The neural correlates of positive self-evaluation and self-related memory

Katharina Pauly,[1,2] Andreas Finkelmeyer,[1,2,3] Frank Schneider,[1,2] and Ute Habel[1,2]

[1]Department of Psychiatry, Psychotherapy, and Psychosomatic Medicine, Medical School, RWTH Aachen University, Aachen, [2]JARA-Brain Jülich Aachen Research Alliance – Translational Brain Medicine, Forschungszentrum Jülich GmbH, Jülich, Germany and [3]Institute of Neuroscience, Newcastle University, Campus for Ageing and Vitality, Newcastle-upon-Tyne, UK

Humans tend to have a positive self-evaluation (PSE). To what extent positive self-perception is interacting with valenced self-related memories is debated. The underlying neural substrates are not adequately explained yet. To explore the cerebral correlates of PSE and its influence on memory, 24 healthy subjects were asked during fMRI to decide in two conditions whether presented positive and negative personality traits characterized their own selves (self-evaluation) or an intimate other (other-evaluation). A lexical condition served as control task. In a subsequent unannounced recognition task, trait adjectives had to be classified as old or new. Activation during positive self- vs positive other-evaluation was found in the medial ventral and dorsolateral prefrontal gyri, the parahippocampus and the supplementary motor area. Memory increased for positive personality traits and traits that had been referred to oneself or the other. In contrast to adjectives of the other-evaluation or lexical condition, recollection of negative vs positive traits of the self-evaluation condition specifically induced increased activation in the hippocampus and several prefrontal and temporal areas. Our data imply a specific network for PSE (although intimate others are perceived similarly). Moreover, memory for traits contradicting PSE resulted in activation increases indicating greater cognitive effort and emotional involvement.

Keywords: self; emotion; memory; recognition; fMRI

INTRODUCTION

Healthy people tend to show a rather positive self-perception, e.g. when self-ascribing mostly positive and scarcely negative personality traits (Beer *et al.*, 2010; Pauly *et al.*, 2011). Positive attributional patterns seem to be beneficial for personal health and self-esteem. Changes in positive self-perception may result in mental disorders such as depression. However, despite plenty of literature on the cerebral correlates of self-reflection and meta-cognition, there is a lack of fMRI studies directly investigating this specific typical human characteristic, namely the neural networks underlying positive self-evaluation (PSE). Self-related information is processed rather deep and elaborate ('self-reference effect'; Symons and Johnson, 1997). Accordingly, self-referred material is better recalled than stimuli processed with respect to semantic, phonemic or structural features (Rogers *et al.*, 1977) or referred to a fictive other person (Kim and Johnson, 2012). However, it further remains unclear to what extent the valence of the self-ascribed material affects the neural correlates of later memory. Only few studies described the neural correlates of self-evaluation and subsequent stimulus recollection within one experiment and sample taking into account the individual behavioral responses.

　Studies so far implicate that the medial prefrontal cortex (mPFC) is a core brain structure for meta-cognitive functions, including self-knowledge (Amodio and Frith, 2006; see also Craik *et al.*, 1999). Healthy individuals consistently demonstrate activation in the mPFC and other midline structures, such as the posterior cingulate cortex, as a function of increasing self-descriptiveness (Moran *et al.*, 2006; see also: Northoff *et al.*, 2006), and while reflecting on their own personality traits compared to when they are engaged in semantic or lexical

tasks (Johnson *et al.*, 2002; Fossati *et al.*, 2003; Schmitz *et al.*, 2004; Ochsner *et al.*, 2005) or tasks requiring judgments about famous people (Kelley *et al.*, 2002). Since only *close* others are also seen in an overly positive light (and therefore more positive than the average peer; e.g. Murray, 1999; Hughes and Beer, 2012), a more specific and conservative comparison would be the contrast of self-evaluation and the evaluation of a *personally close* other person (see also Krienen *et al.*, 2010). This comparison has been associated with stronger dorsolateral (Schmitz *et al.*, 2004) and mPFC activation (Wang *et al.*, 2012). However, fMRI results upto now did not take into account the individual behavioral responses of the subjects.

　While the mPFC responded to self-descriptive material independent of the valence (Moran *et al.*, 2006), the orbitofrontal cortex (OFC) seemed more specifically related to overconfident evaluation of own performance (Beer *et al.*, 2010), but also to positive evaluation (POE) of close others (Hughes and Beer, 2012). Correspondingly, an increase in 'above average' self- and other-ratings was negatively related to medial and left lateral OFC activation (Beer and Hughes, 2010; Hughes and Beer, 2012). The investigation of the neural pattern during the self-ascription and rejection of valenced traits might shed further light on the specific interaction of emotion and self-evaluation.

　Self-relevance has also been investigated in memory research. The episodic retrieval of previously self-referred items was associated with activity in the mPFC (Lou *et al.*, 2004; Macrae *et al.*, 2004; Kim and Johnson, 2012) and the posterior cingulate gyrus. Furthermore, an increasing degree of self-relevance was related to activation increases in the right inferior parietal cortex (Lou *et al.*, 2004). Only a few studies have investigated memory for self-ascribed valenced personality traits allowing for separate analysis of the effects of reference and emotion. Fossati *et al.* (2004) found greater activation in the mPFC and the cerebellum for the recognition of *negative* traits of a previous self-evaluation condition as compared to negative traits that had been judged lexically or according to their social desirability. Analogous to the self-evaluation condition, the comparison with the memory for traits previously referred to an intimate other person would represent a more (self-)specific contrast. It still remains unclear

whether the activation increases similarly apply to the recognition of traits referred to a well-known other, and in what way the trait valence affects the neural substrates of self- and other-related memory.

The present fMRI study, therefore, focused on two main aspects of self-evaluation:

(1) the neural correlates of a typically PSE pattern in comparison to the respective evaluation of an intimate other based on the individual response patterns and
(2) the relationship of PSE and memory as well as valence effects on the cerebral substrates of memory for self-related material.

We further exploratorily investigated potential neurofunctional gender differences.

We expected an essential role of cortical midline structures, mainly the mPFC and the posterior cingulate gyrus, as well as of the OFC for PSE (vs lexical processing), as well as mainly dorsolateral prefrontal cortex (DLPFC) activation for the comparison of PSE and the POE of an intimate other. In addition to mPFC activation, we hypothesized medial temporal and lateral parietal activation for the recognition of previously self-referred items.

METHODS

This study was performed in accordance with the ethical standards of the Declaration of Helsinki. The local institutional review board of the Medical Faculty of RWTH Aachen University approved the protocol. After receiving a detailed description of the study, all participants gave their written informed consent.

Subjects

Twenty-four healthy right-handed (handedness inventory: Oldfield, 1971) participants (12 women), all native German speakers, took part in the study (Table 1). The Structured Clinical Interview for DSM-IV (SCID-I, German version: Wittchen et al., 1997) was performed to exclude actual or lifetime mental illness. Subjects with a medical condition that could influence cerebral metabolism were excluded as well as those with mentally ill first-degree relatives.

Stimuli and tasks

During the fMRI paradigm, personality traits were presented via PRESENTATION software package (Neurobehavioral Systems Inc., San Francisco, CA, USA) by means of a goggles system. Half of the stimuli were positive and the other half negative, with the adjectives not differing in concreteness, imagery, length, frequency of use or intensity of the perceived emotion (see Pauly et al., 2011). The task

started after a brief training phase outside the scanner. Each adjective was presented for 2 s followed by a fixation cross (jittered between 1.5 and 5.5 s; Figure 1).

