



HHS Public Access

Author manuscript

Soc Neurosci. Author manuscript; available in PMC 2021 April 03.

Published in final edited form as:

Soc Neurosci. 2010 ; 5(4): 335–350. doi:10.1080/17470911003615997.

Who is to blame? Neural correlates of causal attribution in social situations

Eva-Maria Seidel,

RWTH Aachen University, Aachen, and Jülich Aachen Research Alliance, Jülich, Germany

Simon B. Eickhoff,

RWTH Aachen University, Aachen, Jülich Aachen Research Alliance, and Research Center Jülich, Jülich, Germany

Thilo Kellermann,

RWTH Aachen University, Aachen, and Jülich Aachen Research Alliance, Jülich, Germany

Frank Schneider,

RWTH Aachen University, Aachen, and Jülich Aachen Research Alliance, Jülich, Germany

Ruben C. Gur,

University of Pennsylvania, and the Philadelphia Veterans Administration Medical Centre, Philadelphia, PA, USA

Ute Habel,

RWTH Aachen University, Aachen, and Jülich Aachen Research Alliance, Jülich, Germany

Birgit Derntl

RWTH Aachen University, Aachen, Jülich Aachen Research Alliance Jülich, Germany, and University of Vienna, Vienna, Austria

Abstract

In everyday life causal attribution is important in order to structure the complex world, provide explanations for events and to understand why our environment interacts with us in a particular way. This study used functional magnetic resonance imaging (fMRI) in 30 healthy subjects to separate the neural correlates of self vs. external responsibility for social events and explore the neural basis of self-serving attributions (internal attributions of positive events and external attributions of negative events). We presented short sentences describing positive and negative social events and asked participants to imagine the event, to decide the main cause and assign it to one of the categories (internal vs. external). fMRI data were analyzed using a 2×2 factorial design with the factors emotional valence and attribution. Internal compared to external attribution revealed activations along the right temporoparietal junction (TPJ). The reverse contrast showed a left lateralized network mainly involving the TPJ, the precuneus and the superior/medial frontal gyrus. These results confirmed the involvement of a fronto-temporoparietal network in differentiating self and external responsibility. Analysis of the self-serving bias yielded activation

in the dorsal anterior cingulate and in the dorsal striatum, suggesting a rewarding value of these attributions.

Keywords

Causal attribution; External; Internal; Social situation; Self-serving

INTRODUCTION

The emergent field of social cognitive neuroscience is dedicated to the investigation of various processes involved in social interactions (for a review see Lieberman, 2007). In contrast to empathy, theory of mind, and agency, the process of causal attribution has received little attention (Blackwood et al., 2003; Harris, Todorov, & Fiske, 2005; Krusemark, Campbell, & Clementz, 2008). However, inferring the cause of events especially in social situations is of high relevance due to its interaction with numerous psychological variables (cf. Terbeck, Chesterman, Fischmeister, Leodolter, & Bauer, 2008): For example, emotional reactions to positive or negative events are attenuated or increased depending on the perceived cause (Ross, McFarland, Conway, & Zanna, 1983; Weiner, Graham, & Chandler, 1982). Moreover, causal attributions strongly influence future expectations and motivations, which have been extensively studied in the context of learned helplessness (Abramson, Seligman, & Teasdale, 1978; Mikulincer, 1986).

In everyday life causal attribution is important in order to structure the complex world, provide explanations for events, and facilitate prediction (Heider, 1958; Kelley, 1967). In particular, automatically generating hypotheses about the cause of events in social situations helps us to understand why our environment interacts with us in a particular way (Weiner, 1985). Here the differentiation of internal vs. external attributions is relevant (Rotter, 1966): We ourselves can feel responsible for an incident (internal attribution) or we can attribute the cause to other persons or the situational circumstances (external). Evidently, these judgments shape and define the idea of oneself and other people, and contribute essentially to social interaction.

The classic attribution theories (Jones & Davis, 1965; Kelley, 1967) suggest that attribution is a systematic process. Contrary, more recent so called dual-process models (Gilbert, Pelham, & Krull, 1988; Trope, 1986) act on the assumption that attributional inference can also be heuristic due to limited information, time, or cognitive resources. The “self-serving-bias” (Snyder & Uranowitz, 1978) is one of these heuristic attributional biases, which is hypothesized to be rewarding, motivationally salient, and favorable to the self-concept by referring positive events to the self and negative events to external causes (Chandler, Lee, & Pengilly, 1997). Besides the differentiation of internal and external responsibility, our study was aimed at elucidating the neural correlates of this attributional heuristic.

To our knowledge there have been only a few neuroimaging studies of attributional processes, which used very different approaches: Pioneering work regarding causal attribution in social situations has been done by Blackwood et al. (2003). Contrasting internal vs. external causal attributions to mentally simulated social events, these authors

report bilateral premotor as well as cerebellar activations. The reverse contrast revealed activation in the left posterior superior temporal sulcus (pSTS). Moreover, these authors observed bilateral dorsal striatal activation to be associated with self-serving attributions, which reflects a rewarding value of these attributions (for a review see Hikosaka, Bromberg-Martin, Hong, & Matsumoto, 2008). Non-self-serving attributions have been associated with increased activation in the angular gyrus (AG), the middle temporal gyrus (mTG) and the orbitofrontal cortex (OFC; Blackwood et al., 2003). Moreover, an EEG study examining achievement-related attributions observed the medial prefrontal cortex (mPFC; Krusemark et al., 2008) to be associated with non-self-serving attributions. Activation of this network known to be involved in self-control (Amodio & Frith, 2006; Elliott, Dolan, & Frith, 2000; Ochsner et al., 2004) suggests that attributing events in a non-self-serving manner is due to a suppression process of the dominant self-enhancement tendencies. Focusing on external dispositional attribution, Harris et al. (2005) showed that additional information (low consensus, low distinctiveness, and high consistency) increased the number of dispositional attributions. These decisions were associated with activity in a widespread mainly fronto-temporal network involving the right superior temporal sulcus (STS), bilateral mPFC, right mTG and right precuneus.

Inspired by the work of Blackwood et al. (2003), we were particularly interested in causal attributions in social situations and thus applied an optimized version of the social situation paradigm in a large sample of healthy females and males. We optimized this paradigm in several regards, e.g., statistical power and robustness of parameter estimation, by increasing the number of events and participants. Moreover, we improved the generalizability of the results by measuring a far more representative sample and implementing a random effects model.

Subjects were confronted with short sentences describing positive and negative social events. Participants were asked to imagine the event happening to them, decide the main cause, and assign it to one of two categories: self (internal) or other person/situation (external). To fulfill this attributional task, first of all participants may facilitate the imagination via episodic memory retrieval. The attributional decision that follows may involve different processes involving self-reflection or self-awareness, feelings of agency, and mentalizing about self and others.

We therefore hypothesize that internal attributions recruit regions known to be involved in self-processing, particularly the mPFC (e.g., D'Argembeau et al., 2005; Fossati et al., 2003), the posterior cingulate cortex (pCC; e.g., Johnson et al., 2002; Kelley et al., 2002), and the temporoparietal junction (TPJ) (e.g., Decety & Sommerville, 2003; Vogeley et al., 2001). External attributions may be associated with regions involved in agency attributions to others, e.g., the lateral and medial inferior parietal cortex (IPC; Farrer & Frith, 2002; Farrer et al., 2003; Ruby & Decety, 2001).

Furthermore, previous studies observed that attributional style is influenced by psychopathology, e.g., an externalizing bias in deluded patients as well as a non-self-serving bias in depression (Diez-Alegria, Vazquez, Nieto-Moreno, Valiente, & Fuentenebro, 2006; Kinderman & Bentall, 1997; Mezulis, Abramson, Hyde, & Hankin, 2004). Additionally, an

association of the self-serving bias and self-esteem has been reported (Chandler et al., 1997; Diez-Alegria et al., 2006). Therefore, this study was also directed at elucidating these relations on the behavioral and, for the first time, on the neuronal level by correlating attributional decisions with questionnaire data on self-esteem and neuroticism in a healthy sample.

METHODS

Subjects

Thirty healthy subjects (15 females) aged 19–33 years (mean age 25.03 years, $SD = 4.09$) without history of neurological or psychiatric disorders participated in the study. All subjects were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971), native German speakers, and had normal or corrected-to-normal vision. They were recruited via advertisements posted at RWTH Aachen University. All subjects were paid for their participation and written informed consent was obtained. The study was approved by the local Institutional Review Board and subjects were treated according to the Declaration of Helsinki (1964) regarding treatment of human research participants.

