Psychological Medicine, Page 1 of 10.
doi:10.1017/S003329171000228X

ORIGINAL ARTICLE

The narcissistic self and its psychological and neural correlates: an exploratory fMRI study

Y. Fan[1,2,3]†, C. Wonneberger[1,2]†, B. Enzi[2], M. de Greck[4], C. Ulrich[5], C. Tempelmann[6], B. Bogerts[2], S. Doering[7,8,9] and G. Northoff[1]*

[1] *Institute of Mental Health Research, University of Ottawa, Canada*
[2] *Department of Psychiatry, Otto-von-Guericke University Magdeburg, Germany*
[3] *Cluster of Excellence 'Language of Emotion' and Dahlem Institute for Neuroimaging of Emotion, Freie Universität Berlin, Berlin, Germany*
[4] *Department of Psychology, Peking University, Beijing, China*
[5] *Department of Psychotherapeutic Medicine, Fachklinikum Uchtspringe, Germany*
[6] *Department of Neurology, Otto-von-Guericke University Magdeburg, Germany*
[7] *Psychosomatics in Dentistry, Department of Prosthodontics and Material Sciences, University of Muenster, Germany*
[8] *Department of Psychosomatics and Psychotherapy, University of Muenster, Germany*
[9] *Department of Psychoanalysis and Psychotherapy, Medical University Vienna, Austria*

Background. The concept of narcissism has been much researched in psychoanalysis and especially in self psychology. One of the hallmarks of narcissism is altered emotion, including decreased affective resonance (e.g. empathy) with others, the neural underpinnings of which remain unclear. The aim of our exploratory study was to investigate the psychological and neural correlates of empathy in two groups of healthy subjects with high and low narcissistic personality trait. We hypothesized that high narcissistic subjects would show a differential activity pattern in regions such as the anterior insula that are typically associated with empathy.

Method. A sample of 34 non-clinical subjects was divided into high ($n=11$) and low ($n=11$) narcissistic groups according to the 66th and 33rd percentiles of their scores on the Narcissism Inventory (NI). Combining the psychological, behavioral and neuronal [i.e. functional magnetic resonance imaging (fMRI)] measurements of empathy, we compared the high and low narcissistic groups of subjects.

Results. High narcissistic subjects showed higher scores on the Symptom Checklist-90 – Revised (SCL-90-R) and the 20-item Toronto Alexithymia Scale (TAS-20) when compared to low narcissistic subjects. High narcissistic subjects also showed significantly decreased deactivation during empathy, especially in the right anterior insula.

Conclusions. Psychological and neuroimaging data indicate respectively higher degrees of alexithymia and lower deactivation during empathy in the insula in high narcissistic subjects. Taken together, our preliminary findings demonstrate, for the first time, psychological and neuronal correlates of narcissism in non-clinical subjects. This might stipulate both novel psychodynamic conceptualization and future psychological–neuronal investigation of narcissism.

Received 4 May 2010; Revised 20 October 2010; Accepted 20 October 2010

Key words: Empathy, fMRI, narcissism, right anterior insula, self.

Introduction

The concept of narcissism originates in psychodynamic theory and is also central in clinical psychiatry. Clinically, the concept of narcissism was developed by Kohut (1971) and Kernberg (1975), and has been the basis for the incorporation of the narcissistic personality disorder into the DSM-III (APA, 1980) (see discussion for more details about the concept of narcissism). Narcissism is also regarded as a trait of the healthy personality (Emmons, 1984) and a prominent behavior pattern in western cultures (Lasch, 1979). Today, narcissism is widely seen as a continuum from healthy self-esteem to severe narcissistic pathology (Watson *et al*. 2002; Ritter & Lammers, 2007).

Patients with narcissistic personality disorder have been reported to show abnormalities not only in their self but also in empathy in terms of reduced affective resonance with other persons (Dimaggio *et al*. 2006). Although, conceptually and clinically, the relationship between self and non-self (i.e. the other) in narcissism

* Address for correspondence: Dr G. Northoff, M.D., Ph.D., F.R.C.P.C., Research Unit Director, Mind, Brain Imaging and Neuroethics, Canada Research Chair, The Michael Smith Chair, ELJB-CIHR, Royal Ottawa Healthcare Group, University of Ottawa Institute of Mental Health Research, 1145 Carling Avenue, Room 6435, Ottawa, ON K1Z 7K4, Canada.
(Email: georg.northoff@rohcg.on.ca)
† These authors contributed equally to this work.

has been widely discussed, the psychological and neuronal manifestations of empathy and its relationship to the narcissistic self remain unclear. Empathy describes the sharing and understanding of the emotional and cognitive states of others; refined neuroimaging techniques have enabled the identification of specific brain regions involved in empathy (Decety & Lamm, 2006; Singer & Lamm, 2009). Core regions identified in the neural network underlying empathy are the anterior insula, the left inferior frontal cortex (including the mirror neurons), and other regions such as the premotor cortex and the dorsolateral prefrontal cortex (DLPFC; Decety & Lamm, 2006; Singer & Lamm, 2009). Although the involvement of these regions in empathy has been demonstrated in several studies, the relationship between this empathy network and narcissism remains unclear. Among these regions implicated in empathy, the anterior insula is most notable because it has been shown to be involved not only in empathy with others but also in focusing on the own self (Enzi *et al*. 2009; Modinos *et al*. 2009). Because subjects with high narcissism show a high self-focus, we can preliminarily hypothesize altered activity in the anterior insula in these subjects when compared to those with low narcissism.

The aim of this exploratory study was to investigate the psychological, behavioral and neuronal manifestations of empathy in a non-clinical sample with high and low narcissism respectively. Self-ratings of narcissism, empathy and clinical symptoms (e.g. somatization, obsessive–compulsive, interpersonal sensitivity, depression, anxiety, anger–hostility, phobic anxiety, paranoid ideation and psychoticism) were applied. These were complemented by a functional magnetic resonance imaging (fMRI) study of empathic perception of emotional faces that have been shown to induce an internal affective representation of the observed face in the observer (Preston & de Waal, 2002; Singer & Lamm, 2009). Based on theoretical conceptualizations and clinical observations, it was hypothesized that non-clinical subjects with high narcissism differ in psychological, behavioral and neuronal measures of empathy when compared to those with low narcissism.

Method

Subjects

Thirty-four subjects (21 females, 13 males, age 40.09 ± 8.71 years, four left-handed) were investigated. Each of them received €30 compensation for voluntary participation in the study, other than the amount they obtained in the reward trials. None of the subjects reported a history of head injury or trauma, central nervous system disease, developmental disorder, medical disorder, substance abuse or psychiatric hospitalization. None were receiving psychotropic medication. All subjects gave written informed consent and the study was approved by the local review board.

Psychological measures

We included a scale on narcissism and also, with regard to our emotional focus, scales for empathy and emotions. In addition, we included a measure for depression, the Beck Depression Inventory (BDI), so as to exclude depression as a possible confounding variable, and also accounted for general non-specific clinical symptoms with the Symptom Checklist-90 – Revised (SCL-90-R). The German versions of the following questionnaires were used: (*a*) the Narcissism Inventory (NI; Denecke & Hilgenstock, 1989), a 163-item questionnaire that covers four dimensions of narcissism: (1) the threatened self, (2) the 'classical' narcissistic self, (3) the idealist self, and (4) the hypochondriac self; (*b*) the 20-item Toronto Alexithymia Scale (TAS-20; Bach *et al*. 1996); (*c*) the Interpersonal Reactivity Index (Paulus, 2009) to assess empathy on the four subscales fantasy, empathic concern, perspective-taking, and personal distress; (*d*) the SCL-90-R (Franke, 1995), and (*e*) the BDI (Hautzinger *et al*. 1994).

Experimental task

We presented emotional pictures of faces and instructed the subjects to empathize with the respective person. After the presentation of the pictures, the subjects were instructed to give an online judgment of their degree of empathy on a visual analogue scale presented on the screen. Following a recent hypothesis (Preston & de Waal, 2002), the mere perception of an emotional face should be sufficient for inducing an empathic response. Hence, we presented no faces (i.e. smoothed faces as the non-empathic control condition; see Supplementary method online for details), although this required additional investigations to control for various confounding factors (discussed later in detail). Every empathy trial began with the 5 s display of an emotional face or the presentation of a control stimulus. Subjects were instructed to empathize with the emotional face presented. Immediately after the presentation of the emotional face, a visual analogue scale was presented and subjects were asked to subjectively judge and evaluate their extent of affective sharing. In consideration of excluding the motor confound associated with the judgment and evaluation phase, the empathy task was preceded by a short finger-tapping task, which identifies motor regions for further exclusive masking analysis. Moreover, to exclude possible confounding variables, such as face and emotion perception, general evaluation and reward (Singer *et al*. 2006), in our imaging results of empathy, we conducted additional and complementary studies focusing on these factors (see online Supplement). This allowed us to control for statistical independence, a recently highlighted methodological issue of imaging studies (Kriegeskorte *et al*. 2009). The main fMRI experiment consisted of six blocks, each 630 s in duration. Blocks 2, 4 and 6 were empathy blocks, and blocks 1, 3 and 5 were reward blocks where we used a reward paradigm, a slightly modified version of the Monetary Incentive Delay Task (Knutson *et al*. 2001*a*, *b*).