During the encoding phase, which lasted ∼14 min, 126 trait adjectives were presented in three recurring conditions in pseudo-randomized order. Subjects were instructed to indicate via button press whether or not the attributes (i) characterized themselves (self-evaluation), (ii) characterized an intimate person (other-evaluation) or (iii) whether the word included the letter 'r' (lexical control task). Items of the same task and valence subcondition were presented in mini-blocks of three trials in an event-related design. Each task was preceded by a brief instruction (5 s). To control for order effects in the recognition phase, two different encoding versions were established: applying the split-half technique and counterbalanced across participants.

After a short break, an unannounced recognition task followed, during which 45 positive and 45 negative adjectives of the encoding phase were presented as well as 90 new valenced personality traits (distractors). In an event-related design (lasting 16 min), subjects had to decide if each stimulus had already been presented during the encoding phase.

After the fMRI measurement, each participant indicated which person he or she referred to during the 'other' condition and rated on an 8-point scale how he or she evaluated this person [−4 (very negative), ..., 4 (very positive)]. The relation to the intimate other person was not further predefined to account for interindividual differences regarding the family background.

MRI data acquisition

Data acquisition took place on a 3 T Phillips MR scanner at the University Hospital of RWTH Aachen University. Structural images were gathered via a standardized Magnetization Prepared Rapid Gradient Echo (MP-RAGE) three-dimensional T1-weighted sequence. Functional images were acquired with echo-planar imaging (EPI; T2*, TR = 2.4 s, TE = 30 ms; voxel size: $3.3 \times 3.3 \times 3$ mm^3, gap: 0.3 mm, 64×64 matrix, FoV: 211×211 mm^2, 37 slices, $\alpha = 90°$). Slices covered the whole brain. Three hundred and sixty volumes were collected during the encoding phase, 400 during the recognition phase.

Data analysis
Behavioral data

Behavioral data were analyzed with SPSS 17.0 for Windows (SPSS Inc., Chicago, USA). A correction of the degrees of freedom was undertaken

Table 1 Demographical and neuropsychological test results for healthy men and women [two-sample t-test: mean ± s.d., t-scores, df and P-values].

Test	Men (mean ± s.d.)	Women (mean ± s.d.)	t	df	P
Age (in years)	33.75 ± 5.75	33.92 ± 8.5	−0.05	22.00	0.957
Education (in years)	13.50 ± 2.94	13.16 ± 3.83	0.24	20.61	0.813
IQ (MWT-B)	115.42 ± 16.67	107.00 ± 11.10	1.46	22.00	0.160
TMT-A (in s)	25.72 ± 12.06	19.94 ± 6.33	1.47	22.00	0.156
TMT-B (in s)	52.22 ± 16.02	46.55 ± 15.71	0.88	22.00	0.391
PERT (percent correct)	80.42 ± 7.37	80.21 ± 8.01	0.07	22.00	0.208
PERT (reaction time)	2460.12 ± 556.17	2279.66 ± 388.26	0.92	22.00	0.367
VLMT learning TP	53.25 ± 6.41	57.17 ± 7.36	−1.39	22.00	0.178
VLMT loss after interference	1.67 ± 1.87	1.17 ± 1.12	0.79	22.00	0.436
VLMT recall TP	11.58 ± 2.39	11.58 ± 2.71	0.00	22.00	1.000
VLMT loss after delay	1.67 ± 1.37	1.58 ± 1.38	0.15	22.00	0.883
VLMT recognition TP	14.25 ± 0.87	14.17 ± 1.59	0.16	22.00	0.875

[MWT-B = Mehrfachwahl-Wortschatz-Intelligenztest—Version B (German multiple choice vocabulary test for crystalline intelligence; Lehrl, 1989); TMT = trail making test (Reitan, 1958); PERT = Penn Emotion Recognition Test (Kohler et al., 2004); VLMT = Verbal Learn and Memory Test—Version A (Helmstaedter et al., 2001; a German version of the California Verbal Learning Test by Delis et al., 2000); TP = true positives/hits].
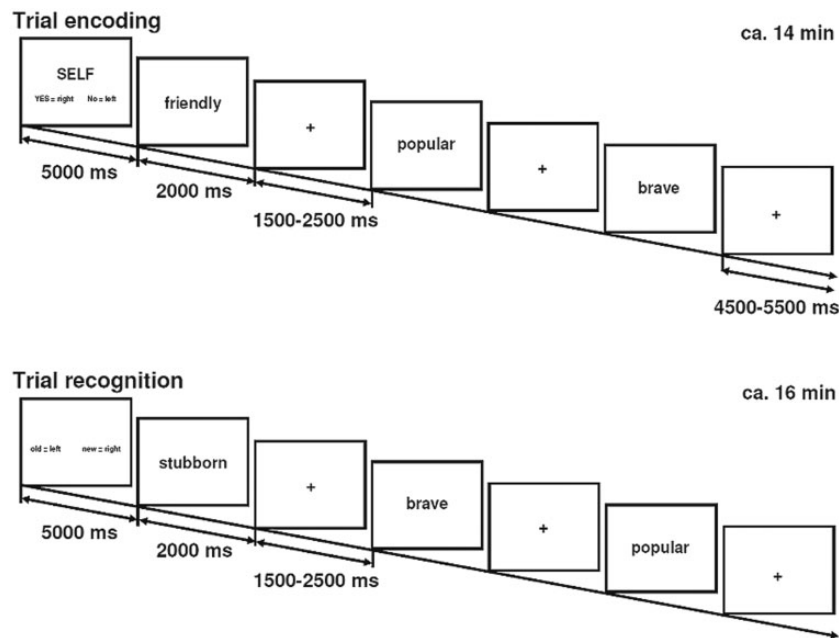
Fig. 1 fMRI paradigm with two runs—one for the encoding and one for the recognition of valenced personality traits. The encoding block was further subdivided in three tasks: self-evaluation regarding the traits, evaluation of a well-known other person and a lexical task.

if Levene's test for equality of variances (*t*-tests) or the Mauchly test on sphericity (ANOVA) revealed significance.

Self- and other-evaluation were assessed by the mean percentage of affirmed positive and negative personality traits for the given responses. A $2 \times 2 \times 2$ repeated measures ANOVA was performed to analyze the effects of the between-subjects factor gender and the within-subjects factors reference (self, other) and valence (positive, negative). Furthermore, a 2×2 ANOVA for positive self- and other-evaluation was calculated on the basis of the amount of affirmed positive and rejected negative traits comparing the sexes. The percentage of correct (true positive and true negative) answers during the lexical control task did not fulfill the assumption of a normal distribution (Kolmogorov–Smirnov test) and was therefore analyzed by means of Mann–Whitney tests regarding the effects of valence and gender.

Results of the recognition phase were investigated analyzing the ratio of correctly recognized words for given responses in a $3 \times 2 \times 2$ ANOVA with the factors reference condition (self-evaluation, other-evaluation, lexical processing), valence and gender. Finally, false alarm rates, i.e. the amount of falsely 'recognized' new words, were analyzed in a 2×2 ANOVA with the factors valence and gender.

FMRI data
fMRI data analysis was accomplished via SPM5 (Wellcome Department of Cognitive Neurology, London). Realignment and stereotaxic normalization ($3 \times 3 \times 3$ mm^3) were followed by smoothing with an 8-mm full-width-at-half-maximum Gaussian blurring kernel. A 1/128 Hz high-pass filter removed low-frequency noise. None of the data sets revealed movement parameters exceeding one voxel size.