The German Version (Borkenau & Ostendorf, 1993) of the “NEO-Five-Factor Inventory” (Costa & McCrae, 1992) was used to obtain a comprehensive personality profile of all participants. Additionally, we applied the German version (Ferring & Filipp, 1996) of the Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965) to evaluate the subjects’ self-esteem.

Moreover, all participants completed a neuropsychological test battery tapping crystallized verbal intelligence (MWT-B; Lehrl, 1996), executive functions (TMT-A/-B; Reitan, 1956) and working memory (digit span, WAIS III; Von Aster, Neubauer, & Horn, 2006), and showed average performance (MWT-B: mean PR = 78.19, $SD = 21.45$; TMT-A: mean PR = 81.55, $SD = 18.03$; TMT-B: mean PR = 86.55, $SD = 8.77$; digit span: mean PR = 53.02, $SD = 26.41$).

Experimental design

All stimuli were presented using goggles (Resonance Technology, Inc., Northridge, CA). Responses were given by button press using a Lumitouch Response System (Photon Control, Inc., Burnaby, Canada). The presentation of stimuli, and recording of responses was performed using Presentation© software package (Neurobehavioral Systems, Inc., Albany, CA).

The experimental setup is shown schematically in Figure 1. Participants were presented with short sentences describing a social event based on the IPSAQ (Kinderman & Bentall, 1996). A comparable paradigm has been used in a previous study by Blackwood et al. (2003).

In order to increase the number of unique events and hence provide a better statistical power for the fMRI analysis, we extended the 32 statements from the IPSAQ to a total of 80 short sentences. These sentences consisted of maximally nine words and described 40 positive (e.g., “A friend sent you a postcard”) and 40 negative (e.g., “A friend ignored you”) social

events. Participants were asked to imagine the event happening to them, to decide the main cause and assign the cause to one of the categories (presented below the sentence): self (internal), other person/situation (external). These stimuli together with the response categories were presented in white letters on a black screen for maximally 10 s or until a response was given, which was on average around 4 s. We reduced the stimulus presentation and therefore possible response time of 30 s in Blackwood et al. (2003) to 10 s. The rather tight time limit was introduced for two reasons: First, it ensures that subjects made a spontaneous response and did not wander off with their thoughts, which would potentially induce confounding activation of brain regions that are *per se* not related to the task. Second, it carries the advantage of eliminating those trials in which subjects apparently did not grasp the described situation or could not make up their mind.

Participants answered by button press whereupon the response categories were not associated with one specific button but changed their position in a randomized way. Presentation of each sentence was followed by a jittered inter-stimulus interval (1.9–4.3 s) with a white fixation cross on a black screen.

Behavioral data analysis

Statistical analyses were performed using SPSS 16.0 and the level of significance was set at $p = .05$. We performed 2 (valence) \times 2 (attribution) ANOVAs with repeated measurements on the number of attributional decisions and on the corresponding reaction times (RTs). For significant effects partial-eta squares are listed as estimates of effect size. Post-hoc *t*-tests were Bonferroni corrected and effect sizes are reported for significant differences (Cohen's *d*). Correlation analyses of neuroticism as well as self-esteem values and attributional data (mean number of attributional decisions for each of the four conditions)/RT data were performed using the Pearson coefficient.

FMRI acquisition parameters

Functional MR images were acquired on a 3.0 T Siemens MRI whole body scanner (SIEMENS Trio) using a standard head coil and foam padding reducing head motion. We used a gradient echo EPI sequence with the following blood-oxygen-level-dependent (BOLD) imaging parameters: TR = 2200 ms, TE = 30 ms, FoV = 200 mm, 36 slices, slice thickness = 3.1 mm, in-plane resolution = 3.1 \times 3.1 mm, flip angle = 90°, and distance factor = 15%. A single experimental session was conducted with a length of about 15 min. Additionally, a high-resolution structural image (3-D Magnetization Prepared Rapid Gradient Echo, MP-RAGE) was acquired at the end of the measurement with the following parameters: TR = 1900 ms; TE = 2.52 ms; TI = 900 ms; flip angle = 9°; 256 matrix; FoV = 250 mm; 176 slices per slab, which took 4 min.

Image preprocessing and statistical analyses

Five dummy scans before the beginning of the experiment were discarded to allow for magnetic saturation. Functional data processing was performed using the Statistical Parametric Mapping software (SPM5; Wellcome Department of Imaging Neuroscience, London) implemented in Matlab (Mathworks, Inc., Sherborn, MA, USA). Functional images were realigned to correct for head movement between scans by an affine registration

(Ashburner & Friston, 2003). Each subject's T1-scans were coregistered to the mean image of the realigned functional images. The mean functional image was subsequently normalized to the Montreal Neurological Institute (MNI) single-subject template (Collins, Neelin, Peters, & Evans, 1994; Evans et al., 1992) using linear proportions and a nonlinear sampling as derived from a segmentation algorithm (Ashburner & Friston, 2005). Normalization parameters were then applied to the functional images and coregistered to the T1-image. Images were resampled at a $1.5 \times 1.5 \times 1.5$ mm voxel size and spatially smoothed using an 8 mm full-width-at-half-maximum Gaussian kernel.

The data were analyzed using a general linear model (Kiebel & Holmes, 2003). For this event-related design, each of the four experimental conditions (internal or external attribution of positive or negative events) was modeled with a separate regressor convolved with the canonical hemodynamic response function and its first-order temporal derivative. We have chosen the onset of the stimulus (sentence and response categories) as the onset time point and the RT as the duration of the particular event. Between-scan movement parameters as estimated during spatial realignment were included as covariates of no interest. Parameter estimates were subsequently calculated for each voxel using weighted least squares to provide maximum likelihood estimates based on the non-sphericity assumption of the data (Kiebel & Holmes, 2003) in order to get identical and independently distributed error terms.

We combined external personal and external situational events to one external attribution category because 9 of our 30 participants made fewer than three external personal attributions, making powerful parameter estimation virtually impossible. For each subject, main effects were computed by applying appropriate baseline contrasts (simple effects) for each condition (i.e., Positive Internal, Positive External, Negative Internal, and Negative External). These first-level individual contrasts were then fed into a second-level group analysis using an ANOVA (factor: condition, blocking factor: subject), thus employing a random effects model (Penny & Holmes, 2003).

We used t -contrasts to analyze the effects of attributional decision, valence (presented in the supplementary material below), as well as attributional bias on the BOLD response. Therefore, we pooled the respective conditions to one internal and one external attributional condition and contrasted both by the following t -contrasts: [(Positive Internal + Negative Internal) – (Positive External + Negative External)] as well as [(Positive External + Negative External) – (Positive Internal + Negative Internal)]. Examining attributional biases, we calculated a self-serving-bias [(Positive Internal + Negative External) – (Positive External + Negative Internal)] and a non-self-serving bias t -contrast [(Positive External + Negative Internal) – (Positive Internal + Negative External)].

Additionally, we performed a correlation analysis between individual first-level contrast images and neuroticism as well as self-esteem scores of the personality questionnaires. Therefore, we used the baseline contrast images for each condition and the self-serving as well as the non-self-serving t -contrast image for each subject. We then applied a simple regression model to each of these first-level contrast images with either the corresponding neuroticism or self-esteem values.

Due to the hypothesized rewarding value of self-serving attributions we expected the striatum to be involved in this process (Blackwood et al., 2003; Hassani, Cromwell, & Schultz, 2001; Hikosaka et al., 2008; Schultz, Tremblay, & Hollerman, 2000). Therefore we performed a regions of interest (ROI) analysis (small volume correction) with the aim of maximizing sensitivity based on a template of the striatum, bilaterally, derived from the WFU PickAtlas software (Wake Forest University, Winston-Salem, NC; Maldjian, Laurienti, Kraft, & Burdette, 2003).

For statistical inference, the contrast images were thresholded at $p < .05$, FWE corrected for multiple comparisons on the cluster level (Worsley et al., 1996). The cluster forming threshold was set to $p < .001$ (uncorrected) and spatial filtering was applied to retain only those clusters that were extended enough to be significant for cluster size at $p < .05$, FWE corrected.

MNI coordinates as well as t -values are listed. MNI coordinates of the local maxima of significant activation were anatomically localized using the SPM Anatomy Toolbox (Eickhoff et al., 2005; available with all published cyto-architectonic maps from www.fz-juelich.de/ime/spm_anatomy_toolbox). Cytoarchitectonic areas are given whenever possible. As Decety and Lamm (2007) pointed out that the TPJ cannot be defined by unequivocal markers, we decided to term activations as TPJ if they include inferior parietal regions, such as the AG or the supramarginal gyrus, and posterior temporal regions at the junction with the parietal cortex.

