschuler EL, Ramachandran VS, Pineda JA (2005): EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Brain Res Cogn Brain Res* 24:190–198.
- Ochsner KN, Knierim K, Ludlow DH, Hanelin J, Ramachandran T, Glover G, Mackey SC (2004): Reflecting upon feelings: An fMRI study of neural systems supporting the attribution of emotion to self and other. *J Cogn Neurosci* 16:1746–1772.
- Olsson A, Ochsner KN (2008): The role of social cognition in emotion. *Trends Cogn Sci* 12:65–71.
- Perner J, Leekam S (2008): The curious incident of the photo that was accused of being false: Issues of domain specificity in development, autism, and brain imaging. *Q J Exp Psychol Colchest* 61:76–89.
- Prinz W (1997): Perception and action planning. *Eur J Cogn Psychol* 9:129–154.
- Ramnani N, Owen AM (2004): Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. *Nat Rev Neurosci* 5:184–194.
- Rizzolatti G, Craighero L (2004): The mirror-neuron system. *Annu Rev Neurosci* 27:169–192.
- Ruby P, Decety J (2001): Effect of subjective perspective taking during simulation of action: A PET investigation of agency. *Nat Neurosci* 4:546–550.
- Ruby P, Decety J (2003): What you believe versus what you think they believe: A neuroimaging study of conceptual perspective-taking. *Eur J Neurosci* 17:2475–2480.
- Sato A, Yasuda A (2005): Illusion of sense of self-agency: Discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition* 94:241–255.
- Satpute AB, Lieberman MD (2006): Integrating automatic and controlled processes into neurocognitive models of social cognition. *Brain Res* 1079:86–97.

- Saxe R, Powell LJ (2006): It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychol Sci* 17:692–699.
- Spengler S, von Cramon DY, Brass M (2009): Was it me or was it you? How the sense of agency originates from ideomotor learning revealed by fMRI. *NeuroImage* 46:290–298.
- Steele JD, Lawrie SM (2004): Segregation of cognitive and emotional function in the prefrontal cortex: A stereotactic meta-analysis. *NeuroImage* 21:868–875.
- Strafella AP, Paus T (2000): Modulation of cortical excitability during action observation: A transcranial magnetic stimulation study. *Neuroreport* 11:2289–2292.
- Swick D, Ashley V, Turken U (2008): Left inferior frontal gyrus is critical for response inhibition. *BMC Neurosci* 9:102.
- Talairach P, Tournoux J (1998): Co-planar stereotactic atlas of the human brain. Stuttgart: Thieme.
- Vogeley K, Bussfeld P, Newen A, Herrmann S, Happe F, Falkai P, Maier W, Shah NJ, Fink GR, Zilles K (2001): Mind reading: Neural mechanisms of theory of mind and self-perspective. *NeuroImage* 14:170–181.
- Ward BD (2000): Simultaneous inference for fMRI data. Retrieved August 23, 2007, from <http://afni.nimh.nih.gov/afni/doc/manual/AlphaSim>.
- Worsley KJ, Friston KJ (1995): Analysis of fMRI time-series revisited again. *NeuroImage* 2:173–181.
- Zysset S, Huber O, Ferstl E, von Cramon DY (2002): The anterior frontomedian cortex and evaluative judgment: An fMRI study. *NeuroImage* 15:983–991.