Before entering the scanner, each subject read detailed information of the paradigm and completed a few trials to familiarize themselves with the task. While lying in the scanner, the stimuli were displayed using the 'Presentation' software package (Neurobehavioral Systems, USA) and were projected onto a matt screen using an LCD projector, visible through a mirror mounted on the headcoil (see online Supplement).

fMRI data acquisition and analysis

The fMRI data were collected in a 1.5-T MR scanner (General Electric Sigma Horizon, USA) using the standard circular polarized headcoil. Using a midsagittal scout image, a stack of 23 slices was aligned parallel to the bicomissural plane. During each functional run, 320 whole-brain volumes were acquired [gradient echo planar imaging (EPI), repetition time (TR) = 2 s, echo time (TE) = 35 ms, flip angle = 80°, field of view = 200×200 mm^2, slice thickness = 5 mm, interslice gap = 1 mm, spatial resolution = $3.125 \times 3.125 \times 5$ mm^3].

Image processing and statistical analyses were carried out according to the general linear model approach using the SPM2 software package running on MATLAB 6.5.1 (The Mathworks Inc., USA). The first five volumes were discarded due to saturation effects. All functional images were slice-time corrected with reference to the first slice acquired, corrected for motion artifacts by realignment to the first volume, and spatially normalized to a standard T1-weighted SPM template. The images were resampled to $2 \times 2 \times 2$ mm^3 and smoothed with an isotropic 6-mm full-width half-maximum Gaussian kernel. The time-series fMRI data were filtered using a high-pass filter with a cut-off at 128 s.

The first-level analysis modeled data of each subject using an event-related design by convolving the blood oxygen level-dependent (BOLD) signal with a canonical SPM hemodynamic response function. The movement parameters (translation and rotation) resulting from realignment were included as regressors to account for any residual effects of head motion. Regionally specific condition effects were tested by performing linear contrasts for each subject between conditions of interest.

To identify the empathy network as such we performed whole-group analysis with all 34 subjects. We thus compared the contrast 'empathy > non-empathy' and elucidated signal changes for the whole group. The contrast images resulting from the first-level analysis were submitted to a second-level random-effects analysis by applying a one-sample *t* test to the images acquired for all subjects. False discovery rate (FDR) correction was applied to control for multiple comparison using a threshold of FDR-corrected $p < 0.01$, cluster size > 50 voxels. The anatomical localization of significant activations was assessed with reference to the standard stereotactic atlas by superimposition of the SPM maps on a standard brain template [Montreal Neurological Institute (MNI)] provided by SPM2. To exclude possible movement confound associated with the judgment phase, activation in the whole-group analysis was then exclusively masked with the motor region found in the finger-tapping contrast. Furthermore, the activation found in this group analysis of empathy contrast was compared with the results of a meta-analysis of previous fMRI studies on empathy (see online Supplement).

The whole-group analysis was followed by group comparison between high and low narcissism groups using a two-sample *t* test. To investigate the neural difference between high and low narcissism groups in their empathic perception, the contrast images of the two groups for the first-level analysis 'empathy > non-empathy' were submitted to a second-level group comparison by applying an independent-sample *t* test. The threshold level was first set to 0.001 uncorrected, and then followed by a small volume correction with an independent region of interest (ROI) in the right anterior insula defined by the meta-analysis results of previous fMRI studies on empathy (see online Supplement). To further reveal the hemodynamic process during empathy, which differs in high and low narcissism groups, the results of this group comparison were selected as ROIs, in which the original BOLD signal changes were extracted and analyzed with the Marseille Region of Interest Toolbox software package (MarsBaR 0.42, http://marsbar.sourceforge.net/) (see online Supplement).

These analyses were supplemented by several analyses to account for possible confounding functions. To independently validate activations of the whole-group analysis for empathy, our results were compared with previous findings of empathy that resulted from a recent meta-analysis (see online Supplement). The possible impact of mere emotion processing was accounted for by comparing an emotional face with a neutral face in the signal changes of ROIs resulting from the group comparison (see online Supplement). Possible differences between high and low narcissism in face processing were tested by comparing signal changes extracted from the bilateral fusiform face area, which were elucidated in the whole-group analysis (see online Supplement). Unspecific task-related effects were controlled for by an additional empathy measure in fMRI in a different group of healthy subjects. There we compared the judgment of empathy with the judgment of skin color, which allowed us to control for unspecific task-related effects of judgment, that is general evaluation. Percentages of signal changes were extracted from this control experiment in ROIs defined by the results of our group comparison, and were compared between empathy and skin-color judgments. It was hypothesized that, if these activations are due to a general evaluation process, they would not differ between the judgment of empathy and the judgment of skin color; however, if they are indeed related to empathy, differences in 'empathy *versus* skin color' would be expected (see online Supplement and the results section for further details and reasoning). Finally, because previous data (Singer *et al*. 2004, 2006) showed a strong overlap of empathy with reward, we also excluded possible reward effects. We therefore tested for the difference between 'anticipation of reward > anticipation of no outcome' between high and low narcissism in the regional coordinates resulting from the high *versus* low narcissistic group comparison (see online Supplement).

Results

Psychological and behavioral data

Eleven subjects scored beyond the 66th percentile of the NI (Denecke & Hilgenstock, 1989) and were regarded as highly narcissistic (six females, five males, age 40.09 ± 10.34 years), the medium group consisted of 11 subjects (eight females, three males, age 40.18 ± 10.08 years), and the low narcissism group had scores below the 33rd percentile (seven females, four males, age 40.18 ± 6.42 years) (two subjects were excluded from the group comparison because of missing data). The high and low narcissism subgroups differed significantly in their narcissism total score (Table 1). No significant differences between the groups occurred with regard to age [$T(20) = -0.025$, $p = 0.98$], gender ($\chi^2 = 0.188$, $p = 0.67$) and general intelligence (Table 1). High narcissism subjects showed significantly higher alexithymia scores than low narcissism subjects (Table 1). No significant differences occurred in the ability to detect emotions in faces or spoken sentences (the Florida Affect Battery) and empathy (the Interpersonal Reactivity Index). This was confirmed by the lack of group differences in the post-scanning results of empathy and other dimensions (Table 1). High narcissism subjects yielded significantly higher psychopathology including depression (Table 1).

Neuronal data

We first performed a one-sample t test across the whole sample ($n = 34$) to detect regions activated in the contrast 'empathy > non-empathy'. This analysis yielded several regions typically associated with empathy (Supplementary Table 1) that were not confounded by movement associated with the judgment phase (Supplementary Fig. 1*a*), and that were further confirmed by a meta-analysis on previous fMRI studies of empathy (Supplementary Fig. 1*b*).

Subjects with high and low narcissism were then compared with each other. Using an SPM-based two-sample t test on the contrast 'empathy > non-empathy', regional signal changes differed significantly between the groups in the right anterior insula, right DLPFC, posterior cingulate cortex, and right lateral premotor cortex (Fig. 1 and Table 2). Signal change analysis for these regions in both conditions revealed that high narcissism subjects showed either decreased deactivation, as in the right anterior insula, or increased activation, as in the right DLPFC, the posterior cingulate cortex and the premotor cortex, during the control (that is, the non-empathic task; Fig. 1).

Methodologically, it is important to note that our fMRI paradigm and our experimental settings were developed for measuring 'affect sharing', a distinct component of the multi-level concept of 'empathy' (Singer & Lamm, 2009). More specifically, we relied here on empathic face perception, which, according to a recent model of empathy (Preston & de Waal, 2002), should be sufficient for inducing an internal affective representation in the observer, which forms the basis of empathic response (Hooker *et al*. 2008). Although we aimed to control for possible confounding factors such as emotion processing, unspecific task-related effects, reward effects and face processing, we are aware of the experimental and conceptual difficulties in this field and refer the reader to a review highlighting the above-mentioned concept of empathy and its differentiation from the closely related concept of mentalizing (Singer, 2006).

Table 1. *Comparison between high and low NI groups in psychological and behavioral data*

	Low NI ($n=11$) Mean (S.D.)	High NI ($n=11$) Mean (S.D.)	Group comparison: low NI *v.* high NI	
			T(df)	p value
Test of cognitive performance (LPS-3)	134.82 (9.74)	136.73 (12.61)	$T(20)=0.397$	0.70
Multiple Choice Vocabulary Intelligence Test (MWT-B)	142.27 (2.97)	138.73 (6.25)	$T(20)=-1.700$	0.11
Narcissism Inventory (NI)				
Threatened self	16.42 (1.30)	23.69 (8.05)	$T(20)=2.952$	0.014
'Classical' narcissistic self	20.98 (4.08)	28.20 (5.34)	$T(20)=3.566$	0.002
Idealist self	26.36 (2.93)	32.19 (4.36)	$T(20)=3.679$	0.002
Hypochondriac self	13.82 (3.49)	20.00 (6.66)	$T(20)=2.726$	0.016
Total score	19.36 (1.61)	26.17 (5.09)	$T(20)=4.237$	0.001
Toronto Alexithymia Scale (TAS-20)	33.36 (6.28)	43.40 (9.54)	$T(19)=2.875$	0.010
Interpersonal Reactivity Index (IRI)				
Fantasy scale	20.73 (2.83)	22.20 (4.64)	$T(19)=0.888$	0.39
Empathic concern	25.82 (3.71)	25.10 (4.48)	$T(19)=-0.401$	0.69
Perspective-taking	23.73 (3.55)	22.40 (2.99)	$T(19)=-0.921$	0.37
Personal distress	16.00 (3.44)	18.80 (3.12)	$T(19)=1.948$	0.07
Post-hoc rating indices for state emotional reaction towards the stimuli				
Empathy	59.33 (15.24)	60.00 (14.65)	$T(20)=-0.104$	0.92
Perspective-taking	56.36 (14.71)	58.64 (14.60)	$T(20)=-0.366$	0.72
Emotional intensity	59.42 (6.22)	56.57 (9.39)	$T(20)=0.839$	0.41
Emotional valence	41.00 (4.34)	38.44 (5.59)	$T(20)=1.199$	0.25
Personal relevance	35.41 (10.86)	43.99 (13.53)	$T(20)=-1.641$	0.12
Sympathy	48.11 (4.81)	47.99 (10.59)	$T(20)=0.035$	0.97
SCL-90-R GSI	41.73 (5.52)	53.50 (6.77)	$T(19)=4.387$	<0.001
BDI global score	1.91 (2.26)	5.20 (4.39)	$T(19)=2.191$	0.041

LPS-3, Achievement Measure System (Leistungsprüfsystem; Horn, 1983); MWT-B, the Multiple Choice Vocabulary Intelligence Test–B (Mehrfachwahl-Wortschatz-Intelligenztest–B; Lehrl *et al.* 1995); SCL-90-R, Symptom Checklist-90 – Revised; GSI, Global Severity Index; BDI, Beck Depression Inventory; S.D., Standard deviation; df, degrees of freedom.