Including gender as a factor in the analysis revealed no significant effects on the neural or behavioral level, which is in accordance with the behavioural meta-analysis of Mezulis et al. (2004).

RESULTS

Behavioral performance

Attributional decisions—The repeated measures ANOVA showed a main effect of attribution, $F(1, 29) = 14.381$, $p = .001$, partial $\eta^2 = 0.332$, with external attributions (mean = 47.7, $SD = 2.03$) having occurred significantly more often than internal attributions (mean = 32.23, $SD = 2.03$). Moreover, we demonstrated a significant valence \times attribution interaction, $F(1, 29) = 34.605$, $p < .001$, partial $\eta^2 = 0.544$, which is shown in Figure 2a.

Post-hoc Bonferroni corrected t -tests revealed a predominance, $t(29) = 7.073$, $p < .001$, $d = 2.58$, of external (mean = 28.8, $SD = 6.81$) compared to internal (mean = 11.20, $SD = 6.81$) attributions when judging the cause of negative events. However, regarding positive events, there was no significant difference, $t(29) = 0.766$, $p = .45$, between the amount of internal (mean = 21.06, $SD = 7.62$) vs. external (mean = 18.83, $SD = 7.62$) attributions. Directly comparing positive and negative events with regard to internal and external attributions, we found that internal decisions were made significantly more often when imagining a positive event and external attributions were more frequent regarding negative events, $t(29) = 5.88$, $p < .001$, $d = 1.37$.

Reaction times—The repeated measures ANOVA on the RT for attributional decisions showed a main effect of valence, $F(1, 29) = 23.594$, $p < .001$, partial $\eta^2 = 0.449$, with reactions to positive events having been faster (mean = 3.37 s, $SD = 0.68$) than reactions to negative events (mean = 3.73 s, $SD = 0.81$). There was also a significant main effect of attribution, $F(1, 29) = 5.168$, $p = .031$, partial $\eta^2 = 0.151$, with internal attributions having been faster (mean = 3.47 s, $SD = 0.75$) than external attributions (mean = 3.63, $SD = 0.75$). Moreover, we observed an interaction of valence and attribution, $F(1, 29) = 6.046$, $p = .020$, partial $\eta^2 = 0.173$, which is illustrated in Figure 2b.

Post-hoc Bonferroni corrected t -tests showed that RT of internal attributions of positive events (mean = 3.2 s, $SD = 0.63$) were faster than external attributions of negative events (mean = 3.72 s, $SD = 0.78$), $t(29) = 6.787$, $p < .001$, $d = 0.73$, and external attributions of positive events (mean = 3.54 s, $SD = 0.77$), $t(29) = 5.352$, $p < .001$, $d = 0.48$. Internal attributions of positive events were also faster compared to internal attribution of negative events (mean = 3.74 s, $SD = 0.97$), $t(29) = 4.252$, $p < .001$, $d = 0.66$.

Correlation analyses of questionnaire data—There were no significant correlations of NEO-FFI data and the number of specific attributional decisions (all p -values $> .064$) as well as the RT of attributional decisions (all p -values $> .077$). Self-esteem (RSEQ) was negatively correlated with the RT of internal attributions of negative events, $r = -0.375$, $p = .041$. There was no significant correlation with the number of specific attributional decisions and self-esteem data (all p -values $> .089$).

Functional data

Internal vs. external attributions—The internal vs. external t -contrast isolated brain areas being more active when referring the cause of social events (irrespective of valence) to oneself compared to external causes. This contrast revealed significant activations along the TPJ (areas PF, PFcm, PFm) in the right hemisphere (see Figure 3). Local maxima were found in the right superior temporal gyrus (STG) and in the right supramarginal gyrus. In the left hemisphere there was a significant activation below resting baseline limited to one local maximum in the STG (areas PPop/PF; see Table 1 for further details) with contrast estimates of the internal conditions being less negative than external conditions. We will not interpret these negative values (as deactivations), but rather all significant differences between parameter estimates that are above the resting baseline will be interpreted.

External vs. internal attributions—The reverse t -contrast revealing brain regions being more active during external causal attributions compared to internal (irrespective of valence) showed a left lateralized parietofrontal network mainly involving lateral and medial parietal areas as well as superior frontal regions (see Figure 4). There was a large cluster with local maxima in the left precuneus, spreading to another maximum in the right precuneus and more occipital to the right cuneus. Another large cluster was localized along the left TPJ (areas PGp/PGa) including local maxima in the left AG and left middle temporal gyrus (TG). The frontal activation cluster comprised local maxima in the left superior, middle, and superior medial frontal gyrus (FG). Contrarily, in the right AG (area PGp; see

Table 1 for further details) activation was significantly lower than the resting baseline with the external conditions being less negative than the internal conditions.

Self-serving attributional bias—The self-serving *t*-contrast [(Positive Internal + Negative External) – (Positive External + Negative Internal)] revealed only one significant activation cluster in the bilateral dorsal anterior cingulate cortex (dACC; $k = 355$, $p = .028$; MNI coordinates of the maximum 11/–21/45, $t = 4.62$). The contrast revealing neural processes associated with the non-self-serving bias [(Positive External + Negative Internal) – (Positive Internal + Negative External)] did not reveal any significant activation.

ROI analysis—The small volume correction based on the striatum template regarding the self-serving bias showed a higher signal of self-serving compared to non-self-serving attributions in three local maxima (see Figure 5). These maxima were located in the right (31/2/2; $k = 28$; $t = 4.76$, $p = .004$, FWE corrected) and left putamen (–26/–13/10; $k = 17$; $t = 4.95$; $p = .002$, FWE corrected) as well as in the right caudate nucleus (14/18/1; $k = 10$; $t = 4.77$; $p = .004$, FWE corrected). The reverse contrast (reflecting non-self-serving attributions) did not reveal any significant activation within this ROI.

Correlation analyses—The only significant association of the BOLD response and personality measures was a positive correlation between neuroticism and external attributions of positive events in the right putamen ($k = 333$, $p = .041$; MNI coordinates of the maximum 34/–8/3).

DISCUSSION

This study used fMRI to separate the neural correlates of self vs. external responsibility for social events and to explore the neural basis of self-serving attributions. To this end, we considerably optimized a previously applied paradigm (Blackwood et al., 2003). Our results are consistent with prior evidence of a fronto-temporoparietal network differentiating self vs. external responsibility (Farrer & Frith, 2002; Ruby & Decety, 2001; Seger, Stone, & Keenan, 2004). However, we did not replicate the findings of Blackwood et al. (2003) concerning the self vs. other distinction, although they applied a comparable paradigm. This might be due to our improvements regarding the generalizability of our results (large sample size, random effects analysis) as well as the increase of statistical power by enhancing the number of events. Nevertheless, the self-serving bias still activated the dorsal striatum in both studies.

Self vs. external responsibility

Directly comparison showed that activation in the right TPJ, including the STG and the supramarginal gyrus, is more associated with internal than with external attributions. There is evidence for a general right hemispheric dominance in self representation (reviewed by Decety & Sommerville, 2003), such as self-face recognition (Uddin, Kaplan, Molnar-Szakacs, Zaidel, & Iacoboni, 2005), imitation (Decety, Chaminade, Grezes, & Meltzoff, 2002) and agency (Chaminade & Decety, 2002). Lesions within the right TPJ cause deficits in self-related processes, e.g., self-face recognition (e.g., Breen, Caine, & Coltheart, 2001; Uddin, Molnar-Szakacs, Zaidel, & Iacoboni, or body-schema disturbances (e.g., Blanke &

Arzy, 2005; Giummarra, Gibson, Georgiou-Karistianis, & Bradshaw, 2008). The recruitment of the right TPJ in our self-referential judgments also fits a finding of Vogeley et al. (2001), who applied so-called self-ascription stories. Integrating our results of a right lateralized TPJ activation during internal causal attribution with previous data emphasizes the fundamental role of this region in shaping the self-concept.

Contrarily, external compared to internal attribution showed higher signal in the bilateral precuneus, the left TPJ (including the AG and the middle TG) and the left superior and medial FG. This widespread network has been observed to be involved in external attribution by previous studies using different paradigms tapping other vs. self processing, e.g., motor imagery (Ruby & Decety, 2001), food preference judgements (Seger et al., 2004), agency (Farrer & Frith, 2002) or action awareness (Farrer et al., 2008). Basically, these tasks and our paradigm differed concerning the processing mechanisms involved, e.g. detecting discrepancy in sensory input and perspective taking. However, the common process engaged is the differentiation between self and other.