For the encoding phase, a first-level model was established including six trial types for affirmed positive and rejected negative items during the self- and the other-evaluation conditions, and for correctly identified adjectives during the positive and negative lexical tasks, as well as a seventh regressor for the short instructions. Contrasts were entered in a random effects flexible factorial design at the second-level contrasting brain activation during PSE, i.e. the combined self-ascription of

positive and rejection of negative personality traits, and the correct lexical processing of positive and negative traits (affirming adjectives with an 'r' while rejecting personality traits without) as well as PSE and the POE, i.e. ascription of positive and rejection of negative personality traits to the intimate person. The comparison of self-evaluation and lexical condition allowed for the investigation of brain networks underlying self-reflection accounting for the unwanted effects of more basal processing steps, such as vision, reading, hearing of the scanner noise and sensorimotor activation. To further investigate which parts of this network are overlapping with activations also found for the POE of well-known others and which areas are distinctly and exclusively activated during self-reflection, we included the POE condition in order to compare self-reflection with a condition in many aspects very similar to self-evaluation, e.g. regarding positive assessment, affective involvement, personal character and knowledge of personal facts.

We further contrasted the results of both genders. In an additional correlation analysis, we correlated PSE brain activation (*vs* correct lexical processing) with the behavioral PSE response pattern.

For the recognition phase, we calculated two flexible factorial analyses on the second level. The first relied on traits that had been affirmed/ascribed or rejected during the encoding/evaluation task. The second flexible factorial analysis differentiated between correctly recognized (and correctly rejected) traits and errors of each encoding condition. Analogous to the analysis of the encoding task, we contrasted the combined recognition of previously affirmed positive and denied negative personality traits related to oneself *vs* the other person. Moreover, we contrasted the trials with correctly recognized adjectives with regard to both valences separately for each of the three reference conditions (self, other, lexical) independent of the answers during the encoding phase. Finally, we contrasted the activation patterns of both genders during the recognition of PSE traits alone, and for the recognition of previously ascribed positive and rejected negative personality traits referred to oneself *vs* the intimate other.

An error probability of 0.001 uncorrected (extent threshold: five voxel) was adopted for all functional analyses.

RESULTS

Behavioral data

Encoding

Men and women did not differ in their positive–negative rating of the intimate other person ($Z = -0.13$; $P = 0.899$) with a median of 3.50 in women and 3.00 in men.

The ANOVA for the self-ascribed personality traits revealed a significant main effect for valence ($F = 511.34$, df = 1, 22; $P < 0.001$) with positive traits being affirmed much more often than the negative ones when referred to self or another person (Figure 2). No significant effects were found for reference, gender or the interaction effects reference × valence, reference × gender, valence × gender or task × valence × gender (all $F < 1.00$). Accordingly, we found no gender, reference or interaction effect for the PSE response pattern (all $F < 1$). The correct reactions during the lexical task revealed no significant effects, neither for gender ($Z = -0.27$; $P = 0.787$) nor valence ($Z = -1.38$; $P = 0.169$).

Recognition

Analyzing hits/true positive answers during recognition yielded a significant main effect for valence ($F = 19.74$, df = 1, 22; $P < 0.001$), with a better recognition performance for positive traits as compared to negative ones (Figure 2). Another significant main effect was found for the previous reference condition of the words ($F = 37.02$, df = 2, 44; $P < 0.001$), but not for gender ($F = 0.07$, df = 1, 22; $P = 0.795$). Post hoc tests revealed better performance for adjectives referred to oneself ($t = 7.42$, df = 23; $P < 0.001$) or an intimate person ($t = 5.85$, df = 23; $P < 0.001$) as compared to lexically processed stimuli with no differences between the first two ($t = 0.23$, df = 23; $P = 0.817$). The interactions of valence × gender ($F = 3.58$, df = 1, 22; $P = 0.072$), reference × valence ($F = 0.80$, df = 2, 44; $P = 0.455$) and reference × valence × gender ($F = 0.42$, df = 2, 44; $P = 0.658$) revealed no significant effects. Gender interacted significantly with the reference condition ($F = 3.94$, df = 2, 44; $P = 0.027$). However, the respective post hoc t-tests did not reach significance with any gender differences for words previously referred to oneself, the close other or processed lexically (all $t < 1.00$).

False alarm rates indicated a main effect of valence ($F = 43.04$, df = 1, 22; $P < 0.001$) with more false-positive reactions for items with positive valence. While there was no effect for gender ($F = 1.50$, df = 1, 22; $P = 0.234$), gender interacted with valence ($F = 6.61$, df = 1, 22; $P = 0.017$). However, no gender differences in the false alarm rates were found post hoc for negative items ($t = -0.04$, df = 22; $P = 0.970$),

and they marginally failed to reach significance for positive traits ($t = -2.00$, df = 22; $P = 0.058$) with slightly more false alarms in women.

fMRI

PSE during encoding

To investigate the neural correlated of PSE, we first compared brain activation for the self-ascription of positive and the rejection of negative personality traits (PSE) *vs* the correct lexical processing of traits. Activation was found in the medial superior PFC, the left inferior OFC, the precuneus extending to the posterior cingulate gyrus (deactivated during the lexical task), the left angular and middle temporal gyri, the right middle temporal pole and the left hippocampus (Table 2 and Figure 3). Due to the fact that on average only a small number of self-referred negative traits was affirmed, and few positive traits were rejected (overall 4–5 items of the 42 stimuli during the self-condition), the opposite contrast (for rejected positive and affirmed negative personality traits) was not calculated.

The contrast of positive traits affirmed and negative traits rejected related to the self *vs* an intimate person yielded activation in the left dorsal PFC, the gyrus rectus, the supplementary motor area (SMA) and the right parahippocampus (Table 2). No significant activation increases were found for the opposite direction.

Furthermore, no gender differences were found during PSE *vs* lexical processing or PSE *vs* POE.

The correlation analysis revealed a negative correlation ($r = -0.70$; $P < 0.001$) between the mean percentage of affirmed positive and rejected negative personality traits and activation in the mPFC ($x = -3$, $y = 66$, $z = 20$, $k_E = 17$, $t = 4.54$) during PSE (*vs* lexical processing; Figure 4).

Recognition

A comparison between the correct recognition of previously affirmed positive and rejected negative personality traits of the self- *vs* other-evaluation condition revealed activation in the left middle temporal lobe, the left lingual gyrus and the right cerebellum (Table 3). No activation was found for the opposite contrast or when comparing men and women.

A closer look at the contrast of correctly recognized negative *vs* positive traits of the self-condition exposed a large cluster of (para-)hippocampus activation, activation in the anterior medial and DLPFC and cerebellum bilaterally, the right OFC, angular gyrus, cuneus and the inferior occipital cortex, as well as the inferior triangular frontal gyrus, the superior and middle temporal gyri, the fusiform gyrus and
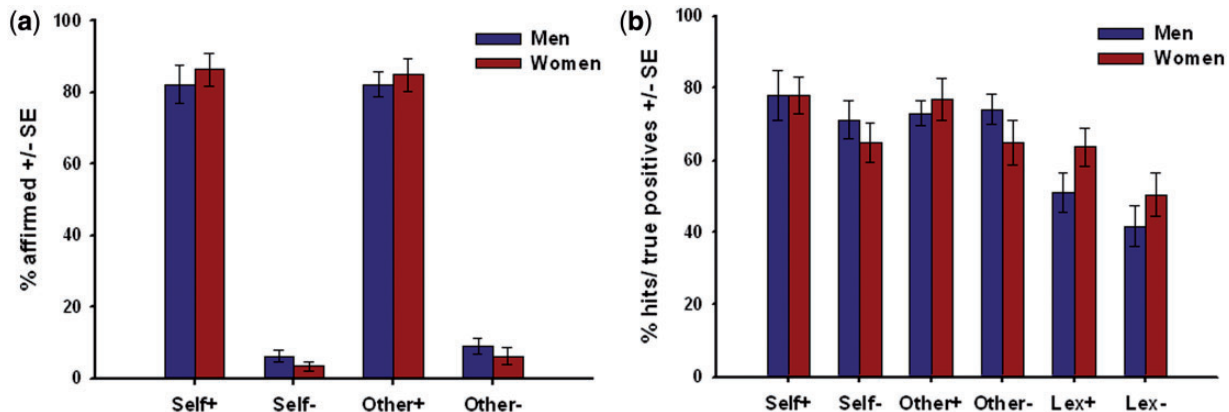
Fig. 2 (**a**) Higher percentage of ascribed positive (+) as compared to negative (−) personality traits referred to oneself or another close person in men and women; (**b**) better recognition performance for personality traits referred to oneself or a close person as compared to adjectives processed lexically (Lex) and for positive traits as compared to negative ones (SE = standard error).

calcarine sulcus of the left side (Table 3 and Figure 5). No significant results were found for the opposite direction (positive > negative self) or the comparison between the correct recognition of negative and positive traits of the two other reference conditions (other, lexical). Again, no gender effects were found.