To rule out the impact of possible confounding factors, such as general face perception (as distinguished from face recognition), general evaluation and reward in these regions, we conducted a series of consecutive analyses (Table 2). First, to exclude differences in face processing, we compared the two groups with regard to neural activity in the fusiform face area with the coordinates for this region taken from the whole-group analysis. This revealed no significant difference between high and low narcissism groups in the activation of the left fusiform face area (−18, −92, −22) [$F(1,20)=1.093$, $p>0.3$] and the right fusiform face area (46, −70, −16) [$F(1,20)=0.481$, $p>0.4$].

To account for the confounding effect of unspecific and general emotion processing (as distinct from affective sharing and thus empathy), the signal changes from four ROIs resulting from 'high narcissism *versus* low narcissism' were entered as functional localizers for the analysis of the same data with regard to empathy for emotional and neutral faces (in half of the subjects who were also subjected to neutral stimuli, i.e. neutral faces). BOLD signal changes in the right anterior insula and the right posterior cingulate cortex showed no significant difference between emotional and neutral empathy conditions. However, signal changes in both regions, that is the right anterior insula and right posterior cingulate cortex (unlike in the right DLPFC and premotor cortex), during emotional and neutral empathy differed from the non-empathic condition, that is the smooth faces (Supplementary Fig. 2*a*). This means that the factor emotion did not play a significant role in the observed activity changes in the right anterior insula and the right posterior cingulate cortex during empathy (because of the lack of difference between emotional and neutral faces). To further explore whether any difference between emotional and neutral empathy correlated with narcissism, the contrast 'emotion >neutral' was

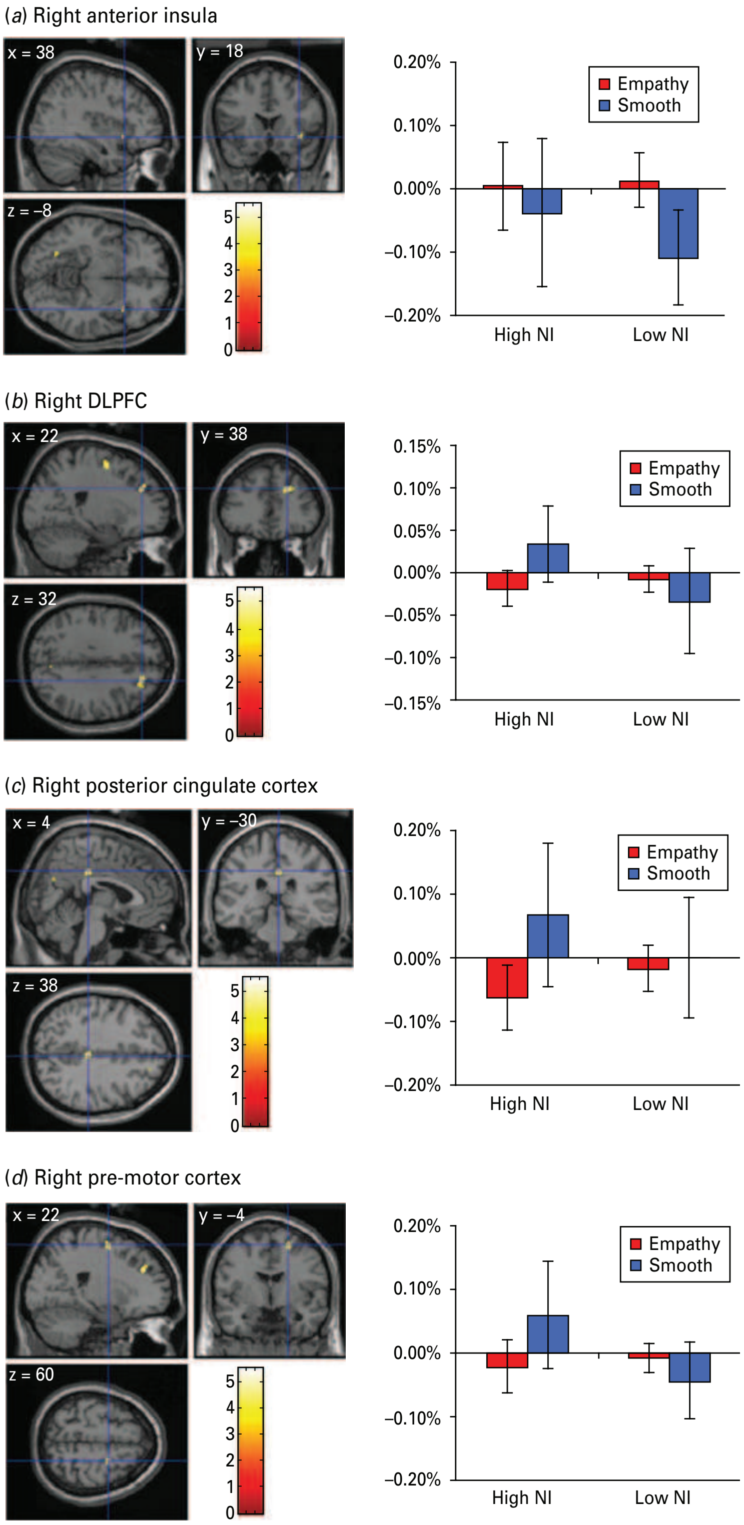


Fig. 1. Signal changes in those regions that differ significantly between high and low narcissistic subjects in empathy *versus* non-empathy: (*a*) right anterior insula; (*b*) right dorsolateral prefrontal cortex; (*c*) right posterior cingulate cortex; (*d*) right premotor cortex. The results are based on a two-sample t test with $p<0.001$ uncorrected, $k>10$; and a small volume correction for the right anterior insula with a 10-mm-radius sphere that showed false discovery rate-corrected p value $=0.017$ and familywise error-corrected p value $=0.033$. Note that high Narcissism Inventory (NI) subjects differ from low NI subjects, especially in the non-empathy condition, showing less task-induced deactivation or even activation.

Table 2. *Brain activation contrast 'empathize with emotional faces > view control faces' for 'low NI > high NI'*

Region	BA	*T* value	MNI coordinates x	y	z	Cluster size	'Empathic reaction > skin color evaluation' (signal change during 6–10 s)	'Anticipation of reward > anticipation of no outcome' (signal change during 6 to 8 s after stimuli onset) Main effect	Interaction with subject group (high/low NI)
Right anterior insula[a]	13/47	4.37	38	18	−8	11	$T(19)=3.316$ $p=0.004$**	$F(1,20)=27.228$ $p<0.001$***	$F(1,20)=0.173$ $p=0.682$
Right dorsolateral prefrontal cortex	9	5.06	22	38	32	93	$T(19)=0.131$ $p=0.897$	$F(1,20)=0.022$ $p=0.885$	$F(1,20)=1.143$ $p=0.298$
Right posterior cingulated cortex	31	5.23	4	−30	38	31	$T(19)=0.912$ $p=0.373$	$F(1,20)=2.240$ $p=0.150$	$F(1,20)=4.005$ $p=0.059$
Right premotor cortex	6	4.77	22	−4	60	23	$T(19)=-0.554$ $p=0.586$	$F(1,20)=2.606$ $p=0.122$	$F(1,20)=0.239$ $p=0.630$

BA, Brodmann area; MNI, Montreal Neurological Institute.

Two-sample *t* test: $p<0.001$ uncorrected; $k>10$.

[a] Small volume correction with region of interest (ROI) in right anterior insula [centered at (38, 22, −1), volume = 216 voxels] defined independently from meta-analysis results of previous functional magnetic resonance imaging (fMRI) studies on empathy (see online Supplement), which showed false discovery rate (FDR)-corrected *p* value = 0.036 and familywise error (FWE)-corrected *p* value = 0.013.

*** $p<0.001$, ** $p<0.01$.

correlated with NI score in a second-level regression analysis, with these subjects who had been subjected to neutral stimuli (see online Supplement). Activations in the right ventromedial prefrontal cortex [r VMPFC, Brodmann area (BA) 10 (18, 60, −4)], the pre-supplementary motor area [pre-SMA, BA 8 (8, 14, 58)] and the left lateral prefrontal cortex [l LPFC, BA 10 (−32, 52, 4)] were found to correlate negatively with NI score (Supplementary Fig. 2*b*), suggesting a difference in general emotion processing associated with narcissism. However, the right anterior insula and the right posterior cingulate cortex were not found to be involved in this comparison between emotional and neutral faces.

To exclude unspecific task-related effects of general evaluation (Singer, 2006; Kriegeskorte *et al*. 2009) in the empathy task, we applied the coordinates differing between the high and low narcissism groups to a data set, where we directly compared empathy perception/judgment with perception/judgment of skin color (see online Supplement). The comparison of empathy *versus* skin color judgment yielded a significant difference for the coordinates from the right anterior insula whereas the other regions from the high *versus* low narcissism comparison did not show any significant difference (Table 2). This supports the assumption that the difference in the right anterior insula between high and low narcissism is due to empathy rather than to some non-specific task-related effect of a general evaluation function.