Studies investigating self-processing without direct comparison to a third person reported the precuneus to be crucial for self-related mental operations (e.g., Kjaer, Nowak, & Lou, 2002; Lou et al., 2004). However, studies implementing first vs. third person tasks found the precuneus to be more active in the third person condition (Farrer & Frith, 2002; Ruby & Decety, 2001; Seger et al., 2004). Moreover, lesions within the left upper medial FG (Brainin, Seiser, & Matz, 2008) but also within the IPC (e.g., Bundick & Spinella, 2000; Gondim, Oliveira, & Cruz-Flores, 2005) have been associated with the so-called alien hand syndrome, which is characterized by perceiving one limb as belonging to or being controlled by others. These results underline the assumption that this network (especially the precuneus) is important in differentiating the self from others or the external world, which is further supported by our data on causal attributions.

There is also converging evidence that the precuneus (reviewed by Cavanna & Trimble, 2006), the TPJ (Maguire & Mummery, 1999) and the superior FG (Abraham, Schubotz, & von Cramon, 2008) are involved in autobiographic memory processes. The left AG seems to establish an “online” representation of episodic information (Vilberg & Rugg, 2009) and to be recruited during recollection as well as prospection of personal events (Abraham et al., 2008). The superior FG has been shown to be responsible for recombination processes of prospection and memory of personal events (Abraham et al., 2008). These memory-related operations seem to be core processes when one is establishing an attributional decision. Consequently, our data suggest that autobiographical memory is required more during external as compared to internal causal attribution. This speculation is supported by the well established “self-reference effect” leading to an eased retrieval of self-relevant information (Symons & Johnson, 1997).

Generally, we conclude that attributing causes for social events somehow relies on personal memories. However, there must be something special about the other vs. self distinction within this neural network since its activation has also been reported by Farrer and Frith (2002) and Farrer et al. (2008) using simple motor tasks without any memory requirements. Furthermore, this assumption is also supported by studies examining conceptual and

emotional perspective-taking (Ruby & Decety, 2003, 2004). Contrasting the other vs. the self-perspective revealed a fronto-temporoparietal network that widely overlapped with our findings. However, our findings differ from the results of Blackwood et al. (2003) despite the comparable paradigm, which might be due to some improvements in our study design (e.g., sample size, number of events).

Taking our findings and previous reports together, the IPC seems to play a crucial role in differentiating self from other. However, results of previous studies are inconsistent regarding the lateralization of the self vs. the other condition. Two previous studies (Chaminade & Decety, 2002; Decety et al. 2002) found the same lateralization pattern, we reported, such that the other condition is associated with a left lateralized IPC activation and the self condition is right lateralized. However, Ruby and Decety (2001) report an inverse lateralization effect. Farrer and Frith (2002) as well as Farrer et al. (2003) reported bilateral IPC for the other condition; however, they did not find any IPC activation to be associated with the self condition but instead reported insula activation. Adding to the complexity, the tasks are very different, ranging from imagining oneself and another person performing an action (Ruby & Decety, 2001), to imitating another person vs. being imitated by the other person (Decety et al., 2002), to simple visuomotor-feedback incongruency tasks (e.g., Farrer et al., 2003). Our task adds a social situation setting to this self vs. other comparison.

Beyond the general involvement of the IPC in agency decisions, we observed that two parts of the IPC can be separated with regard to the self vs. other differentiation. Whereas the supramarginal gyrus was more involved in internal attributions, the AG was more activated during external attributions. Adding to the inconsistencies on lateralization effects, there is also a heterogeneous picture regarding the involvement of the AG (Farrer & Frith, 2002; Farrer et al., 2003; Ruby & Decety, 2001) and the supramarginal gyrus (e.g., Chaminade & Decety, 2002; Decety et al., 2002; Ruby & Decety, 2001) in the differentiation of self vs. other. Therefore, we regard our finding as a hint of a functional differentiation and lateralization within the inferior parietal cortex regarding the self vs. other distinction that awaits future replication in comparably large samples.

Integrating our results in the X- and C-system framework of social cognition postulated by Lieberman (2007), internal attributions are associated with the reflexive, automatic X-system whereas all three activation clusters specific for external attributions lie within the reflective, regulatory C-system. This approach suggests that internal attribution is a more automatic process and external attribution involves more intentional and controlled mechanisms. Similarly, Seger et al. (2004) interpreted the superior FG activation together with longer RT as a hint for other related decisions to be more complicated compared to self-referential judgments.

Altogether, our data showed that internal attribution results in a circumscribed activation cluster whereas external attribution was associated with a more extended network. Likewise, previous studies inducing discrepancy in action effects (Spengler, von Cramon, & Brass, 2009) or imagining harmful actions (Kedia, Berthoz, Wessa, Hilton, & Martinot, 2008) reported that the other condition showed a higher signal in a widespread fronto-temporoparietal activation whereas there were no suprathreshold clusters associated with the

self-condition. Ruby and Decety (2001) concluded from their findings that a vivid representation of the self is important in order to discriminate between self and other, which is supported by results of comparable studies reporting overlapping activations during self and other judgments (e.g., Decety & Sommerville, 2003; Ochsner et al., 2005; Seger et al., 2004). We hypothesized to observe significant activation associated with the internal vs. external contrast in the mPFC as well as pCC based on previous studies on self processing (D'Argembeau et al., 2005; Fossati et al., 2003; Johnson et al., 2002; Kelley et al., 2002), which was not supported by our data. However, a closer look at the tasks applied revealed that they differ regarding the other condition. In our task this condition referred to a friend whereas other studies used a famous reference person (D'Argembeau et al., 2005; Kelley et al., 2002) or a semantic control condition (Fossati et al., 2003; Johnson et al., 2002). Apart from an activation overlap due to the discrimination process between self and other responsibility (Ruby & Decety, 2001), the lacking mPFC and pCC activation in our study might be due to these task differences.

Self-serving bias

Analysis of behavioral data confirmed our expectations of a high frequency of self-serving attributions, which were based on previous studies in healthy samples (cf. meta-analysis by Mezulis et al., 2004). Our RT data revealed that self-serving attributions are the fastest and internal attributions of negative events the slowest attributions. These findings reflect a dominance and automaticity of self-serving attributions on the behavioral level.

Our whole-brain analysis of self-serving compared to non-self-serving attributions revealed bilateral dACC activation. Previous studies (Holroyd & Coles, 2008; Santesso et al., 2008) reported the dACC to be engaged in reinforcement learning and especially in monitoring behavior and selecting the most appropriate action alternatives. The activation of the dACC in self-serving attributional decisions reflects a selection process choosing attributions that are more favorable than others. Moreover, due to the faster RT of self-serving compared with non-self-serving attributions, we suppose that this selection process acts rather automatically.

Consistent with our hypothesis and the results of Blackwood et al. (2003), the ROI analysis in the striatum showed self-serving specific activity in the dorsal part of the striatum. The ventral striatum has long been thought to be the main reward processing node, but recent findings suggest that activity within the ventral striatum is associated with craving (Delgado, 2007), reward anticipation (Knutson, Fong, Bennett, Adams, & Homme, 2003) and is more sensitive to short-term rewards (Tanaka et al., 2007). The dorsal striatum has recently been assumed to guide motivated behavior and select actions that have a high (long-term) reward value (reviewed by Hikosaka et al., 2008; Tanaka et al., 2007). Activation of the dorsal reward system in our data suggests that attributing social events in a self-serving manner has a rewarding effect. Additionally, dorsal striatal activation and the involvement of the dACC reflects that choosing a self-serving attribution is a selection of the highest rewarding alternative.

An attributional decision is not rewarding *per se*, for example in terms of money. The absence of activation of other parts of the reward circuit (amygdala, ACC, insula, mPFC,

OFC) that have been observed using directly rewarding stimuli (for reviews see Berridge & Kringelbach, 2008; Leknes & Tracey, 2008; O'Doherty, 2004), argues for a more indirect and delayed rewarding value of self-serving attributions. It can be speculated that a learning process turns these attributions to rewarding elements. Consequently, this learning process leads to an automation of self-serving attributions which is reflected in our RT data.

However, one has to keep in mind that we did not observe an association of self-esteem and attributions on the neuronal or on the behavioral level, which could be due to our homogeneous sample with an average self-esteem (Ferring & Frith, 1996). The sole hint of a connection was a negative correlation with RT of internal attributions of negative events.

Moreover, our correlation analysis revealed a first hint of an association of attributing positive events to external causes (i.e. non-self-serving attributions) and neuroticism values in the right putamen. This association with the reward system seems interesting, but discrepant. High neuroticism values describe persons that are anxious, depressive, insecure, timid, and shy. Our “neurotic” participants perceived positive social events as rewarding even when they attributed them to external causes. We are currently not able to resolve this discrepancy and propose further examination. This may have implications for the understanding of depressive cognition, which is characterized by a non-self-serving bias, e.g., attributing positive events to external causes (Diez-Algeria et al., 2006). However, it is important to note that this study only gathered neuroticism scores in a healthy student sample which cannot be exactly equated with depressive tendencies, but is considered to be a marker of vulnerability for depression (e.g., Haas, Constable, & Canli, 2008). In future studies it would be more valuable to assess depressive symptom severity directly as well as its correlation with the neural correlates of biased attribution in a sample of depressed patients.