Table 2 Brain activation during the self-evaluation *vs* lexical baseline (flexible factorial analysis; $P < 0.001$ uncorrected, extent threshold: five voxels; MNI coordinates)

Region	Side	x	y	z	k_E	t
PSE > lexical[a]						
Superior mPFC	L	−3	59	16	1247	8.35*
Inferior orbitofrontal gyrus	L	−50	26	−7	376	7.84*
Posterior cingulate gyrus, precuneus	L	−3	−56	23	283	6.69*
Angular gyrus	L	−53	−73	30	144	4.65*
Middle temporal gyrus	L	−56	−13	−20	8	3.62*
	L	−53	−30	−10	9	3.38*
Middle temporal pole	R	36	17	−33	81	4.00*
Hippocampus	L	−23	−13	−23	10	3.58*
PSE > POE[b]						
Dorsolateral prefrontal gyrus	L	−56	17	36	12	3.81
Gyrus rectus (medial ventral frontal gyrus)	R	7	23	−20	8	3.76
Supplementary motor area	R	3	23	66	19	3.57
Parahippocampus	R	16	−3	−26	10	3.80

[a]PSE (positive traits affirmed and negative traits rejected) *vs* lexical baseline.
[b]PSE *vs* POE of an intimate other person.
*Also found for $P < 0.05$ FDR corrected for multiple comparisons.

## DISCUSSION

We always strive for a positive image of ourselves, which is of considerable relevance for our self-esteem and well-being. In our study, we sought to investigate the neural correlates of PSE usually found in healthy individuals and to analyze its specificity in comparison to the evaluation of personality traits of another intimate well-known person. We further examined the effects of reference condition and valence on later recognition and the underlying cerebral substrates.

Self-evaluation during encoding

Subjects affirmed most of the positive traits while rejecting most of the negative characteristics (see also Pauly *et al.*, 2011). This was true for both self-evaluation and evaluation of an close other, quite in keeping with the general tendency, in healthy humans, toward positive self-perception (Beer and Hughes, 2010; Beer *et al.*, 2010) and perception of intimate others (Neff and Karney, 2005), which seem to interact with each other (Murray 1999; Murray *et al.*, 2002). Accordingly, our neurofunctional results cannot just be traced back to a more positive evaluation of oneself than of the other person.

The comparison between PSE and POE revealed activation in the SMA, a key region for self-initiated actions (Jenkins *et al.*, 2000; Wu *et al.*, 2011). The SMA is considered to be part of the network underlying the feeling of agency and the ability to differentiate between actions caused by ourselves and others (Yomogida *et al.*, 2010).
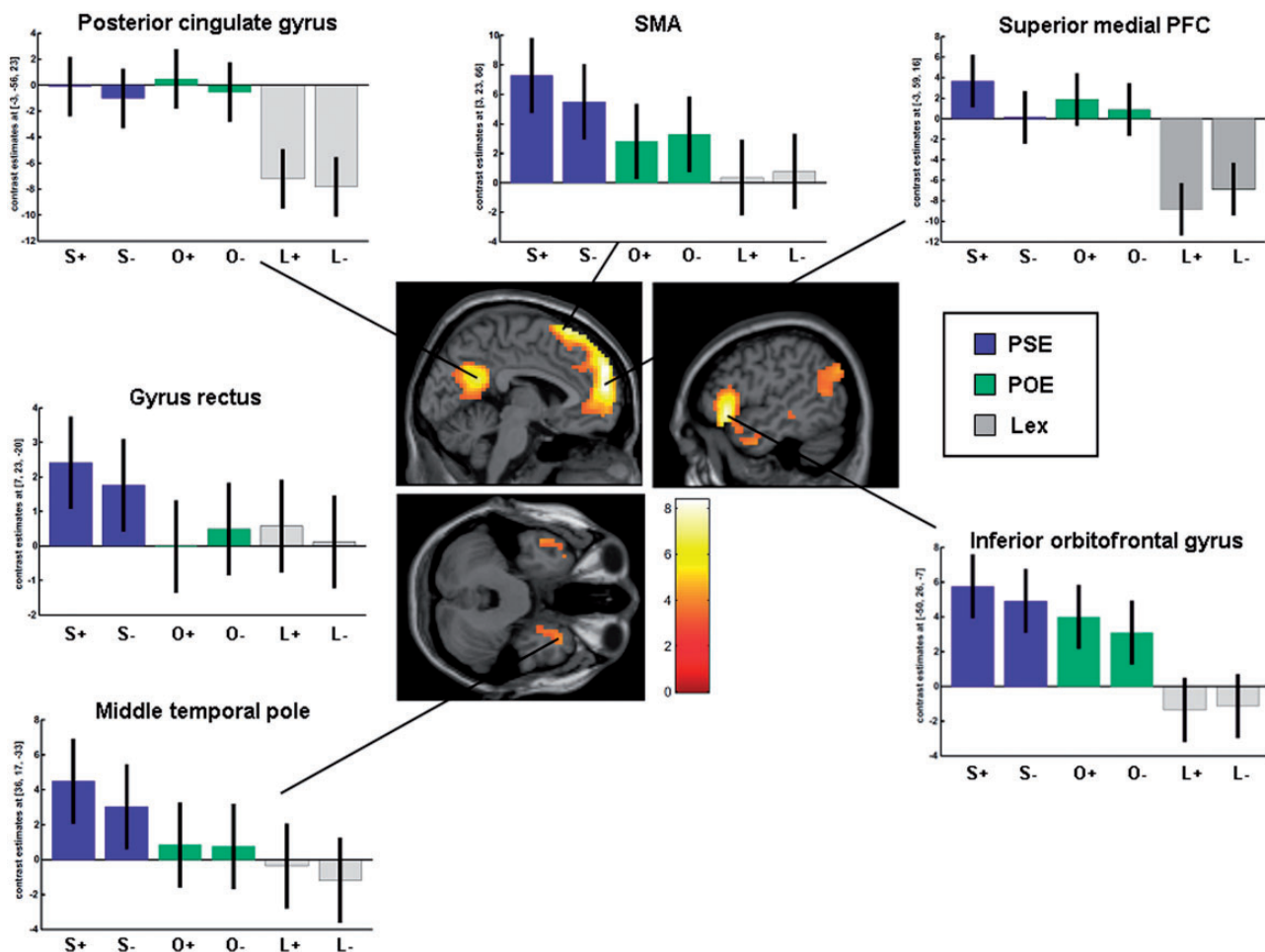
Fig. 3 Brain activation during PSE [i.e. self-ascribed positive (S[+]) and rejected negative traits (S[−])] *vs* lexical processing of positive (L[+]) and negative (L[−]) personality traits: activation in the mPFC, including the supplementary motor area (SMA), the posterior cingulate gyrus, left inferior orbitofrontal gyrus and the right middle temporal pole. Parameter estimates for the encoding condition are presented [POE of an intimate other person, i.e. positive traits ascribed to the intimate other (0[+]) or rejected negative traits (0[−])].