We also tested for possible rewarding effects of empathy. Singer *et al*. (2004, 2006) showed a strong neural and psychological overlap between reward and empathy. It is not clear whether the differences between high and low narcissistic groups regarding their empathic response are due to empathy or are related to the rewarding effects of empathic stimuli (i.e. that emotional faces have a significant meaning to the perceiver). Thus we included a reward task to control for possible effects of reward related to the empathic stimuli. Using a reward task (Knutson *et al*. 2001*a*, *b*) enabled us to test whether the neuronal activity differences between the two groups during empathy might be due to reward effects in the same regions. Using the coordinates from the group difference during empathy to the data acquired during reward did not show any significant interaction between the two groups with regard to the anticipation of reward (compared to the anticipation of no outcome) (Table 2).

Finally, to further confirm the association between narcissism and the empathy-related activity in the right anterior insula, we examined whether the medium NI group showed intermediate findings in the right anterior insula (see online Supplement). The percentage of signal changes of all three groups of subjects, extracted from the ROI of the right anterior insula defined independently by the meta-analysis results of previous fMRI studies on empathy, were found to correlate negatively with the NI scores [$r(30) = -0.370$, $p < 0.05$] while controlling for the impact of age and intelligence [partial correlation $r(30) = -0.374$, $p < 0.05$] (Supplementary Fig. 3).

Taken together, our neuroimaging results demonstrate neuronal differences between high and low narcissism subjects, most notably in the right anterior insula, that cannot be traced back to confounding functions such as emotion processing, face perception, general evaluation and reward.

Discussion

We have investigated the relationship between emotions, empathy and narcissism in an exploratory study on a non-clinical sample that combined psychological, behavioral and neuronal measures. Psychologically, highly narcissistic subjects yielded higher scores of alexithymia, general psychopathology and depression. By contrast, no significant psychological differences were observed with regard to empathy and *post-hoc* ratings for emotional reaction to the stimuli. Neuronally, highly narcissistic subjects showed differences predominantly in the right anterior insula that, as they were carefully controlled for, could not be traced back to other possible confounding functions. Taken together, these results indicate psychological and neuronal group differences in highly narcissistic subjects when compared to low narcissistic subjects. This may contribute to a better understand of the physiological basis of narcissism in general.

On a continuum from non-clinical narcissism to narcissistic personality disorder (Watson *et al*. 2002), our study sample is comparable to that of a control sample of healthy volunteers described by Denecke & Hilgenstock (1989). The mean narcissism total score in our sample was 22.5, compared to 23.0 in the above-mentioned group. Our high narcissism group with a mean score of 26.2 is located in the middle between the controls and a group of patients with narcissistic personality disorder (mean = 29.2) (Denecke & Hilgenstock, 1989). Daig *et al*. (2010) reported a mean score of 27.8 for a group of psychosomatic outpatients. Thus, our low narcissism group can be regarded as narcissistic below average, whereas the high narcissism group is above average but still below clinical severity.

Our high narcissism subjects exhibited more alexithymia, which parallels the clinically often described emotional abnormalities in narcissistic subjects (Kohut, 1971; Kernberg, 1975; Dimaggio *et al*. 2006). The co-occurrence and interdependency of narcissistic traits and alexithymia in eating-disordered patients have been reported by Lawson *et al*. (2008). The correlation of narcissism and lack of empathy is even more established, as documented in clinical–conceptual accounts (Kohut, 1971; Kernberg, 1975), experimental investigation (Watson *et al*. 1984) and the DSM-IV (APA, 1994) diagnostic criteria for narcissistic personality disorder.

Our main neural finding concerns the observation of neuronal differences in the right anterior insula in particular and also in other regions such as the DLPFC, the posterior cingulate cortex, and the lateral premotor cortex. The right anterior insula has indeed been typically associated with empathy (Singer *et al*. 2004, 2006; Singer & Lamm, 2009; Lamm & Singer, 2010). The fact that high narcissistic subjects show differences in this region when compared to low narcissistic subjects further supports the assumption of changes in empathy in narcissism. However, the exact mechanisms associated with insula activity in empathy remain unclear. One prominent theory assumes simulation of the other's affective state to be crucial in empathy (Singer & Lamm, 2009; Lamm & Singer, 2010); but whether the differential insula activity in our groups may be related to different degrees of simulation of the other's affective state remains unclear.

Unlike in the other regions (i.e. the DLPFC, lateral premotor cortex and posterior cingulate cortex), neuronal activity changes in the right anterior insula in high narcissism subjects could not be traced back to related functions such as general evaluation, face perception or reward. This indicates that the right anterior insula must play a specific role in the constitution of empathy in people with narcissistic personality traits, which is in line with the characterization of the psychological function of the right anterior insula by Craig (2009). Based on its connectivity pattern, Craig assumes that the right anterior insula, as distinguished from the left one, is involved in representing the self, more specifically the bodily self or the 'material me' as he calls it. Because our high narcissism subjects show increased preoccupation with the self, as indicated by the narcissism scores, their altered activity in the insula may thus result from increased focus on the own self as, for instance, during mind-wandering in the non-empathic control condition. This is compatible with recent imaging studies showing recruitment of the right anterior insula during tasks focusing on the self (Enzi *et al*. 2009; Modinos *et al*. 2009).

We should also consider the fact that the main signal changes in the insula and the other regions stem from differences, for instance lower deactivation, in the non-empathic control condition rather than the empathy condition itself. However, the origin of the difference in the non-empathic control condition remains unclear. Because of the absence of an identifiable external stimulus in the empathic control condition, it is possible that activity differences in this condition might be related to differences in internal activity rather than externally stimulus-induced activity. This interpretation should, however, be considered with caution because our design did not include a true resting-state period. Moreover, the psychological correlates of the resting state in general and in narcissism remain unclear. Daydreaming (Singer, 1966) and mind-wandering (Smallwood & Schooler, 2006) may be central; how these are involved in narcissism is a subject for future investigation. To lend further support to our speculative hypothesis, future investigation of the neuronal relationship between empathy, mind-wandering and self-related fantasies in narcissism is necessary.

Questions may also be raised concerning how our psychological and neuronal findings relate to the concept of narcissism as presupposed in psychoanalytic theory and clinical psychiatry. The concept of narcissism went through transference from pure intra-psychic nature of the self (and the ego as originally postulated by Freud, 1905, 1910) into a more relational or intersubjective 'self-objects' framework (as reviewed in Kohut, 1971, 1977, 1984; Gehrie, 2009; Hartmann, 2009) that includes the other (non-self) as a constitutive element (Kohut, 1984; Fosshage, 2009), which, on the one hand, reflects its character as a continuum from 'mature narcissism' to 'archaic or immature narcissism', such as the narcissistic personality disorder in DSM-IV (APA, 1994; see also Russ *et al*. 2008), and, on the other hand, is in accordance with recent accounts that link narcissism to emotions and empathy (Hartmann, 2009). Both our psychological and neural findings lend further support to such a relational view of narcissism, which may consequently be closely related to changes in intersubjective relationships to other persons. The apparent difficulty in accessing the own emotions, as suggested by our finding of increased alexithymia, may lead to problems in simulating the other's affective states as our neural findings may be interpreted. The lack of accessing the other's emotions may in turn induce a shift from inter- to intra-subjective relationships, where the subject aim to relate to themselves through their narcissistic fantasies. This, however, is somewhat speculative at this point and requires further more detailed empirical investigation.

Finally, our findings also lend support to the concept of narcissism as a continuum between healthy and pathological forms. Narcissism is regarded as a trait of the healthy personality (Emmons, 1984) and a prominent behavior pattern in western cultures (Lasch, 1979). Our findings indicate that, even within a healthy sample, different degrees of narcissism as a personality trait seem to go along with neural and psychological differences. What may be of interest in the future is to investigate pathological forms of narcissism. Based on our findings reported here, we may expect even higher degrees of alexithymia and lower deactivation in, for instance, the right anterior insula during empathy. This remains to be investigated in the future.

Note

Supplementary material accompanies this paper on the Journal's website (http://journals.cambridge.org/psm).

Acknowledgments

We thank the staff from the Department of Neurology at the University of Magdeburg for their generous support. This study was supported financially by grants from the Hope of Depression Research Foundation (HDRF), the German Research Foundation (SFB-779, A6), the CIHR and the EJLB-Michael Smith Foundation to G.N.

Declaration of Interest

None.

References

APA (1980). *Diagnostic and Statistical Manual of Mental Disorders*, 3rd edn. American Psychiatric Press: Washington, DC.

APA (1994). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. American Psychiatric Press: Washington, DC.

Bach M, Bach D, de Zwaan M, Serim M, Bohmer F (1996). Validation of the German version of the 20-item Toronto Alexithymia Scale in normal persons and psychiatric patients [in German]. *Psychotherapie, Psychosomatik, Medizinische Psychologie* **46**, 23–8.

Craig AD (2009). How do you feel – now? The anterior insula and human awareness. *Nature Reviews Neuroscience* **10**, 59–70.

Daig I, Burkert S, Fischer F, Kienast T, Klapp BF, Fliege H (2010). Development and factorial validation of a short version of the Narcissism Inventory (NI-20). *Psychopathology* **43**, 150–158.

Decety J, Lamm C (2006). Human empathy through the lens of social neuroscience. *Scientific World Journal* **6**, 1146–1163.