Limitations

Our behavioral data analysis revealed that external attributions were more frequent than internal attributions, and mainly triggered by negative events. Concerning RT, we observed that generally internal attributions were faster than external attributions and reactions to positive events were faster than reactions to negative events, mainly triggered by internal attributions of positive events, which were the fastest decisions. These behavioral effects might have had an impact on our imaging data. Regarding our design, it is important to differentiate between differences in the number of attributions, which affect the number of events for the fMRI analysis, and effects on RT data, which affect the duration of these events. Spending less time on one condition would probably result in a weaker BOLD response, which would be the case for internal attributions of positive events. However, an increased number of events (external attributions for negative events) would improve statistical power. This leaves us with the problem of differentiating between the effects of weaker BOLD response in one condition and more statistical power in another condition.

We did not include these two conditions separately in our main contrasts (internal > external; external > internal) but created a mean for internal and external attributions. Moreover, internal attributions of positive events and external attributions of negative events constitute the self-serving-bias, a major aim of this study, which we analyzed with a separate *t*-contrast

[(Positive Internal + Negative External) – (Positive External + Negative Internal)]. Thus, we pooled these two conditions to one self-serving condition and analyzed the mean of this effect. We were interested in the neural process underlying this typical attributional bias in healthy subjects, which is reflected in our behavioral data. Therefore, we accepted this possible drawback in our study design.

CONCLUSION

The present study explored the neural correlates of attributing responsibility for social events to the self vs. external causes. Internal attributions were associated with a focal cluster within the right TPJ. External attributions recruited a more widespread left-lateralized fronto-temporoparietal network, which confirmed the previously reported involvement of this neural network in differentiating self and external responsibility. In summary, external attribution seems to rely more on the retrieval of personal memories and to recruit more controlled cognitive mechanisms, whereas internal attribution seems to proceed rather automatically. Nevertheless, our data and prior results suggest that there is a considerable overlap between neural and cognitive processes underlying self vs. other judgements, with a vivid representation of the self being important in order to establish a differentiation. Taking a closer look at the so-called self-serving bias, our data reflect a dominance and automaticity of self-serving attributions on the behavioral level. Activation of the dACC and the dorsal striatum during self-serving attributions suggest that these decisions have a rewarding effect that is due to a reinforcement learning process.

SUPPLEMENTARY MATERIAL

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was funded by the Medical Faculty of RWTH Aachen University (START 690811). EMS was supported by the Interdisciplinary Centre for Clinical Research (IZKF) within the Faculty of Medicine at the RWTH Aachen University (NWW11-SP3). SBE was supported by the Human Brain Project (R01-MH074457-01A1) and the Helmholtz-Initiative on Systems Biology. BD and UH were supported by the German Research Foundation (DFG, IRTG 1328) and RCG was supported by NIMH grant MH 60722.

REFERENCES

- Abraham A, Schubotz RI, & von Cramon DY (2008). Thinking about the future versus the past in personal and non-personal contexts. *Brain Research*, 1233, 106–119. [PubMed: 18703030]
- Abramson LY, Seligman MEP, & Teasdale JD (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology*, 87(1), 49–74. [PubMed: 649856]
- Amodio DM, & Frith CD (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7(4), 268–277. [PubMed: 16552413]
- Ashburner J, & Friston KJ (2003). Rigid body registration. In Frackowiak RS, Friston KJ, Frith C, Dolan R, Price CJ, Zeki S, et al. (Eds.), *Human brain function* (pp. 635–655). London: Academic Press.
- Ashburner J, & Friston KJ (2005). Unified segmentation. *NeuroImage*, 26(3), 839–851. [PubMed: 15955494]
- Berridge KC, & Kringelbach ML (2008). Affective neuroscience of pleasure: Reward in humans and animals. *Psychopharmacology*, 199(3), 457–480. [PubMed: 18311558]

- Blackwood NJ, Bentall RP, Ffytche DH, Simmons A, Murray RM, & Howard RJ (2003). Self-responsibility and the self-serving bias: An fMRI investigation of causal attributions. *NeuroImage*, 20(2), 1076–1085. [PubMed: 14568477]
- Blanke O, & Arzy S (2005). The out-of-body experience: Disturbed self-processing at the temporo-parietal junction. *The Neuroscientist*, 11(1), 16–24. [PubMed: 15632275]
- Borkenau P, & Ostendorf F (1993). NEO-Fünf-Faktoren-Inventar (NEO-FFI) nach Costa und McCrae. Göttingen, Germany: Hogrefe.
- Brainin M, Seiser A, & Matz K (2008). The mirror world of motor inhibition: The alien hand syndrome in chronic stroke. *Journal of Neurology, Neurosurgery and Psychiatry*, 79(3), 246–252.
- Breen N, Caine D, & Coltheart M (2001). Mirrored-self misidentification: Two cases of focal onset dementia. *Neurocase*, 7(3), 239–254. [PubMed: 11459919]
- Bundick T, & Spinella M (2000). Subjective experience, involuntary movement, and posterior alien hand syndrome. *Journal of Neurology, Neurosurgery and Psychiatry*, 68(1), 83–85.
- Cavanna AE, & Trimble MR (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain*, 129, 564–583. [PubMed: 16399806]
- Chaminade T, & Decety J (2002). Leader or follower? Involvement of the inferior parietal lobe in agency. *NeuroReport*, 15, 1975–1978.
- Chandler TA, Lee MS, & Pengilly JW (1997). Self-esteem and causal attributions. *Genetic Social and General Psychology Monographs*, 123(4), 479–491.
- Collins DL, Neelin P, Peters TM, & Evans AC (1994). Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, 18(2), 192–205. [PubMed: 8126267]
- Costa PT, & McCrae RR (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual. Odessa, FL: Psychological Assessment Resources.
- D'Argembeau A, Collette F, Van der Linden M, Laureys S, Del Fiore G, Degueldre C, et al. (2005). Self-referential reflective activity and its relationship with rest: A PET study. *NeuroImage*, 25(2), 616–624. [PubMed: 15784441]
- Decety J, Chaminade T, Grezes J, & Meltzoff AN (2002). A PET exploration of the neural mechanisms involved in reciprocal imitation. *NeuroImage*, 15, 265–272. [PubMed: 11771994]
- Decety J, & Lamm C (2007). The role of the right temporo-parietal junction in social interaction: How low-level computational processes contribute to meta-cognition. *The Neuroscientist*, 13(6), 580–593. [PubMed: 17911216]
- Decety J, & Sommerville JA (2003). Shared representations between self and other: A social cognitive neuroscience view. *Trends in Cognitive Sciences*, 7(12), 527–533. [PubMed: 14643368]
- Delgado MR (2007). Reward-related responses in the human striatum. *Annals of the New York Academy of Science*, 1104, 70–88.
- Diez-Alegria C, Vazquez C, Nieto-Moreno M, Valiente C, & Fuentenebro F (2006). Personalizing and externalizing biases in deluded and depressed patients: Are attributional biases a stable and specific characteristic of delusions? *British Journal of Clinical Psychology*, 45, 531–544.
- Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, 25(4), 1325–1335. [PubMed: 15850749]
- Elliott R, Dolan RJ, & Frith CD (2000). Dissociable functions in the medial and lateral orbitofrontal cortex: Evidence from human neuroimaging studies. *Cerebral Cortex*, 10(3), 308–317. [PubMed: 10731225]
- Evans AC, Marrett S, Neelin P, Collins L, Worsley K, Dai W, et al. (1992). Anatomical mapping of functional activation in stereotactic coordinate space. *NeuroImage*, 1(1), 43–53. [PubMed: 9343556]
- Farrer C, Franck N, Georgieff N, Frith CD, Decety J, & Jeannerod A (2003). Modulating the experience of agency: A positron emission tomography study. *NeuroImage*, 18(2), 324–333. [PubMed: 12595186]
- Farrer C, Frey SH, Van Horn JD, Tunik E, Turk D, Inati S, et al. (2008). The angular gyrus computes action awareness representations. *Cerebral Cortex*, 18(2), 254–261. [PubMed: 17490989]