Interestingly, already imagined movements activate the SMA underlining its key role in self-consciousness and self-relatedness (e.g. Kimberley *et al.*, 2006). Moreover, SMA activation was found during online self-evaluation of confidence during decision making (Beer *et al.*, 2010) and when reflecting on one's positive traits in the present, the past and the future (D'Argembeau *et al.*, 2010). Hence, not only the imagination of own actions, but also the reflection on self-related traits (associated with self-initiated actions), results in SMA activation differentiating self from other. This is underlined by

Table 3 Brain activation (flexible factorial analyses; $P < 0.001$ uncorrected, extent threshold: five voxels; MNI coordinates) during the recognition phase

| Region	Side	x	y	z	k_E	t
PSE > POE[a]						
Middle temporal lobe	L	−53	−46	13	5	3.31
Lingual gyrus	L	−7	−82	−13	10	3.60
Cerebellum	R	13	−79	−17	6	3.34
Self-negative > self-positive						
Superior medial frontal cortex	L	−7	53	26	23	3.79*
Anterior mPFC	R	3	53	46	103	4.46*
	R	10	43	56	5	3.38*
Medial frontal gyrus	R	20	17	40	13	3.43*
DLPFC	R	26	46	46	63	4.00*
	R	46	30	46	19	3.80*
Middle frontal gyrus	L	−30	46	23	5	3.34*
Inferior triangular frontal gyrus	L	−43	26	26	41	3.57*
Middle orbitofrontal gyrus	R	40	50	−3	5	3.33*
Inferior orbitofrontal gyrus	R	53	36	−7	8	3.65*
Angular gyrus	R	53	−56	23	20	3.54*
Cuneus	R	10	−66	23	25	3.77*
Superior temporal gyrus	L	−66	−20	−3	19	3.47*
Middle temporal gyrus	L	−66	−46	0	6	3.55*
Hippocampus, parahippocampus	R	33	−10	−17	2823	5.14*
Fusiform gyrus	L	−40	−23	−20	7	3.73*
Inferior occipital cortex	R	40	−66	−10	20	3.98*
Calcarine sulcus	L	−13	−92	0	13	3.43*
Cerebellum	R	13	−46	−36	66	4.03*
	L	−30	−89	−20	20	3.84*
	L	−17	−40	−40	9	3.67*

[a]The correct recognition of previously affirmed positive and denied negative personality traits related to the own *vs* the other person (PSE *vs* POE).
[b]Correctly recognized negative *vs* positive traits of the self-evaluation condition.
*Also found for $P < 0.05$ FDR corrected for multiple comparisons.

the fact that not only OFC and lateral inferior prefrontal activation, but also activation in the SMA was found for the interaction of self-relevance and valence (Moran *et al.*, 2006).

The comparison of PSE and positive other-evaluation also revealed activation in hypothesized areas, namely in the (left) DLPFC and ventral mPFC (gyrus rectus). Although the results have to be interpreted with caution, given the small cluster sizes, this corroborates the notion that the DLPFC plays a key role in the delimitation of specifically self-related evaluation processes from other evaluative functions (Schmitz *et al.*, 2004). Activation may not only be restricted to *positive* self-perception but also was found during self-criticism (Longe *et al.*, 2010), pointing to a lesser extent to valence-specific effects than to self-specificity. Moreover, also the stimulus characteristics during self-reflection may exert influence. In this context, the study of Beer and Hughes (2010) could show that, when judging own personality traits relative to an average peer, activation in the ventral mPFC (including the gyrus rectus) was stronger for specific as compared to broad personality traits. In addition to its role in the differentiation of self- and other-related evaluative processes, we suggest that the ventral PFC and parahippocampal activation indicate a greater emotional involvement (Royet *et al.*, 2000; Takahashi *et al.*, 2004; Ochsner *et al.*, 2005) during (positive) evaluation of the self as compared to the evaluation of others. In line with our findings, parahippocampus, superior dorsal and medial ventral PFC were all involved when contrasting statements about the personal past and possible future with non-personal thinking (Abraham *et al.*, 2008).

Larger effects for PSE *vs* lexical processing as compared to PSE *vs* POE, especially in cortical midline structures, can be traced back to largely overlapping neural networks for self-evaluation and evaluation of close others (Schmitz *et al.*, 2004; Ochsner *et al.*, 2005; Vanderwal *et al.*, 2008; Grigg and Grady, 2010) with only some studies reporting increased activation in the mPFC (Heatherton *et al.*, 2006) or other prefrontal areas including the superior frontal cortex (Vanderwal *et al.*, 2008; Benoit *et al.*, 2010) for self-reference. Correspondingly, anterior mPFC activation was correlated with the perceived similarity between oneself and others while mentalizing the feelings of the counterpart (Mitchell *et al.*, 2005).