Denecke FW, Hilgenstock B (1989). *The Narcissism Inventory* [in German]. Hans Huber: Bern.

Dimaggio G, Nicolò G, Popolo R, Semerari A, Carcione A (2006). Self-regulatory dysfunctions in personality disorders: the role of poor self-monitoring and mindreading. *Applied Psychology: An International Review* **55**, 397–407.

Emmons RA (1984). Factor analysis and construct validity of the Narcissistic Personality Inventory. *Journal of Personality Assessment* **48**, 291–300.

Enzi B, de Greck M, Prosch U, Tempelmann C, Northoff G (2009). Is our self nothing but reward? Neuronal overlap and distinction between reward and personal relevance and its relation to human personality. *PLoS ONE* **4**, e8429.

Fosshage JL (2009). Some key features in the evolution of self psychology and psychoanalysis. *Annals of the New York Academy of Sciences* **1159**, 1–18.

Franke G (1995). *SCL-90-R. Die Symptom-Checkliste von Derogatis – Deutsche Version* [SCL-90-R. The Symptom Check-List from Derogatis – German version]. Beltz: Goettingen.

Freud S (1905). *Three Essays on the Theory of Sexuality*, Vol. 7. Hogarth: London.

Freud S (1910). Leonardo da Vinci and a memory of his childhood. In *The Standard Edition of the Complete Psychological Works of Sigmund Freud* (ed. S. Freud), pp. 59–137. Hogarth Press: London.

Gehrie MJ (2009). The evolution of the psychology of the self: toward a mature narcissism. *Annals of the New York Academy of Sciences* **1159**, 31–50.

Hartmann HP (2009). Psychoanalytic self psychology and its conceptual development in light of developmental psychology, attachment theory, and neuroscience. *Annals of the New York Academy of Science* **1159**, 86–105.

Hautzinger M, Bailer M, Worall H, Keller F (1994). *The Beck Depression Inventory (BDI)* [in German]. Huber: Bern.

Hooker CI, Verosky SC, Germine LT, Knight RT, D'Esposito M (2008). Mentalizing about emotion and its relationship to empathy. *Social Cognitive and Affective Neuroscience* **3**, 204–217.

Horn W (1983). *Leistungsprüfsystem L-P-S, Handanweisung* [Achievement Measurement System, User's Guide]. Hogrefe: Göttingen.

Kernberg OF (1975). *Borderline Conditions and Pathological Narcissism*. Jason Aronson: New York.

Knutson B, Adams CM, Fong GW, Hommer D (2001*a*). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience* **21**, RC159.

Knutson B, Fong GW, Adams CM, Varner JL, Hommer D (2001*b*). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport* **12**, 3683–3687.

Kohut H (1971). *The Analysis of the Self*. International Universities Press: New York.

Kohut H (1977). *The Restoration of the Self*. International Universities Press: New York.

Kohut H (1984). *How Does Analysis Cure?* The University of Chicago Press: Chicago.

Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience* **12**, 535–540.

Lamm C, Singer T (2010). The role of anterior insular cortex in social emotions. *Brain Structure and Function* **214**, 579–591.

Lasch C (1979). *The Culture of Narcissism: American Life in an Age of Diminishing Expectations*. WW Norton: New York.

Lawson R, Waller G, Sines J, Meyer C (2008). Emotional awareness among eating-disordered patients: the role of narcissistic traits. *European Eating Disorders Review* **16**, 44–48.

Lehrl S, Triebig G, Fischer B (1995). Multiple choice vocabulary test MWT as a valid and short test to estimate premorbid intelligence. *Acta Neurologica Scandinavica* **91**, 335–345.

Modinos G, Ormel J, Aleman A (2009). Activation of anterior insula during self-reflection. *PLoS ONE* **4**, e4618.

Paulus C (2009). *Saarbruecker Persoenlichkeits-Fragebogen zur Empathie (SPF) [Saarbruecker Personality Questionnaires for Empathy (SPF)]. Based on the Interpersonal Reactivity Index (IRI). V5.5.* University of Saarbruecken: Saarbruecken.

Preston SD, de Waal FB (2002). Empathy: its ultimate and proximate bases. *Behavioral and Brain Sciences* **25**, 1–20; discussion 20–71.

Ritter K, Lammers CH (2007). Narcissism – variable of personality and personality disorder [in German]. *Psychotherapie, Psychosomatik, Medizinische Psychologie* **57**, 53–60.

Russ E, Shedler J, Bradley R, Westen D (2008). Refining the construct of narcissistic personality disorder: diagnostic criteria and subtypes. *American Journal of Psychiatry* **165**, 1473–1481.

Singer JL (1966). *Daydreaming: An Introduction to the Experimental Study of Inner Experience*. Random House: New York.

Singer T (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neuroscience and Biobehavioral Reviews* **30**, 855–863.

Singer T, Lamm C (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences* **1156**, 81–96.

Singer T, Seymour B, O'Doherty J, Kaube H, Dolan RJ, Frith CD (2004). Empathy for pain involves the affective but not sensory components of pain. *Science* **303**, 1157–1162.

Singer T, Seymour B, O'Doherty JP, Stephan KE, Dolan RJ, Frith CD (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* **439**, 466–469.

Smallwood J, Schooler JW (2006). The restless mind. *Psychological Bulletin* **132**, 946–958.

Watson PJ, Grisham SO, Trotter MV, Biderman MD (1984). Narcissism and empathy: validity evidence for the Narcissistic Personality Inventory. *Journal of Personality Assessment* **48**, 301–305.

Watson PJ, Sawrie SM, Greene RL, Arredondo R (2002). Narcissism and depression: MMPI-2 evidence for the continuum hypothesis in clinical samples. *Journal of Personality Assessment* **79**, 85–109.

ONLINE DATA SUPPLEMENT for Fan *et al.*:

The narcissistic self and its psychological and neural correlates. An exploratory fMRI study

SUPPLEMENTAL METHODS

fMRI paradigm

Empathy blocks: The empathy block started with a short finger tapping task, which makes it possible to identify the brain region associated with hand movement of each subject in the primary motor cortex. Through a short introduction of 6 s presented on the screen, subjects were told to repeatedly press a button with their index finger 32 times, at the speed indicated by the change of brightness of the fixation cross, and at a frequency of approximately 2 button presses per second. The empathy task started immediately after the finger tapping task with again the presentation of a 6 s instruction. Subjects were asked to empathize with the presented emotional face, which was expressed by the instruction phrase “please try to share the emotional state of the shown person”. A total number of 40 empathy trials were presented in a randomized order. Every empathy trial began with the 5 s lasting sole display of an emotional face or the presentation of a control stimulus. Immediately after the presentation of the emotional face a subjective evaluation task was presented and subjects were asked to rate their ability to empathize by moving a virtual bar of a visual analogue scale. Prior to the following empathy trial a short inter trial interval was presented, which lasted for 2 or 3 s.

Stimuli: The emotional face stimuli were taken from the ‘Japanese and Caucasian Facial Expressions of Emotion (JACFEE) and Neutral Faces (JACNeuF)’ battery provided by Matsumoto and Ekman (1988). A number of 8 different faces were shown in the emotional expression of angry, disgusted, happy, and neutral, resulting in a total number of 32 different faces. Each emotion was expressed by 4 Caucasian and 4 Japanese actors (male:female = 1:1). Considered automatic empathic responses might be induced by the mere presentation of emotional faces, even without a specific instruction to empathize (Yamada and Decety, 2009), our control condition was presenting smoothed pictures with unrecognizable contents, which was transformed from neutral face using a smoothing function (Gerlach *et al.*, 2002). The stimuli were used repeatedly over three empathy blocks. The presentation of different emotions enabled the subject to empathize with a broad range of emotions. Sixteen out of 34 subjects erroneously watched contempt instead of neutral pictures, though none the subjects were aware of this, they were not included into the analysis concerning neutral faces. To acquire additional information of how the emotional stimuli were subjectively perceived by the subjects, an evaluation of all emotional stimuli was performed after the fMRI session. All subjects watched the stimuli and gave subjective ratings of their impression of (i) self-relatedness, (ii) emotional intensity, (iii) valence, (iv) cognitive empathy, (v) emotional empathy, and (vi) sympathy concerning the shown person.

Reward blocks: The reward task was a modified version of the “Monetary Incentive Delay Task” (MID) as introduced by Knutson and colleagues (Knutson *et al.*, 2001a,b). Every reward block started with the 6 s presentation of an instruction. In each reward block a total number of 60 trials was presented, which included 20 trials with reward anticipation, 20 trials with punishment anticipation and 20 trials with the anticipation of no outcome. Every trial required a button press of the subject with the index finger of their right hand within a certain time during the presentation of the target image (a black square in the center of the screen). The length of this time period was determined in accordance with the average reaction time obtained in the pre-scan trial run, allowing the difficulty of the task to be modulated according to the individual's ability, and varied between 0.2 and 0.5 s. Furthermore, we applied an adapting algorithm, to ensure that in approximately 60% of all trials the required response was successful. Prior to this target image being displayed, a symbol indicating what the possible outcomes of the task would be – either reward, punishment, or no-outcome – was shown for 0.3 s, followed by a 2.25-2.75 s anticipation period. The trial type indicator took the form of a black circle with a small white circle within it at one of the cardinal points. Each position represented a different trial type (e.g., a circle in the 'north' position would represent a reward trial). During the anticipation period a light gray colored cross was displayed in the center of the screen.