- Farrer C, & Frith CD (2002). Experiencing oneself vs another person as being the cause of an action: The neural correlates of the experience of agency. *NeuroImage*, 15(3), 596–603. [PubMed: 11848702]
- Ferring D, & Filipp S-H (1996). Messung des Selbstwertgefühls: Befunde zur Reliabilität, Validität und Stabilität der Rosenberg-Skala. *Diagnostica*, 42, 284–292.
- Fossati P, Hevenor SJ, Graham SJ, Grady C, Keightley ML, Craik F, et al. (2003). In search of the emotional self: An fMRI study using positive and negative emotional words. *American Journal of Psychiatry*, 160(11), 1938–1945.
- Gilbert DT, Pelham BW, & Krull DS (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology*, 54(5), 733–740.
- Giummarra MJ, Gibson SJ, Georgiou-Karistianis N, & Bradshaw JL (2008). Mechanisms underlying embodiment, disembodiment and loss of embodiment. *Neuroscience and Biobehavioral Reviews*, 32(1), 143–160. [PubMed: 17707508]
- Gondim FAA, Oliveira GR, & Cruz-Flores S (2005). Position-dependent levitation of the dominant arm after left parietal stroke: An unreported feature of posterior alien limb syndrome? *Movement Disorders*, 20(5), 632–633. [PubMed: 15726577]
- Haas BW, Constable RT, & Canli T (2008). Stop the sadness: Neuroticism is associated with sustained medial prefrontal cortex response emotional facial expressions. *NeuroImage*, 42, 385–392. [PubMed: 18511299]
- Harris LT, Todorov A, & Fiske ST (2005). Attributions on the brain: Neuro-imaging dispositional inferences, beyond theory of mind. *NeuroImage*, 28(4), 763–769. [PubMed: 16046148]
- Hassani OK, Cromwell HC, & Schultz W (2001). Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *Journal of Neurophysiology*, 85(6), 2477–2489. [PubMed: 11387394]
- Heider F (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Hikosaka O, Bromberg-Martin E, Hong S, & Matsumoto M (2008). New insights on the subcortical representation of reward. *Current Opinion in Neurobiology*, 18(2), 203–208. [PubMed: 18674617]
- Holroyd CB, & Coles MGH (2008). Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behaviour. *Cortex*, 44(5), 548–559. [PubMed: 18387587]
- Johnson SC, Baxter LC, Wilder LS, Pipe JG, Heiserman JE, & Prigatano GP (2002). Neural correlates of self-reflection. *Brain*, 125, 1808–1814. [PubMed: 12135971]
- Jones EE, & Davis KE (1965). From acts to dispositions: The attribution process in person perception. *Advances in Experimental Social Psychology*, 2(4), 219–266.
- Kedia G, Berthoz S, Wessa M, Hilton D, & Martinot JL (2008). An agent harms a victim: A functional neuroimaging study on specific moral emotions. *Journal of Cognitive Neuroscience*, 20(10), 1788–1798. [PubMed: 18211233]
- Kelley HH (1967). Attribution theory in social psychology. In Levine D (Ed.), *Nebraska symposium on motivation*. Lincoln, NE: University of Nebraska Press.
- Kelley WM, Macrae CN, Wyland CL, Caglar S, Inati S, & Heatherton TF (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, 14(5), 785–794. [PubMed: 12167262]
- Kiebel S, & Holmes AP (2003). The general linear model. In Frackowiak RSJ, Friston KJ, Frith C, Dolan R, Price CJ, Zeki S, et al. (Eds.), *Human brain function* (pp. 725–760). London: Academic Press.
- Kinderman P, & Bentall RP (1996). A new measure of causal locus: The internal, personal and situational attributions questionnaire. *Personality and Individual Differences*, 20(2), 261–264.
- Kinderman PB, & Bentall RP (1997). Causal attributions in paranoia and depression: Internal, personal and situational attributions for negative events. *Journal of Abnormal Psychology*, 106(2), 341–345. [PubMed: 9131855]
- Kjaer TW, Nowak M, & Lou HC (2002). Reflective self-awareness and conscious states: PET evidence for a common midline parietofrontal core. *NeuroImage*, 17(2), 1080–1086. [PubMed: 12377180]
- Knutson B, Fong GW, Bennett SM, Adams CM, & Homme D (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: Characterization with rapid event-related fMRI. *NeuroImage*, 18(2), 263–272. [PubMed: 12595181]

- Krusemark EA, Campbell WK, & Clementz BA (2008). Attributions, deception, and event related potentials: An investigation of the self-serving bias. *Psychophysiology*, 45(4), 511–515. [PubMed: 18282197]
- Lehrl S (1996). Der MWT– ein Intelligenztest für die ärztliche Praxis. *Praxis für Neurologie und Psychiatrie*, 7, 488–491.
- Leknes S, & Tracey I (2008). Science & society: A common neurobiology for pain and pleasure. *Nature Reviews Neuroscience*, 9(4), 314–320. [PubMed: 18354400]
- Lieberman MD (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, 58, 259–289.
- Lou HC, Luber B, Crupain M, Keenan JP, Nowak M, Kjaer TW, et al. (2004). Parietal cortex and representation of the mental self. *Proceedings of the National Academy of Sciences of the United States of America*, 101(17), 6827–6832. [PubMed: 15096584]
- Maguire EA, & Mummery CJ (1999). Differential modulation of a common memory retrieval network revealed by positron emission tomography. *Hippocampus*, 9(1), 54–61. [PubMed: 10088900]
- Maldjian JA, Laurienti PJ, Kraft RA, & Burdette JH (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19(3), 1233–1239. [PubMed: 12880848]
- Mezulis AH, Abramson LY, Hyde JS, & Hankin BL (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin*, 130(5), 711–747. [PubMed: 15367078]
- Mikulincer M (1986). Attributional processes in the learned helplessness paradigm: Behavioral effects of global attributions. *Journal of Personality and Social Psychology*, 51(6), 1248–1256. [PubMed: 3806360]
- Ochsner KN, Beer JS, Robertson ER, Cooper JC, Gabrieli JDE, Kihlstrom JF, et al. (2005). The neural correlates of direct and reflected self-knowledge. *NeuroImage*, 28(4), 797–814. [PubMed: 16290016]
- Ochsner KN, Knierim K, Ludlow DH, Hanelin J, Ramachandran T, Glover G, et al. (2004). Reflecting upon feelings: An fMRI study of neural systems supporting the attribution of emotion to self and other. *Journal of Cognitive Neuroscience*, 16(10), 1746–1772. [PubMed: 15701226]
- O’Doherty JP (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current Opinion in Neurobiology*, 14(6), 769–776. [PubMed: 15582382]
- Oldfield RC (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologica*, 9, 97–113.
- Penny WD, & Holmes AP (2003). Random effects analysis. In Frackowiak RSJ, Friston KJ, Frith C, Dolan R, Price CJ, Zeki S, et al. (Eds.), *Human brain function* (pp. 843–850). London: Academic Press.
- Reitan RM (1956). *Trail Making Test: Manual for administration, scoring and interpretation*. South Tucson, AZ: Reitan Neuropsychology Laboratory.
- Rosenberg M (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Ross M, McFarland C, Conway M, & Zanna MP (1983). Reciprocal relation between attitudes and behavior recall: Committing people to newly formed attitudes. *Journal of Personality and Social Psychology*, 45(2), 257–267.
- Rotter JB (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80(1), 1–28.
- Ruby P, & Decety J (2001). Effect of subjective perspective taking during simulation of action: A PET investigation of agency. *Nature Neuroscience*, 4(5), 546–550. [PubMed: 11319565]
- Ruby P, & Decety J (2003). What you believe versus what you think they believe: A neuroimaging study of conceptual perspective-taking. *European Journal of Neuroscience*, 17(11), 2475–2480.
- Ruby P, & Decety J (2004). How would you feel versus how do you think she would feel? A neuroimaging study of perspective-taking with social emotions. *Journal of Cognitive Neuroscience*, 16(6), 988–999. [PubMed: 15298786]
- Santesso DL, Dillon DG, Birk JL, Holmes AJ, Goetz E, Bogdan R, et al. (2008). Individual differences in reinforcement learning: Behavioral, electrophysiological, and neuroimaging correlates. *NeuroImage*, 42(2), 807–816. [PubMed: 18595740]