Inhibiting the mPFC temporarily by means of transcranial magnetic stimulation reduced the overly positive self-perception of subjects in ratings of their own desirable and undesirable traits (Kwan *et al.*, 2007). In this regard, we found a negative correlation between the behavioral PSE pattern and activation in the mPFC, i.e. with an
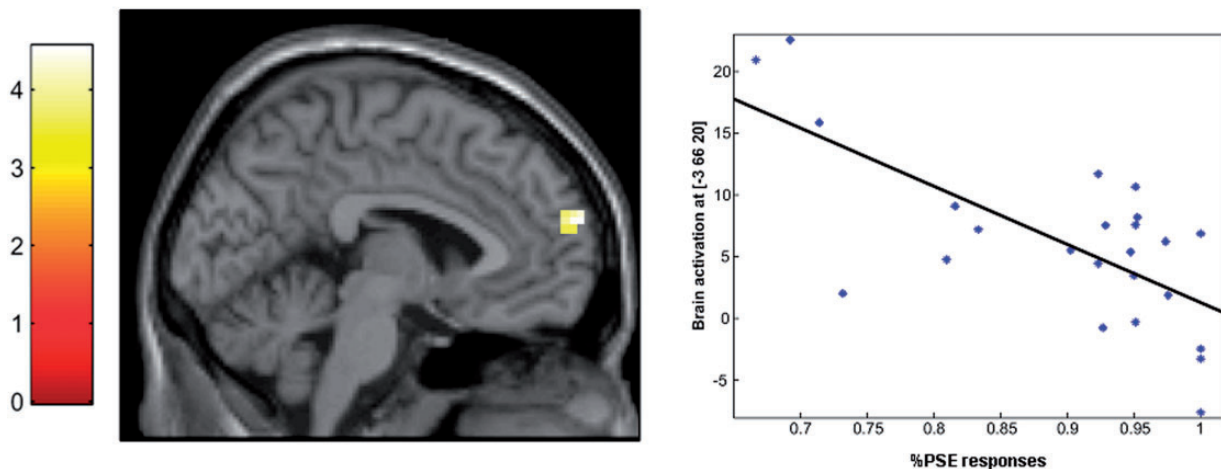
Fig. 4 Negative correlation between brain activation during self-ascription of positive and rejection of self-referred negative personality traits (PSE) *vs* lexical processing contrast and a behavioral PSE response pattern in the anterior mPFC.
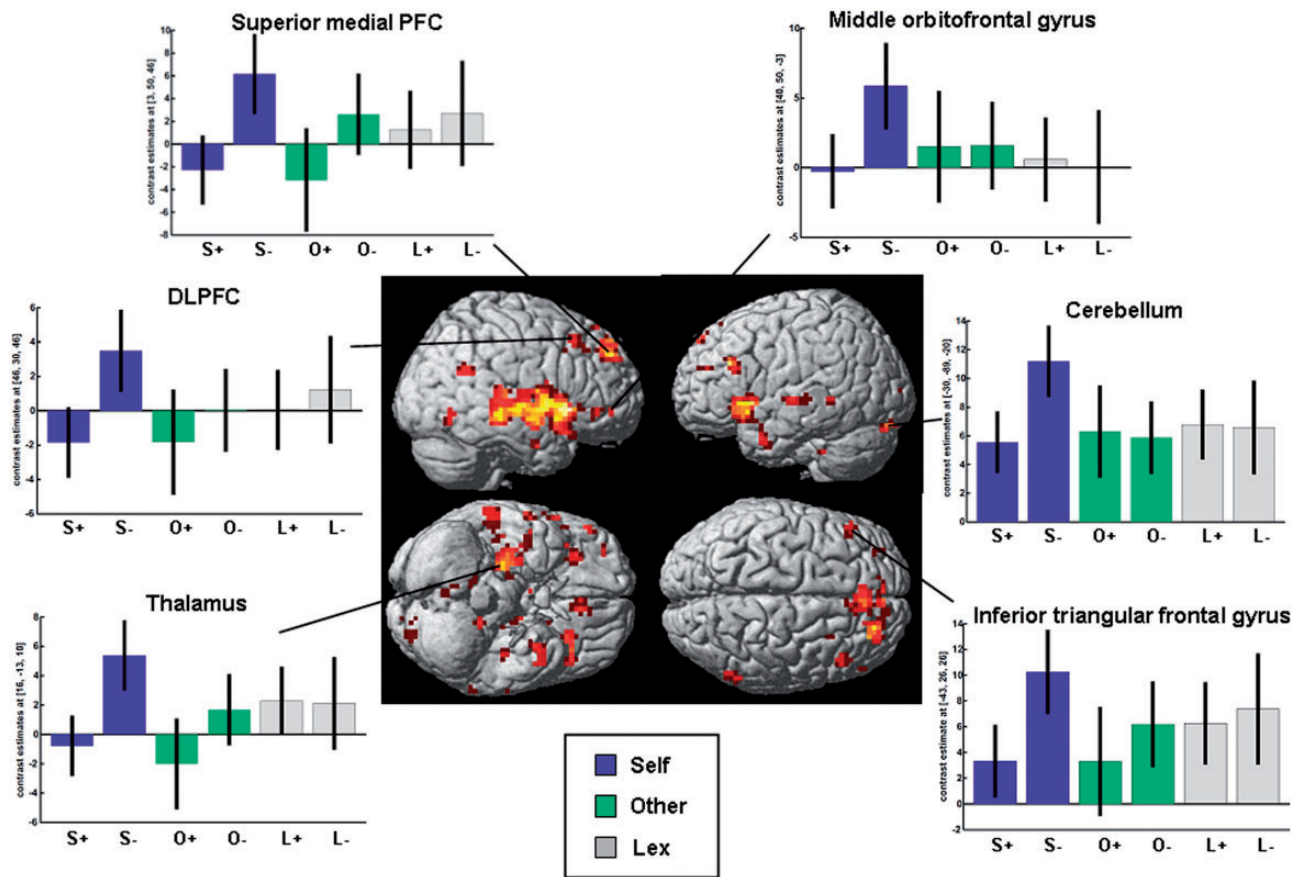
Fig. 5 Brain activation for correctly recognized negative (S⁻) *vs* positive (S⁺) personality traits of the self-evaluation condition: activation in the medial and DLPFC, the orbitofrontal gyrus and the cerebellum. Parameter estimates for the correct recognition of trait adjectives of the evaluation and lexical processing condition are presented (0⁺ = other positive, 0⁻ = other negative; L⁺ = lexical positive, L⁻ = lexical negative).

increase of PSE responses activation in the mPFC decreased, underlining its central role in the modulation of self-perception.

Valence-related differences in referential memory

In consistence with our own earlier findings (Pauly *et al.*, 2011; see also Danion *et al.*, 2003), but in contrast to Fossati *et al.* (2004; see also: Buchanan *et al.*, 2001), positive adjectives were better remembered than negative ones. Results so far imply a higher predictive power of emotional arousal than for valence itself (Bradley and Lang, 2000). Since our negative and positive adjectives were matched according to their valence intensity, a better memory for positive personality traits may be related to an inconsistency-negativity neglect model, implying a worse memory for information that contradicts positive self-beliefs (Sedikides and Green, 2000). In line with this explanation, we also found more false alarms for positive personality traits. Moreover, in accordance with the self-reference effect, adjectives were better remembered if they had been referred to oneself previously, but also when referred to a well-known other person, as compared to purely lexically processed words (see also Miall, 1986; Pauly *et al.*, 2011).

The direct comparison of memory processes during the recognition of PSE material and stimuli that had been answered in favor of the intimate other revealed activation in small clusters of the left middle temporal lobe and the right cerebellum. The middle temporal lobe is not only involved in self-referential processes, as described above, but also associated with memory processes. Accordingly, the middle temporal gyrus and the cerebellum were both identified as core areas of autobiographical memory (in addition to mainly prefrontal

areas: Ryan *et al.*, 2001; Svoboda *et al.*, 2006). Self-related processes are an essential aspect of autobiographical memory. While the middle temporal cortex was linked to semantic aspects of autobiographical memory, the role of the cerebellum is rather unclear being also involved in several higher order cognitive functions (Ravizza *et al.* 2006; Garrard *et al.*, 2008), such as reflection (D'Argembeau *et al.*, 2005), decision making under uncertainty (Blackwood *et al.*, 2004) and self-responsibility (Blackwood *et al.*, 2003).

Further decomposing the valence effects of self-referential memory, the direct comparison of correctly recognized negative *vs* positive traits of the self-condition not only revealed activation in several lateral and medial prefrontal areas and the cerebellum, as also found by Fossati *et al.*, 2004 during the recognition of negative traits of a self-reference task, but also in a broader network comprising the (para-) hippocampus, the angular gyrus and the middle and superior temporal gyri (with no significant activation increases for the opposite contrast). Interestingly, these activations were specific to the correct recognition of traits of the self-evaluation condition. The comparison of brain activation for negative and positive traits of the other-evaluation or the lexical condition revealed no activation differences. This underlines the specificity of increased effort and brain activation in a large network of autobiography- and emotion-related areas (see also: Richardson *et al.*, 2004) during the recollection of previously self-referred negative traits. All mentioned areas are part of the complex brain network involved in autobiographical memory—important not only when falling back on past experiences (Svoboda *et al.*, 2006) while judging one's personality but obviously also during recognition

of such self-related traits. The largest activation cluster was found in the hippocampus and parahippocampus. The hippocampus is indeed a key region for memory processes, including recollection of recent and remote vivid autobiographical memories (Cabeza and St Jacques, 2007; Rabin *et al.*, 2009), and is also linked to more general episodic retrieval (Schacter and Wagner, 1999). Correspondingly, the anterior (para-) hippocampus has been related to the retrieval of (visual) scenes (Rombouts *et al.*, 2001). While objective recollection (i.e. relying on source memory) mainly increased lateral PFC and inferior parietal activation, subjective recollection (e.g. 'remember *vs* know') mainly involved (para-)hippocampus and anterior mPFC activation, but also activation in the angular gyrus (Spaniol *et al.*, 2009; Kim, 2010). It may be speculated that the increased effort during recollection of negative *vs* positive traits of the self-condition, as reflected in heightened activation in brain networks related to autobiographical memory and retrieval success, points to a deeper encoding of the negative traits, despite their rejection, due to an augmented emotional involvement in self-evaluation concerning unfavorable personality characteristics. Alternatively, negative traits result in greater resistance, which has to be overcome during recall. Moreover, due to the tendencies of PSE and the more frequent self-ascription of positive (as compared to negative) traits, during encoding the correct recollection of positive traits of the self-evaluation (but also other-evaluation) condition might be facilitated.