In reward trials, completing the task successfully resulted in the subject winning € 1, whilst failure meant that they would neither win nor lose anything. In punishment trials, the subject was about to lose € 1, which could be prevented by a response within the required time period.
Finally, in no-outcome trials no money was either won or lost, regardless of whether the subject responded within the required time period or not. Subjects were, however, instructed to still respond to the cue as quickly as possible. An equal number of reward, punishment, and no-outcome trials were displayed in each of the three reward/punishment runs in a pseudo-random order, giving a total of 60 instances of each trial type. Each trial was followed by a feedback stage during which the subject was informed of the outcome. The amount of money won or lost in the preceding trial was displayed, along with the running total for their winnings, for a period of 1.65 s. Trials were separated by a 4-5 s inter trial interval. The total amount of money won during the whole experiment was provided to the subjects as reimbursement for their participation in the experiment.

Region of interest (ROI) analysis

Our SPM analysis was complemented by the ROI analysis of BOLD signal changes. Using the Marseille Region of Interest Toolbox software package (MarsBaR 0.42, http://marsbar.sourceforge.net/), we applied a sphere with a radius of 5 mm centered at the peak voxels of our ROIs, and extracted the raw fMRI data. The extracted data then underwent a linear interpolation, onset adapting and normalizing procedure using the software package Practical Data Extraction and Reporting Language (PERL; www.perl.org) to account for intersubject differences. Our regional signal changes were further corrected referring to the signal changes of the preceding fixation cross period in order to exclude possible baseline shifts on subsequent stimuli-induced signal changes. The mean normalized regional signal changes (4 to 10 s after stimulus onset) were finally calculated and entered statistical analysis of interested conditions.

Confirmatory fMRI study

To exclude the confounding effect of unspecific general evaluation which was possibly involved in the comparison between empathizing emotional faces and perceiving smoothed control, the brain regions differing between the high and low narcissism groups were selected as ROIs and applied to the data of a second fMRI study, with a comparable empathy task. The main difference concerns the control task. In the confirmatory fMRI study we decided to use a control task which allowed us to compare empathy with a general evaluation of facial stimuli (skin color). The control task of this study was hence similar to the empathy task and contained the evaluation of the skin color of the emotional face stimuli.

Subjects: We investigated 20 healthy Chinese subjects (11 female, 9 male, mean age: 23, range: 21-26). After a detailed explanation of the study design and any potential risks, all subjects gave their written informed consent. All of the subjects were Chinese students. The study was approved by the institutional review board of the University of Peking, China.

Experimental Design: The fMRI experiment was divided into 7 blocks of 312 s duration each. Prior to entering the scanner each subject read a detailed information of the paradigm and completed a couple of trial runs in order to familiarize fully with them. While lying in the scanner, the stimuli were displayed using the ‘Presentation’ software package (Neurobehavioral Systems, Albany, CA), and were projected onto a matt screen via an LCD projector, visible through a mirror mounted on the headcoil. Every block started with 10 s pause to control for epi-saturation effects. A total number of 24 trials were presented in a randomized order. 12 trials were empathy trials, 12 trials were skin color evaluation trials. The actual task, empathy or skin color evaluation, was indexed by the 0.5 s presentation of cue which consisted of a black circle with a small white circle within it at two positions. The white circle in the ‘North’ position indexed an empathy trial, the white circle in the ‘South’ position cued the skin color evaluation trials. The cue was followed by a blank screen for 1 s. Subsequently the emotional face picture was displayed for 4 s. Subjects were instructed to feel inside the depicted emotional face during empathy trials or to concentrate on the skin color of the presented face during skin color evaluation trials. The face picture was followed by the presentation of a visual analogue scale. By virtually moving a red bar with left and right button presses the subjects were instructed to give an intra scanner rating of how good they felt able to empathize with the emotional face, respectively how dark or bright they rate the skin color of the emotional face. With a third button subjects had to confirm their rating. After confirmation the color of the bar turned into gray. The duration of the evaluation phase was 3.5 s. Prior to the next trial an inter trial interval was included lasting 1.2, 1.4, 1.6 or 1.8 s. After every 6 trials a baseline trial was included, which consisted of the mere presentation of the fixation cross, lasting for 6 or 7 s.

Stimuli: Our emotional face stimuli consisted of 12 emotional faces. Four faces, 2 female and 2 male, containing neutral emotional expressions were taken from the “Japanese and Caucasian Facial Expressions of Emotion (JACFEE) and Neutral Faces (JACNeuF)” battery provided by Matsumoto and Ekman (1988). Eight additional faces of Chinese subjects were photographed by our own group. The pictures were taken in front of comparable backgrounds and under comparable conditions to match them as close as possible to the 4 pictures taken from the JACNeuF battery. These additional 8 pictures contained the emotions angry and neutral and were taken from 4 female and 4 male Chinese students. Every stimulus was presented twice during each block, once during empathy, once during skin color evaluation. During the whole experiment every stimulus was hence presented 14 times.

fMRI data acquisition: The study was conducted using a GE 3 Tesla Magnetic Resonance Imaging Scanner (24 slices parallel to the AC-PC plane, slice thickness 5 mm, TR 2000 ms, TE 30 ms, flip angle $\alpha = 90°$, 64x64 voxel per slice with 3.75x3.75x5 mm). Functional data were acquired in seven scanning sessions containing 156 volumes per session for each subject.

Data analysis: The statistical analysis of the fMRI data was performed using SPM2 and Matlab 6.5.1 (The Mathworks Inc., Natick, MA, USA). fMRI data were slice time corrected with regard to the first slice acquired and movement corrected by realignment to the first volume. The functional images were normalized to a standard brain and resampled to a 2x2x2 mm voxel size. Smoothing was performed using a 6x6x6 mm FWHM-Kernel. We built sphere-shaped ROIs (5 mm radium) according to the most significant coordinates resulting from the group comparison high vs. low narcissism in the main fMRI study. We used these ROIs to extract raw data from confirmatory fMRI study. We then calculated regional signal changes for the empathy and the skin color evaluation condition. As described above, we used the software package PERL to apply a linear interpolation, onset adapting, normalizing and baseline correction. The resulting mean normalized regional signal changes (4 to 10 s after stimulus onset) for empathy and skin color evaluation were compared using a paired t-test.

Confirmatory meta-analysis of 40 previous empathy studies

The results of this meta-analysis of 40 previous empathy studies (Akitsuki and Decety, 2009, Benuzzi *et al.*, 2008, Blakemore *et al.*, 2005, Botvinick *et al.*, 2005, Carr *et al.*, 2003, Chakrabarti *et al.*, 2006, Cheng *et al.*, 2007, de Gelder *et al.*, 2004, Decety *et al.*, 2009, Farrow *et al.*, 2001, Grosbras and Paus, 2006, Gu and Han, 2007a,b, Hennenlotter *et al.*, 2005, Hynes *et al.*, 2006, Jabbi *et al.*, 2007, Jackson *et al.*, 2006a, Jackson *et al.*, 2005, Keysers *et al.*, 2004, Kim *et al.*, 2009, Lamm *et al.*, 2007, Lamm and Decety, 2008, Lawrence *et al.*, 2006, Leslie *et al.*, 2004, Moriguchi *et al.*, 2007, Morrison and Downing, 2007, Morrison *et al.*, 2004, Morrison *et al.*, 2007, Nummenmaa *et al.*, 2008, Olsson *et al.*, 2007, Prehn-Kristensen *et al.*, 2009, Saarela *et al.*, 2007, Schulte-Ruther *et al.*, 2007, Seitz *et al.*, 2008, Simon *et al.*, 2006, Singer *et al.*, 2004, Singer *et al.*, 2006, Vollm *et al.*, 2006, Warren *et al.*, 2006, Wicker *et al.*, 2003) served as an independent validation of the activation of the whole group analysis 'empathy > non-empathy'.

Literature search: The relevant empathy papers were collected through a step-wise procedure. First we performed a standard search in two databases, PubMed (www.pubmed.gov) and ISI Web of Science (https://apps.isiknowledge.com), using keywords ['empathy' OR 'empathic' OR 'emotion contagion' OR 'affective theory of mind' OR 'affective mentalizing'] combined with ['fMRI' OR 'magnetic resonance imaging']. Second, we reviewed the reference lists of the relevant articles obtained in the first step, and used the 'related article' function of the PubMed database to identify additional papers. Finally, the reference lists of several review articles were inspected for further relevant studies (de Vignemont and Singer, 2006, Decety and Jackson, 2004,

Decety and Lamm, 2006, Eslinger, 1998, Jackson *et al.*, 2006b, Seitz *et al.*, 2006).
Studies were considered empathy-relevant if their paradigms met the following criteria:

1) The perception of others' sensory or affective state shows activation similar to the experience, imitation or generation of the similar state in oneself.

2) The perception of others' sensory or affective state shows activation correlating to the disposition-measurement of one's empathy.

3) The task required the subjects to empathize with other individuals and make judgments about others' feelings.

Other inclusion criteria were applied:

1) Only studies measuring healthy adults were included. Data of the healthy control group in patient studies were included if detailed statistical analysis was performed.

2) Only studies measuring neural activity in the whole brain were included; studies reporting only selected regions of interest were excluded.

3) Presentation of results has been limited to regional activation changes (as revealed by task comparison or image subtraction method, parametric designs or brain-behavioral comparison). Data on changes in functional or effective connectivity have been excluded.

4) Only activation data were included, whereas deactivation data were not considered.