- Schultz W, Tremblay L, & Hollerman JR (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral Cortex*, 10(3), 272–283. [PubMed: 10731222]
- Seger CA, Stone M, & Keenan JP (2004). Cortical activations during judgments about the self and an other person. *Neuropsychologia*, 42(9), 1168–1177. [PubMed: 15178169]
- Snyder M, & Uranowitz SW (1978). Reconstructing the past: Some cognitive consequences of person perception. *Journal of Personality and Social Psychology*, 36(9), 941–950.
- Spengler S, von Cramon DY, & Brass M (2009). Was it me or was it you? How the sense of agency originates from ideomotor learning revealed by fMRI. *NeuroImage*, 46(1), 290–298. [PubMed: 19457378]
- Symons CS, & Johnson BT (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, 121(3), 371–394. [PubMed: 9136641]
- Tanaka S, Schweighofer N, Asahi S, Shishida K, Okamoto Y, Yamawaki S, et al. (2007). Serotonin differentially regulates short- and long-term prediction of rewards in the ventral and dorsal striatum. *PLoS ONE*, 2(12), 1–7.
- Terbeck S, Chesterman P, Fischmeister FPS, Leodolter U, & Bauer H (2008). Attribution and social cognitive neuroscience: A new approach for the “online-assessment” of causality ascriptions and their emotional consequences. *Journal of Neuroscience Methods*, 173(1), 13–19. [PubMed: 18599124]
- Trope Y (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, 93(3), 239–257.
- Uddin LQ, Kaplan JT, Molnar-Szakacs I, Zaidel E, & Iacoboni M (2005). Self-face recognition activates a fronto-parietal “mirror” network in the right hemisphere: An event-related fMRI study. *NeuroImage*, 25, 926–935. [PubMed: 15808992]
- Uddin LQ, Molnar-Szakacs I, Zaidel E, & Iacoboni M (2006). rTMS to the right inferior parietal lobule disrupts self–other discrimination. *Social Cognitive and Affective Neuroscience*, 1, 65–71. [PubMed: 17387382]
- Vilberg KL, & Rugg MD (2009). Functional significance of retrieval-related activity in lateral parietal cortex: Evidence from fMRI and ERPs. *Human Brain Mapping*, 30(5), 1490–1501. [PubMed: 18649352]
- Vogel K, Bussfeld P, Newen A, Herrmann S, Happe F, Falkai P, et al. (2001). Mind reading: Neural mechanisms of theory of mind and self-perspective. *NeuroImage*, 14(1), 170–181. [PubMed: 11525326]
- Von Aster M, Neubauer A, & Horn R (2006). *Hamburg-Wechsler-Intelligenz-Test für Erwachsene III*. Frankfurt, Germany: Harcourt.
- Weiner B (1985). Spontaneous causal thinking. *Psychological Bulletin*, 97(1), 74–84. [PubMed: 3983298]
- Weiner B, Graham S, & Chandler C (1982). Pity, anger, and guilt: An attributional analysis. *Personality and Social Psychology Bulletin*, 8(2), 226–232.
- Worsley KJ, Marret S, Neelin P, Vandal AC, Friston KJ & Evans AC (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, 4, 58–73. [PubMed: 20408186]

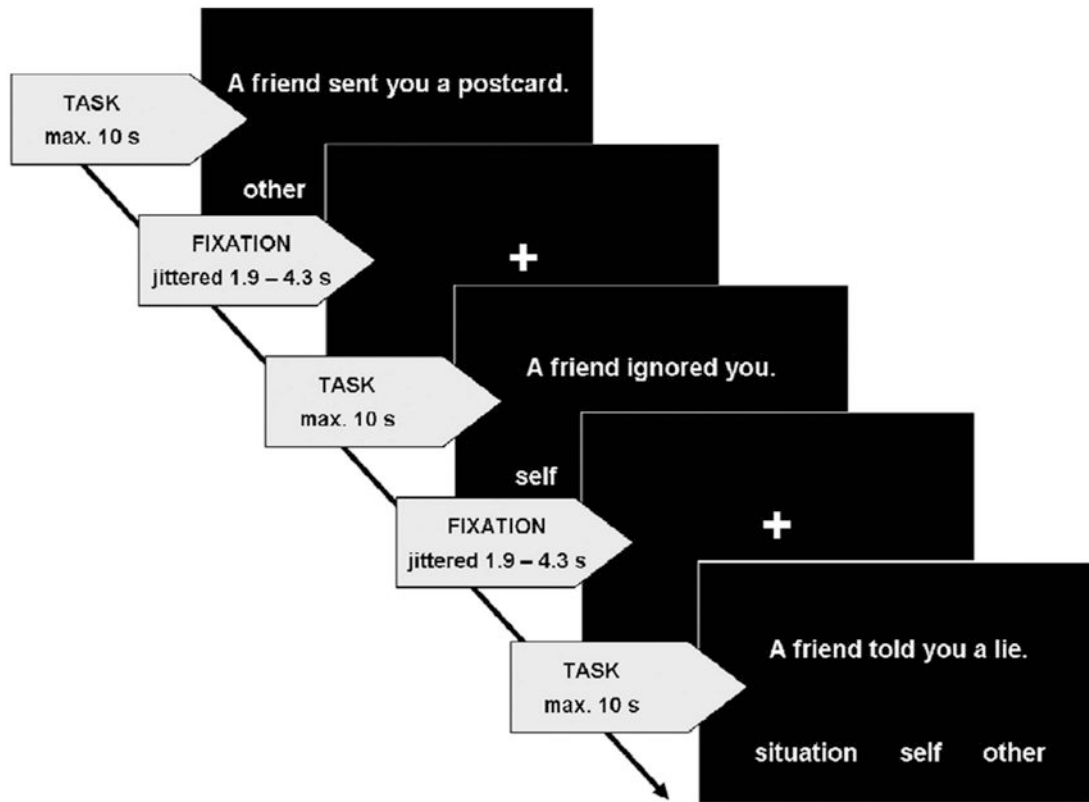


Figure 1.
Visualization of the fMRI paradigm.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

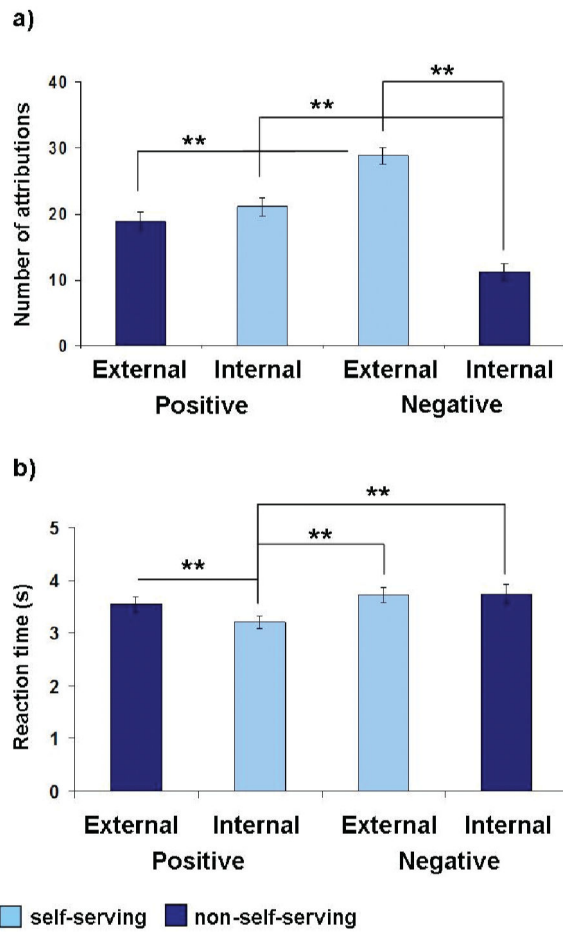


Figure 2. Behavioral results: Main effect of attribution and valence by attribution interaction regarding the mean number of attributional decisions (a) and main effect of valence and valence by attribution interaction on the mean reaction times (b).