CONCLUSIONS

Our data underline the essential role of the reference condition and the valence of the processed personality traits during self-evaluation for their later recognition and its underlying neural correlates. Like many other imaging studies, we acknowledge the problem of small sample size—especially as compared to behavioral studies. Correspondingly, some of the weaker effects and comparisons between subgroups have to be considered as exploratory.

While the mPFC is related to modulating effects of PSE, the differentiation between favorable evaluation of oneself or an intimate other is mainly linked to regions associated with self-consciousness and emotional involvement. The memory of self-related negative traits, finally, might interfere with attitudes of positive self-perception and call for increased effort and activation in brain networks of autobiographical memory and retrieval success. We found no behavioral (see also: Lameiras Fernández and Rodríguez Castro, 2003; Garaigordobil *et al.*, 2008; Somerville *et al.*, 2010) or neurofunctional gender differences regarding PSE or self-referential memory.

Self-evaluative processes clearly have clinical implications as certain mental disorders affect self-concept and self-related optimism, such as depression (e.g. Stone *et al.*, 2001) or schizophrenia (Fannon *et al.*, 2009; Pauly *et al.*, 2011). The investigation of underlying neural changes would be a worthwhile enterprise.

Conflict of Interest

None declared.

REFERENCES

Abraham, A., Schubotz, R.I., von Cramon, D.Y. (2008). Thinking about the future versus the past in personal and non-personal contexts. *Brain Research*, 1233, 106–19.

Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature*, 7, 268–77.

Beer, J.S., Hughes, B.L. (2010). Neural systems of social comparison and the "above-average" effect. *NeuroImage*, 49, 2671–9.

Beer, J.S., Lombardo, M.V., Bhanji, J.P. (2010). Roles of medial prefrontal cortex and orbitofrontal cortex in self-evaluation. *Journal of Cognitive Neuroscience*, 22, 2108–19.

Benoit, R.G., Gilbert, S.J., Volle, E., Burgess, P.W. (2010). When I think about me and simulate you: medial rostral prefrontal cortex and self-referential processes. *NeuroImage*, 50, 1340–9.

Blackwood, N.J., Bentall, R.P., Ffytche, D.H., Simmons, A., Murray, R.M., Howard, R.J. (2003). Self-responsibility and the self-serving bias: an fMRI investigation of causal attributions. *NeuroImage*, 20, 1076–85.

Blackwood, N., Ffytche, D., Simmons, A., Bentall, R., Murray, R., Howard, R. (2004). The cerebellum and decision making under uncertainty. *Cognitive Brain Research*, 20, 46–53.

Bradley, M.M., Lang, P.J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, 37, 204–15.

Buchanan, T.W., Denburg, N.L., Tranel, D., Adolphs, R. (2001). Verbal and nonverbal emotional memory following unilateral amygdale damage. *Learning & Memory*, 8, 326–35.

Cabeza, R., St Jacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends in Cognitive Science*, 11, 219–27.

Craik, F.I.M., Moroz, T.M., Moscovich, M., et al. (1999). In search of the self: a positron emission tomography study. *Psychological Science*, 10, 26–34.

Danion, J.-M., Kazes, M., Huron, C., Karchouni, N. (2003). Do patients with schizophrenia consciously recollect emotional events better than neutral events? *American Journal of Psychiatry*, 160, 1879–81.

D'Argembeau, A., Collette, F., Van der Linden, M., et al. (2005). Self-referential reflective activity and its relationship with rest: a PET study. *NeuroImage*, 25, 616–24.

D'Argembeau, A., Stawarczyk, D., Majerus, S., Collette, F., Van der Linden, M., Salmon, E. (2010). Modulation of medial prefrontal and inferior parietal cortices with thinking about past, present, and future selves. *Social Neuroscience*, 5, 187–200.

Delis, D.C., Kramer, J.H., Kaplan, E., Ober, B.A. (2000). *California Verbal Learning Test*, 2nd edn. San Antonio, TX: Psychological Corporation.

Fannon, D., Hayward, P., Thompson, N., Green, N., Surguladze, S., Wykes, T. (2009). The self or the voice? Relative contributions of self-esteem and voice appraisal in persistent auditory hallucinations. *Schizophrenia Research*, 112, 174–80.

Fossati, P., Hevenor, S.J., Graham, S.J., et al. (2003). In search of the emotional self: an fMRI study using positive and negative emotional words. *American Journal of Psychiatry*, 160, 1938–45.

Fossati, P., Hevenor, S.J., Lepage, M., et al. (2004). Distributed self in episodic memory: neural correlates of successful retrieval of self-encoded positive and negative personality traits. *NeuroImage*, 22, 1596–604.

Garaigordobil, M., Pérez, J.I., Mozaz, M. (2008). Self-concept, self-esteem and psychopathological symptoms. *Psicothema*, 20, 114–23.

Garrard, P., Martin, N.H., Giunti, P., Cipolotti, L. (2008). Cognitive and social cognitive functioning in spinocerebellar ataxia – a preliminary characterization. *Journal of Neurology*, 255, 398–405.

Grigg, O., Grady, C.L. (2010). The default network and processing of personally relevant information: converging evidence from task-related modulations and functional connectivity. *Neuropsychologia*, 48, 3815–23.

Heatherton, T.F., Wyland, C.L., Macrae, C.N., Demos, K.E., Denny, B.T., Kelley, W.M. (2006). Medial prefrontal activity differentiates self from close others. *Social Cognitive and Affective Neuroscience*, 1, 18–25.

Helmstaedter, C., Lendt, M., Lux, S. (2001). *VLMT Verbaler Lern und Merkfähigkeitstest. Test und Testhandbuch*. Göttingen: Beltz Test GmbH.

Hughes, B.L., Beer, J.S. (2012). Orbitofrontal cortex and anterior cingulate cortex are modulated by motivated social cognition. *Cerebral Cortex*, 22, 1372–81.

Jenkins, I.H., Jahanshahi, M., Jueptner, M., Passingham, R.E., Brooks, D.J. (2000). Self-initiated versus externally triggered movements II. The effect of movement predictability on regional cerebral blood flow. *Brain*, 123, 1216–28.

Johnson, S.C., Baxter, L.C., Wilder, L.S., Pipe, J.G., Heiserman, J.E., Prigatano, G.P. (2002). Neural correlates of self-reflection. *Brain*, 125, 1808–14.

Kelley, W.M., Macrea, C.N., Wyland, C.L., Caglar, S., Inati, S., Heatherton, T.F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, 14, 785–94.

Kim, H. (2010). Dissociating the roles of the default-mode, dorsal, and ventral networks in episodic memory retrieval. *NeuroImage*, 50, 1648–57.

Kim, K., Johnson, M.K. (2012). Extended self: medial prefrontal activity during transient association of self and objects. *Social Cognitive and Affective Neuroscience*, 7, 199–207.

Kimberley, T.J., Khandekar, G., Skraba, L.L., Spencer, J.A., Van Gorp, E.A., Walker, S.R. (2006). Neural substrates for motor imagery in severe hemiparesis. *Neurorehabilitation and Neural Repair*, 20, 268–77.