5) For conversion between the two coordinate systems, the Montreal Neurological Institute (MNI) standard brains and the Talairach space, we used the algorithm developed by Brett *et al.* (2001)

Multilevel kernel density analysis (MKDA): Following the literature search, we included 40 studies, yielding 50 contrasts (664 peak coordinates in total). The activation consistency across these studies was evaluated using a quantitative meta-analysis method MKDA (Wager *et al.*, 2007, 2009). First, an activation-indicator map was calculated for each particular statistic contrast, by convolving the peak coordinates with a 10 mm spherical kernel. The indicator map was threshold at a maximum value of 1 so that multiple nearby peaks were not counted as multiple activations. Second, the weighted average of all indicator maps provided a summary map, where the weights were related to the sample size of a study and the multiple comparison correction it used. Finally, the statistical threshold was established for the summary map using the Monte Carlo procedure, which scrambled foci as clusters and produced a random map. The Monte Carlo procedure was performed for 5000 iterations and the result was reported at a threshold of FWE corrected $p < 0.05$.

To compare the activation of the whole group analysis and the results of the meta analysis, we superimposed both activation images on a standard brain template (MNI) using the software 'MRIcron' (www.sph.sc.edu/comd/rorden/MRicron/) (see supplementary Figure 1).

References

Akitsuki, Y. & Decety, J. (2009). Social context and perceived agency affects empathy for pain: an event-related fMRI investigation. *Neuroimage* **47**, 722-34.

Benuzzi, F., Lui, F., Duzzi, D., Nichelli, P. F. & Porro, C. A. (2008). Does it look painful or disgusting? Ask your parietal and cingulate cortex. *Journal of Neuroscience* **28**, 923-31.

Blakemore, S. J., Bristow, D., Bird, G., Frith, C. & Ward, J. (2005). Somatosensory activations during the observation of touch and a case of vision-touch synaesthesia. *Brain* **128**, 1571-83.

Botvinick, M., Jha, A. P., Bylsma, L. M., Fabian, S. A., Solomon, P. E. & Prkachin, K. M. (2005). Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. *Neuroimage* **25**, 312-9.

Brett, M., Leff, A. P., Rorden, C. & Ashburner, J. (2001). Spatial normalization of brain images with focal lesions using cost function masking. *Neuroimage* **14**, 486-500.

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 USA* **100**, 5497-502.

Chakrabarti, B., Bullmore, E. & Baron-Cohen, S. (2006). Empathizing with basic emotions: common and discrete neural substrates. *Social Neuroscience* **1**, 364-84.

Cheng, Y., Lin, C. P., Liu, H. L., Hsu, Y. Y., Lim, K. E., Hung, D. & Decety, J. (2007). Expertise modulates the perception of pain in others. *Current Biology* **17**, 1708-13.

de Gelder, B., Snyder, J., Greve, D., Gerard, G. & Hadjikhani, N. (2004). Fear fosters flight: a mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proceedings of the National Academy of Sciences USA* **101**, 16701-6.

de Vignemont, F. & Singer, T. (2006). The empathic brain: how, when and why? *Trends in Cognitive Sciences* **10**, 435-41.

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

Decety, J. & Lamm, C. (2006). Human empathy through the lens of social neuroscience. *The Scientific World Journal* **6**, 1146-63.

Decety, J., Michalska, K. J., Akitsuki, Y. & Lahey, B. B. (2009). Atypical empathic responses in adolescents with aggressive conduct disorder: a functional MRI investigation. *Biological Psychology* **80**, 203-11.

Eslinger, P. J. (1998). Neurological and neuropsychological bases of empathy. *European Neurology* **39**, 193-9.

Farrow, T. F., Zheng, Y., Wilkinson, I. D., Spence, S. A., Deakin, J. F., Tarrier, N., Griffiths, P. D. & Woodruff, P. W. (2001). Investigating the functional anatomy of empathy and forgiveness. *Neuroreport* **12**, 2433-8.

Gerlach, C., Aaside, C. T., Humphreys, G. W., Gade, A., Paulson, O. B. & Law, I. (2002). Brain activity related to integrative processes in visual object recognition: bottom-up integration and the modulation influence of stored knowledge. *Neuropsychologia* **40**, 1254-1267.

Grosbras, M. H. & Paus, T. (2006). Brain networks involved in viewing angry hands or faces. *Cerebral Cortex* **16**, 1087-96.

Gu, X. & Han, S. (2007a). Attention and reality constraints on the neural processes of empathy for pain. *Neuroimage* **36**, 256-67.

Gu, X. & Han, S. (2007b). Neural substrates underlying evaluation of pain in actions depicted in words. *Behavioural Brain Research* **181**, 218-23.

Hennenlotter, A., Schroeder, U., Erhard, P., Castrop, F., Haslinger, B., Stoecker, D., Lange, K. W. & Ceballos-Baumann, A. O. (2005). A common neural basis for receptive and expressive communication of pleasant facial affect. *Neuroimage* **26**, 581-91.

Hynes, C. A., Baird, A. A. & Grafton, S. T. (2006). Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. *Neuropsychologia* **44**, 374-83.

Jabbi, M., Swart, M. & Keysers, C. (2007). Empathy for positive and negative emotions in the gustatory cortex. *Neuroimage* **34**, 1744-53.

Jackson, P. L., Brunet, E., Meltzoff, A. N. & Decety, J. (2006a). Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia* **44**, 752-61.

Jackson, P. L., Meltzoff, A. N. & Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *Neuroimage* **24**, 771-9.

Jackson, P. L., Rainville, P. & Decety, J. (2006b). To what extent do we share the pain of others? Insight from the neural bases of pain empathy. *Pain* **125**, 5-9.

Keysers, C., Wicker, B., Gazzola, V., Anton, J. L., Fogassi, L. & Gallese, V. (2004). A touching sight: SII/PV activation during the observation and experience of touch. *Neuron* **42**, 335-46.

Kim, J. W., Kim, S. E., Kim, J. J., Jeong, B., Park, C. H., Son, A. R., Song, J. E. & Ki, S. W. (2009). Compassionate attitude towards others' suffering activates the mesolimbic neural system. *Neuropsychologia* **47**, 2073-81.

Knutson, B., Adams, C. M., Fong, G. W. & Hommer, D. (2001a). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience* **21**, RC159.

Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L. & Hommer, D. (2001b). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport* **12**, 3683-7.

Lamm, C., Batson, C. D. & Decety, J. (2007). The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience* **19**, 42-58.

Lamm, C. & Decety, J. (2008). Is the extrastriate body area (EBA) sensitive to the perception of pain in others? *Cerebral Cortex* **18**, 2369-73.

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**, 1173-84.

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**, 601-7.

Matsumoto, D. & Ekman, P. (1988). Japanese and Caucasian Facial Expressions of Emotion and Natral Faces (JACFEE and JACNeuF). Available from Human Interaction Laboratory, University of California, San Francisco, 401 Parnassus Avenue, San Francisco, CA, 94143.

Moriguchi, Y., Decety, J., Ohnishi, T., Maeda, M., Mori, T., Nemoto, K., Matsuda, H. & Komaki, G. (2007). Empathy and judging other's pain: an fMRI study of alexithymia. *Cerebral Cortex* **17**, 2223-34.

Morrison, I. & Downing, P. E. (2007). Organization of felt and seen pain responses in anterior cingulate cortex. *Neuroimage* **37**, 642-51.

Morrison, I., Lloyd, D., di Pellegrino, G. & Roberts, N. (2004). Vicarious responses to pain in anterior cingulate cortex: is empathy a multisensory issue? *Cognitive, Affective & Behavioral Neuroscience* **4**, 270-8.

Morrison, I., Peelen, M. V. & Downing, P. E. (2007). The sight of others' pain modulates motor processing in human cingulate cortex. *Cerebral Cortex* **17**, 2214-22.

Nummenmaa, L., Hirvonen, J., Parkkola, R. & Hietanen, J. K. (2008). Is emotional contagion special? An fMRI study on neural systems for affective and cognitive empathy. *Neuroimage* **43**, 571-80.

Olsson, A., Nearing, K. I. & Phelps, E. A. (2007). Learning fears by observing others: the neural systems of social fear transmission. *Social Cognitive and Affective Neuroscience* **2**, 3-11.

Prehn-Kristensen, A., Wiesner, C., Bergmann, T. O., Wolff, S., Jansen, O., Mehdorn, H. M., Ferstl, R. & Pause, B. M. (2009). Induction of empathy by the smell of anxiety. *PLoS one* **4**, e5987.

Saarela, M. V., Hlushchuk, Y., Williams, A. C., Schurmann, M., Kalso, E. & Hari, R. (2007). The compassionate brain: humans detect intensity of pain from another's face. *Cerebral Cortex* **17**, 230-7.

Schulte-Ruther, M., Markowitsch, H. J., Fink, G. R. & Piefke, M. (2007). Mirror neuron and theory of mind mechanisms involved in face-to-face interactions: a functional magnetic resonance imaging approach to empathy. *Journal of Cognitive Neuroscience* **19**, 1354-72.

Seitz, R. J., Nickel, J. & Azari, N. P. (2006). Functional modularity of the medial prefrontal cortex: involvement in human empathy. *Neuropsychology* **20**, 743-51.

Seitz, R. J., Schafer, R., Scherfeld, D., Friederichs, S., Popp, K., Wittsack, H. J., Azari, N. P. & Franz, M. (2008). Valuating other people's emotional face expression: a combined functional magnetic resonance imaging and electroencephalography study. *Neuroscience* **152**, 713-22.

Simon, D., Craig, K. D., Miltner, W. H. & Rainville, P. (2006). Brain responses to dynamic facial expressions of pain. *Pain* **126**, 309-18.

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**, 1157-62.

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J. & Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* **439**, 466-9.

Vollm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., Deakin, J. F. & Elliott, R. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *Neuroimage* **29**, 90-8.

Wager, T. D., Lindquist, M. & Kaplan, L. (2007). Meta-analysis of functional neuroimaging data: Current and future directions. *Social Cognitive and Affective Neuroscience* **2**, 150-158.