INTERNAL > EXTERNAL

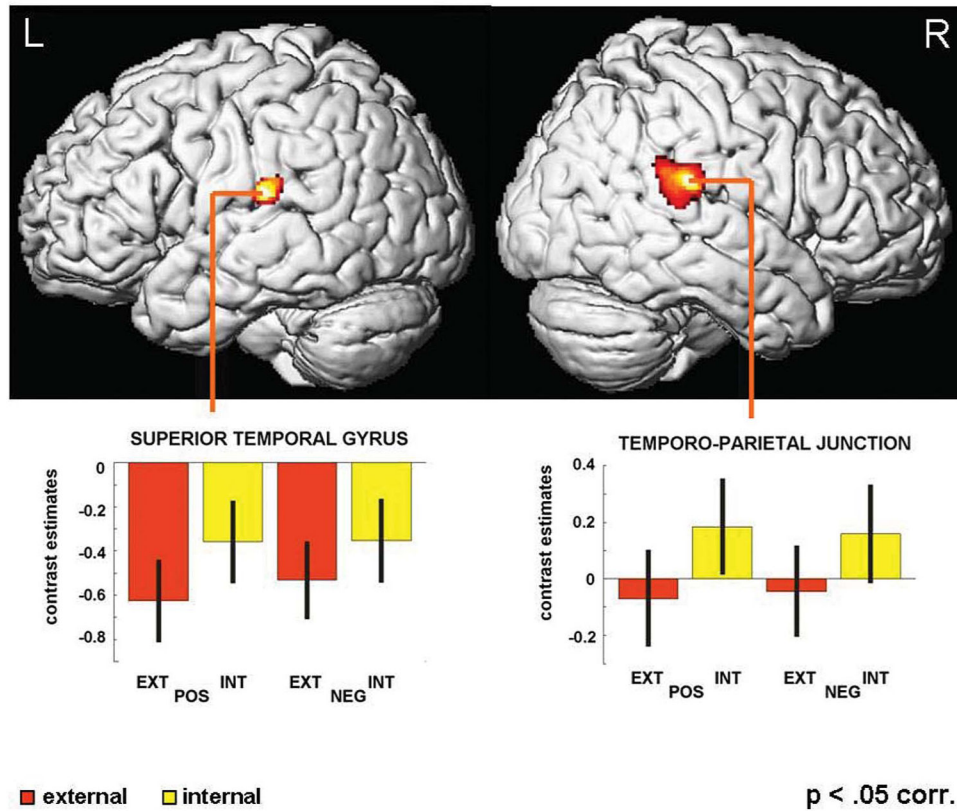


Figure 3. Results of the whole-brain analysis contrasting internal vs. external attributional decisions, revealing differential activity in the bilateral STG.

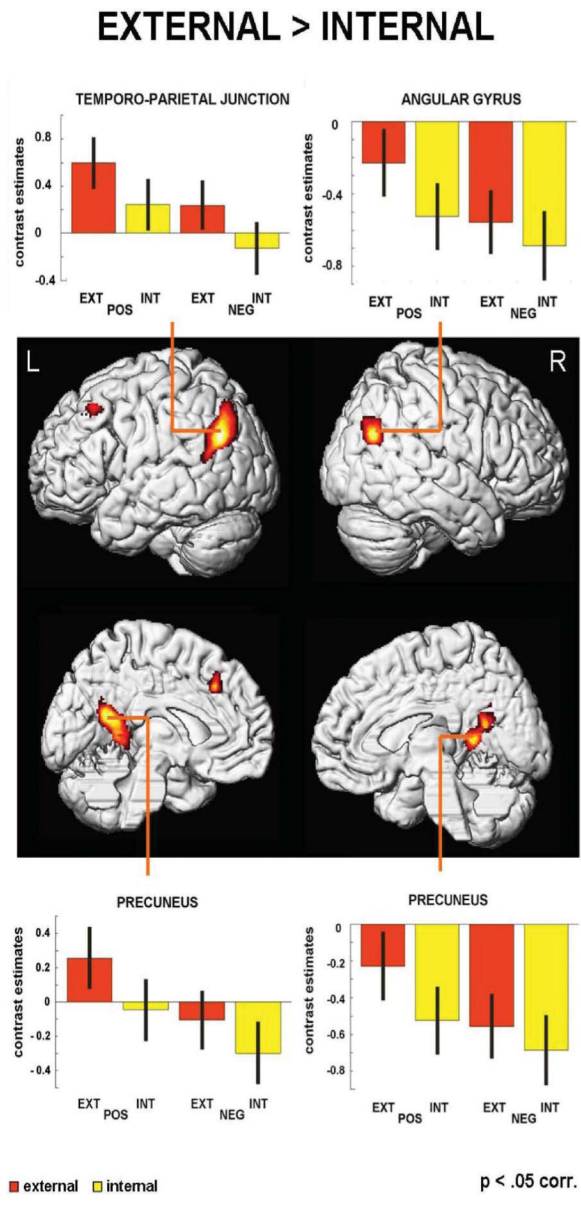


Figure 4. Whole-brain analysis of external vs. internal attributional decisions with differential activity in lateral and medial IPC as well as superior FG.

SELF-SERVING BIAS STRIATUM

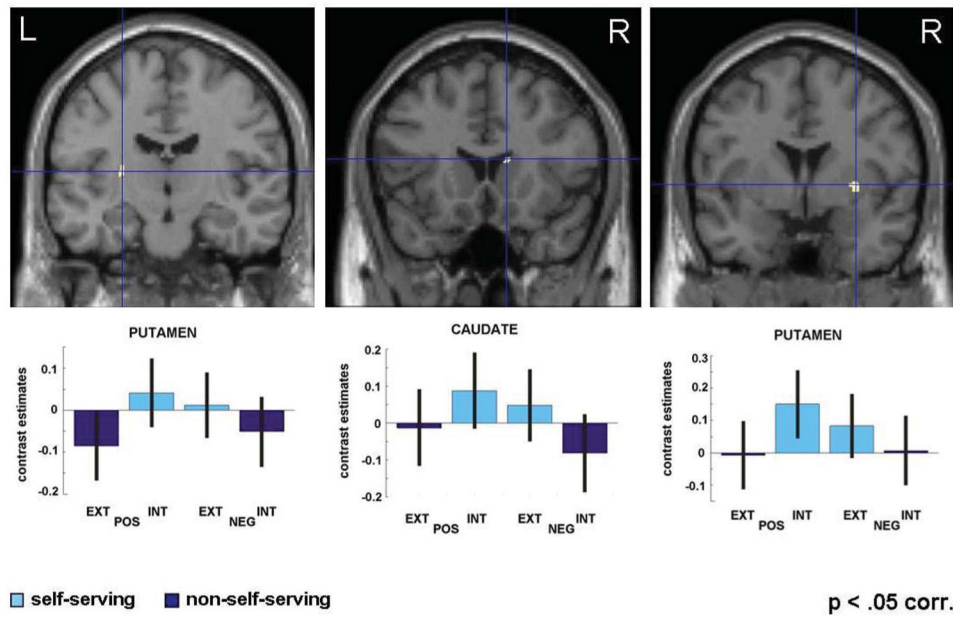


Figure 5. ROI analysis in the striatum revealed three local maxima specifically associated with self-serving attributional decisions.

Regions of significant differences in activation between the specific contrasts including MNI coordinates, cluster size (*k*), *t*-value and *p*-value

TABLE 1

Contrast	MNI						L/R	Region
	X	Y	Z	t-value	k	p-value		
INT > EXT	59	-37	25	5.33	848	<.001	R	TPJ
	-60	-30	22	4.75	312	.047	L	Superior temporal gyrus
EXT > INT	-3	-55	17	5.42	2062	<.001	L	Precuneus
	-42	-68	34	6.50	1887	<.001	L	TPJ
	-22	21	46	4.53	585	.002	L	Superior frontal gyrus
	46	-71	32	5.63	496	.005	R	Angular gyrus

Notes: All *p*-values are cluster level corrected, $p < .05$. Only the highest peak is included in the case of several confluent peaks.

TABLE 2

Regions of significant valence effects are given including MNI coordinates, cluster size (k), t -value and p -value

Contrast	MNI						k	p-value	L/R	Region
	X	Y	Z	t-value	Y	Z				
POS > NEG	-1	-48	22	8.41	8926	<.001	L	Precuneus		
	-40	-66	36	8.46	4043	<.001	L	Angular gyrus		
	-5	44	-10	7.64	3768	<.001	L	Rectal gyrus		
	-23	27	45	6.65	2865	<.001	L	Superior frontal gyrus		
	37	36	6	5.41	1414	<.001	R	Inferior frontal gyrus		
	27	-26	-23	5.66	1172	<.001	R	Fusiform gyrus		
	39	-68	38	4.29	1230	<.001	R	Angular gyrus		
	-30	-33	-22	6.47	1083	<.001	L	Fusiform gyrus		
	-58	-46	-14	4.89	942	<.001	L	Inferior temporal gyrus		
	41	-68	-37	5.60	765	<.001	R	Cerebellum (Crus II)		
NEG > POS	-33	36	-15	6.03	557	.003	L	Inferior frontal gyrus (p. orbitalis)		
	-47	37	10	4.23	407	.015	L	Inferior frontal gyrus (p. triangularis)		
	20	-22	12	3.66	387	.019	R	Thalamus		
	22	25	47	3.89	312	.047	R	Superior frontal gyrus		
	1	46	32	5.08	679	<.001	R	Superior medial frontal gyrus		
	-25	-75	-34	5.47	405	.015	L	Cerebellum (Crus II)		
	19	60	21	4.62	325	.040	R	Superior frontal gyrus		

Notes: All p -values are cluster level corrected, $p < .05$. Only the highest peak is included in the case of several confluent peaks.