Kohler, C.G., Turner, T.H., Gur, R.E., Gur, R.C. (2004). Recognition of facial emotions in neuropsychiatric disorders. *CNS spectrums*, 9, 267–74.

Krienen, F.M., Tu, P.-C., Buckner, R.L. (2010). Clan mentality: evidence that the medial prefrontal cortex responds to close others. *Journal of Neuroscience*, 30, 13906–15.

Kwan, V.S., Barrios, V., Ganis, G., et al. (2007). Assessing the neural correlates of self-enhancement bias: a transcranial magnetic stimulation study. *Experimental Brain Research*, 182, 379–85.

Lameiras Fernández, M., Rodríguez Castro, Y. (2003). Age and sex differences in self-esteem among Spanish adolescents. *Psychological Reports*, 93, 876–8.

Lehrl, S. (1989). *Mehrfachwahl-Wortschatz-Intelligenztest MWT-B Manual*. Erlangen: Perimed Fachbuch-Verlagsgesellschaft.

Longe, O., Maratos, F.A., Gilbert, P., et al. (2010). Having a word with yourself: Neural correlates of self-criticism and self-reassurance. *NeuroImage*, 49, 1849–56.

Lou, H.C., Luber, B., Crupain, M., et al. (2004). Parietal cortex and representation of the mental self. *Proceedings of the National Academy of Sciences USA*, 101, 6827–32.

Macrae, C.N., Moran, J.M., Heatherton, T.F., Banfield, J.F., Kelley, W.M. (2004). Medial prefrontal activity predicts memory for self. *Cerebral Cortex*, 14, 647–54.

Miall, D.S. (1986). Emotion and the self: the context of remembering. *British Journal of Psychology*, 77, 389–97.

Mitchell, J.P., Banaji, M.R., Macrae, C.N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, 18, 1306–15.

Moran, J.M., Macrae, C.N., Heatherton, T.F., Wyland, C.L., Kelley, W.M. (2006). Neuroanatomical evidence for distinct cognitive and affective components of self. *Journal of Cognitive Neuroscience*, 18, 1586–94.

Murray, S.L. (1999). The quest for conviction: motivated cognition in romantic relationships. *Psychological Inquiry*, 10, 23–34.

Murray, S.L., Rose, P., Bellavia, G.M., Holmes, J.G., Kusche, A.G. (2002). When rejection stings: how self-esteem constrains relationship-enhancement processes. *Journal of Personality and Social Psychology*, 83, 556–73.

Neff, L.A., Karney, B.R. (2005). To know you is to love you: the implications of global adoration and specific accuracy for marital relationships. *Journal of Personality and Social Psychology*, 88, 480–97.

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., Panksepp, J. (2006). Self-referential processing in our brain - A meta-analysis of imaging studies on the self. *NeuroImage*, 31, 440–57.

Ochsner, K.N., Beer, J.S., Robertson, E.R., et al. (2005). The neural correlates of direct and reflected self-knowledge. *NeuroImage*, 28, 797–814.

Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, 9, 97–113.

Pauly, K., Kircher, T., Weber, J., Schneider, F., Habel, U. (2011). Self-reference, emotion and memory performance in schizophrenia. *Psychiatry Research*, 186, 11–7.

Rabin, J.S., Gilboa, A., Stuss, D.T., Mar, R.A., Rosenbaum, R.S. (2009). Common and unique neural correlates of autobiographical memory and theory of mind. *Journal of Cognitive Neuroscience*, 22, 1095–1111.

Ravizza, S.M., McCormick, C.A., Schlerf, J.E., Justus, T., Ivry, R.B., Fiez, J.A. (2006). Cerebellar damage produces selective deficits in verbal working memory. *Brain*, 129, 306–20.

Reitan, R.M. (1958). Validity of the Trailmaking Test as an indication of organic brain damage. *Perceptual & Motor Skills*, 8, 271–6.

Richardson, M.P., Strange, B.A., Dolan, R.J. (2004). Encoding of emotional memories depends on amygdala and hippocampus and their interactions. *Nature Neuroscience*, 7, 278–85.

Rombouts, S.A.R.B., Barkhof, F., Witter, M.P., Machielsen, W.C.M., Scheltens, P. (2001). Anterior medial temporal lobe activation during attempted retrieval of encoded visuospatial scenes: an event-related fMRI Study. *NeuroImage*, 14, 67–76.

Rogers, T.B., Kuiper, N.A., Kirker, W.S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, 35, 677–88.

Royet, J.-P., Zald, D., Versace, R., et al. (2000). Emotional responses to pleasant and unpleasant olfactory, visual and auditory stimuli: a Positron Emission Tomography study. *Journal of Neuroscience*, 20, 7752–9.

Ryan, L., Nadel, L., Keil, K., et al. (2001). Hippocampal complex and retrieval of recent and very remote autobiographical memories: evidence from functional magnetic resonance imaging in neurologically intact people. *Hippocampus*, 11, 707–14.

Schacter, D.L., Wagner, A.D. (1999). Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus*, 9, 7–24.

Schmitz, T.W., Kawahara-Baccus, T.N., Johnson, S.C. (2004). Metacognitive evaluation, self-relevance, and the right prefrontal cortex. *NeuroImage*, 22, 941–7.

Sedikides, C., Green, J.D. (2000). On the self-protective nature of inconsistency-negativity management: using the person memory paradigm to examine self-referent memory. *Journal of Personality and Social Psychology*, 79, 906–22.

Somerville, L.H., Kelley, W.M., Heatherton, T.F. (2010). Self-esteem modulates medial prefrontal cortical responses to evaluative social feedback. *Cerebral Cortex*, 20, 3005–13.

Spaniol, J., Davidson, P.S.R., Kim, A.S.N., Han, H., Moscovitch, M., Grady, C.L. (2009). Event-related fMRI studies of episodic encoding and retrieval: Meta-analyses using activation likelihood estimation. *Neuropsychologia*, 47, 1765–79.

Stone, E.R., Dodrill, C.L., Johnson, N. (2001). Depressive cognition: a test of depressive realism versus negativity using general knowledge questions. *Journal of Psychology*, 135, 583–602.

Svoboda, E., McKinnon, M.C., Levine, B. (2006). The functional neuroanatomy of autobiographical memory: a meta-analysis. *Neuropsychologia*, 44, 2189–208.

Symons, C.S., Johnson, B.T. (1997). The self-reference effect in memory: a meta-analysis. *Psychological Bulletin*, 121, 371–94.

Takahashi, H., Yahata, N., Koeda, M., Matsuda, T., Asai, K., Okuboc, Y. (2004). Brain activation associated with evaluative processes of guilt and embarrassment: an fMRI study. *NeuroImage*, 23, 967–74.

Vanderwal, T., Hunyadi, E., Gruppe, D.W., Connors, C.M., Schultz, R.T. (2008). Self, mother and abstract other: an fMRI study of reflective social processing. *NeuroImage*, 41, 1437–46.

Wang, G., Mao, L., Ma, Y., et al. (2012). Neural representations of close others in collectivistic brains. *Social Cognition and Affective Neuroscience*, 7, 222–9.

Wittchen, H.U., Zaudig, M., Fydrich, T. (1997). *SKID-I/II: Strukturiertes klinisches Interview für DSM-IV*. Göttingen: Hogrefe.

Wu, T., Wang, L., Hallett, M., Chen, Y., Li, K., Chan, P. (2011). Effective connectivity of brain networks during self-initiated movement in Parkinson's disease. *NeuroImage*, 55, 204–15.

Yomogida, Y., Sugiura, M., Sassa, Y., et al. (2010). The neural basis of agency: An fMRI study. *NeuroImage*, 50, 198–207.