Wager, T. D., Lindquist, M., Nichols, T. E., Kober, H. & Van Snellenberg, J. X. (2009). Evaluating the consistency and specificity of neuroimaging data using meta-analysis. *Neuroimage* **45**, S210-S221.

Warren, J. E., Sauter, D. A., Eisner, F., Wiland, J., Dresner, M. A., Wise, R. J., Rosen, S. & Scott, S. K. (2006). Positive emotions preferentially engage an auditory-motor "mirror" system. *Journal of Neuroscience* **26**, 13067-75.

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**, 655-64.

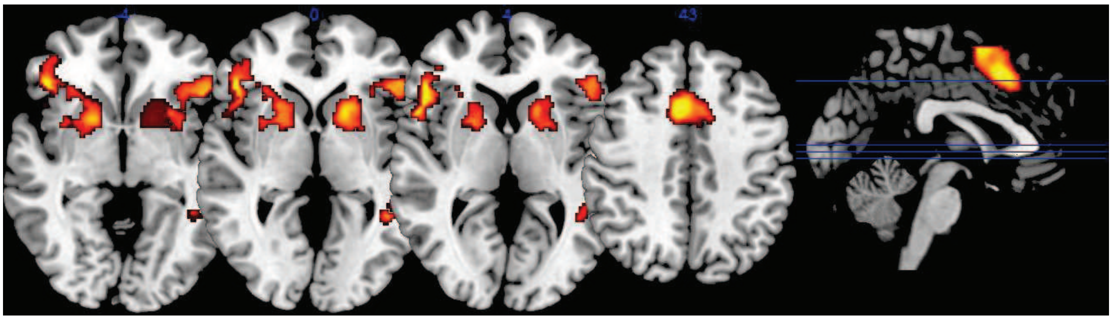
Yamada, M. & Decety, J. (2009). Unconscious affective processing and empathy: an investigation of subliminal priming on the detection of painful facial expressions. *Pain* **143**, 71-75.

Supplementary TABLE 1: Activation resulted from the contrast ‘viewing emotional face > viewing control face’ corresponding to Supplementary Figure 1

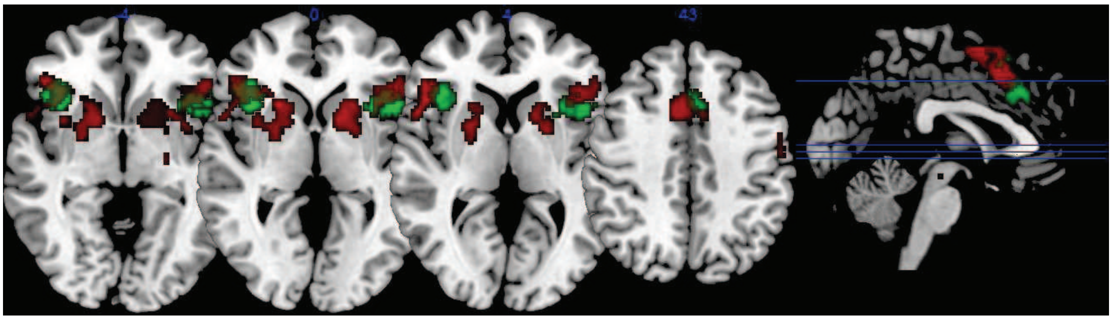
Activation of ‘viewing emotion > viewing smooth’	Brodmann area	T-value	MNI coordinate			Cluster size
			x	y	z	
Right presupplementary motor area	6	6.55	4	2	62	854
° Left supragenual anterior cingulate cortex	32	6.23	-2	14	46	
° Right supplementary motor area	6	5.24	6	-6	60	
Left inferior frontal gyrus	44	8.82	-54	16	8	532
° Left inferior frontal gyrus	47	6.9	-48	28	0	
° Left anterior insula / inferior frontal gyrus	47	5.6	-38	28	-6	
Left putamen		7.15	-22	18	0	371
° Left putamen		6.6	-16	8	-6	
° Left putamen		5.13	-22	6	6	
Right putamen		6.56	18	8	-2	624
° Right anterior insula / inferior frontal gyrus	47	5.29	44	24	2	
° Right putamen		5.11	22	4	6	
Right parahippocampal gyrus		7.63	24	-12	-18	325
° Right amygdala		5.03	30	0	-20	
° Right lateral globus palidus		4.74	26	-14	-6	
Right fusiform gyrus	19	7.07	46	-70	-16	268
° Right fusiform gyrus	18	7.02	22	-90	-18	
° Right cerebellum		5.19	38	-68	-26	
Left fusiform gyrus	18	6.02	-18	-92	-22	70
° Left fusiform gyrus	18	5.65	-26	-88	-24	
° Left cerebellum		4.76	-36	-80	-24	
Right primary motor cortex	4	5.62	50	-12	50	104
° Right pre-motor cortex	9	4.73	56	4	36	

FDR corrected $p<0.01$, cluster size>50.

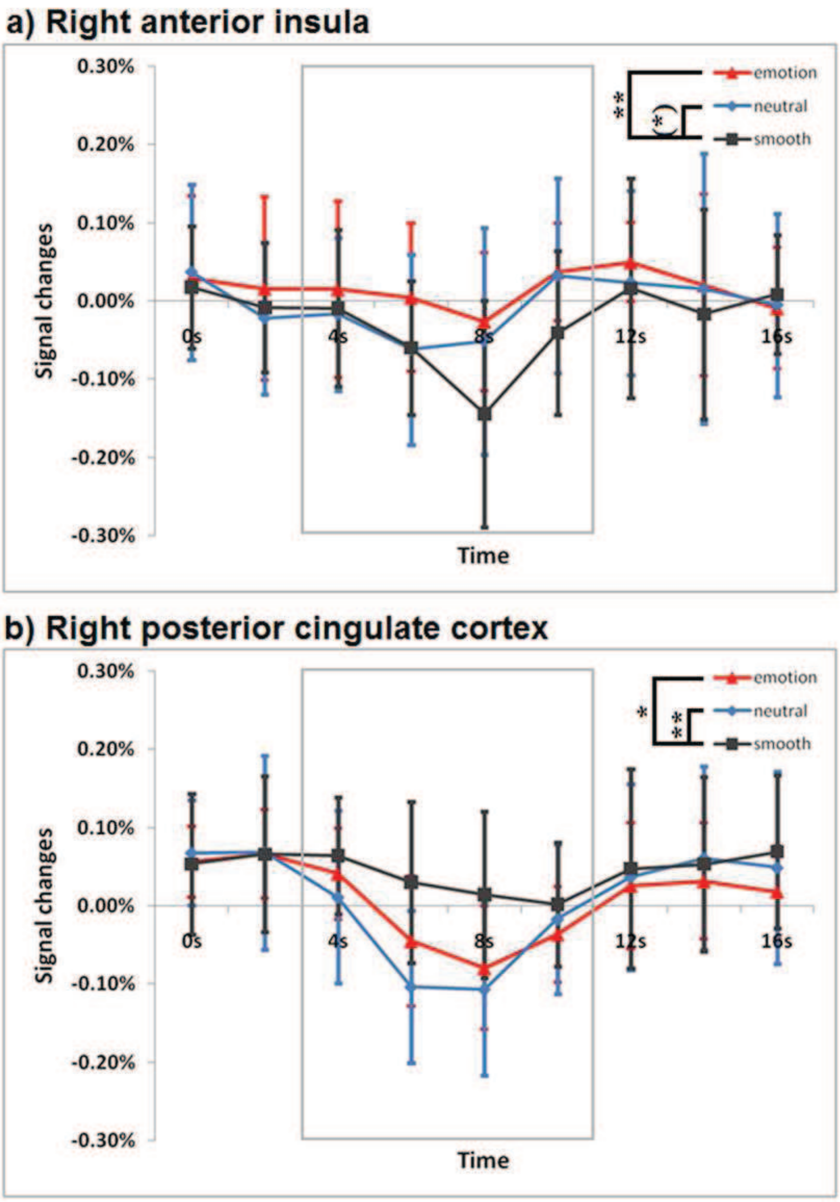
Supplementary FIGURE 1a: Empathy activation exclusively masked with motor activation



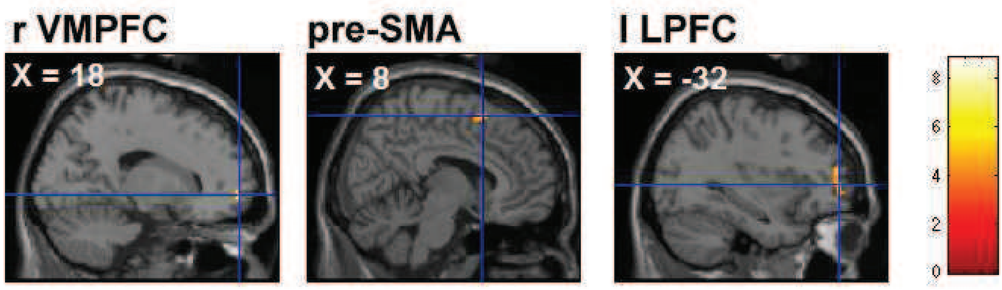
Supplementary FIGURE 1b: Comparison between empathy activation (red) and meta-analysis results of previous neuroimaging studies on empathy (green)



Supplementary FIGURE 2a: Signal changes of emotion, neutral, and smoothened conditions in right anterior insula and right posterior cingulated cortex



Supplementary FIGURE 2b: 2nd-level-regression analysis (NI score as regressor) for the contrast 'emotional face > neutral face'



Supplementary FIGURE 3: Percent of signal changes from right anterior insula correlated with NI score for all three groups of subjects.