In search of the depressive self: extended medial prefrontal network during self-referential processing in major depression

Cédric Lemogne, Guillaume le Bastard, Helen Mayberg, Emmanuelle Volle, Jean-François Alliaire, and Philippe Fossati

Major depression is associated with an excessive self-focus, a tendency to engage oneself in self-referential processing. The medial frontal gyrus (MFG) is central to self-referential processing. This study aimed to explore the neural bases of this excessive self-focus and to disambiguate the role of the MFG in the pathophysiology of major depression. We presented 15 depressed patients and 15 healthy subjects with personality traits during functional magnetic resonance imaging and asked them to judge whether each trait described them (‘self’ condition) or a generally desirable trait (‘general’ condition). Both patients and healthy subjects activated the MFG in ‘self’ vs ‘general’ condition. However, the activation of the dorsal part of the MFG and of the dorsolateral prefrontal cortex (DLPFC) in ‘self’ vs ‘general’ condition was unique to patients. Additionally, patients displayed an increased functional connectivity between the MFG, the dorsal anterior cingulate cortex and the DLPFC. These results provide evidence for an extended medial prefrontal network during self-referential processing in major depression, suggesting the involvement of a greater cognitive control.

Keywords: depression; emotion; mood; prefrontal cortex; self

INTRODUCTION

According to the cognitive theory of depression, depressed patients suffer from two kinds of cognitive biases (Clark and Beck, 1999). First, they allocate more attention to negative stimuli. Second, they display systematic errors in the cognitive appraisal of experience. For instance, major depression is associated with an excessive self-focus, a tendency to engage oneself in self-referential processing (Mor and Winquist, 2002). Self-referential processing concerns the appraisal of stimuli as strongly related to one’s own person (Northoff et al., 2006). The neural bases of the increased attention to negative stimuli are under intense scrutiny, whereas those of the excessive self-focus remain mostly unexplored in depressed patients (Northoff, 2007). In healthy subjects, self-referential processing relies on the cortical midline structures, including the medial frontal gyrus (MFG) (Gusnard et al., 2001; Kelley et al., 2002; Fossati et al., 2003; Northoff et al., 2006). For instance, Fossati et al. (2003) presented healthy subjects with personality traits and asked them to judge whether each trait described them (‘self’ condition) or whether it described a generally desirable trait (‘general’ condition). Activation in the MFG was unique to the ‘self’ condition.

In the present study, we used the same task with functional magnetic resonance imaging (fMRI) to explore the neural bases of the excessive self-focus associated with major depression and to disambiguate the role of the MFG in its pathophysiology. The MFG is central to the neural networks that are impaired in depression (Mayberg, 2003; Drevets et al., 2008), and both increased and decreased activation of the MFG have been reported in depressed patients (Fitzgerald et al., 2008). Additionally, aberrant activation of the MFG was observed in healthy individuals with high scores of neuroticism (Keightley et al., 2003; Haas et al., 2008), a marker of vulnerability for depression. The MFG underlies emotional regulation through cognitive appraisal in both healthy (Ochsner and Gross, 2005) and depressed subjects (Johnstone et al., 2007).

Here, we hypothesized that depressed patients would display an aberrant activation of the MFG during self-referential processing. Because there is evidence that self-focus is not only quantitatively increased in depression, but also qualitatively distinct (Mor and Winquist, 2002;
Watkins and Teasdale, 2004), we additionally hypothesized that depressed patients would activate a different set of brain regions during self-referential processing, compared to healthy controls.

MATERIALS AND METHODS

Subjects
All subjects were native French-speaking, right-handed and gave written informed consent after complete description of the study. They were screened for DSM-IV diagnoses with the Mini International Neuropsychiatric Interview (Sheehan et al., 1998). Severity of depression was assessed using the Montgomery–Asberg Depression Rating Scale (MADRS) and the Beck Depression Inventory (BDI) (Beck et al., 1961; Montgomery and Asberg, 1973). Patients with a major depression were recruited from the Pitié-Salpêtrière Hospital psychiatry department. Exclusion criteria were history of manic episode, psychotic features, medical disorders or medication likely to affect cognition, substance-related disorders or electroconvulsive therapy in the previous 12 months. Healthy subjects with no history of psychiatric disorders were recruited from the community. The Ethics Committee for Biomedical Research of the Pitié-Salpêtrière Hospital approved the study.

Task design
Stimuli were visually presented words: 140 personality traits, either positive (e.g. ‘generous’) or negative (e.g. ‘greedy’), and 50 non-human neutral attributes, either a color (e.g. ‘green’) or not (e.g. ‘adjacent’).

There were three judgment conditions: ‘self’, ‘general’ and a control condition. In both ‘self’ and ‘general’ conditions, the subjects were presenting with personality traits. In the ‘self’ condition, subjects judged whether the trait described them. In the ‘general’ condition, subjects judged whether the trait was socially desirable. For each subject, personality traits were randomly allocated to either ‘self’ or ‘general’ condition. In the control condition, subjects were presenting with neutral non-human attributes. They judged whether the attribute described a color. In all conditions, subjects gave a ‘yes’ or ‘no’ response for each word by pushing a button with the right or the left thumb, respectively.

The task encompassed one practice run, performed outside the scanner, and four scanning runs. Each run contained three blocks. Each block was associated with only one condition. The order of the conditions was counterbalanced across the runs, in order to avoid presenting the same condition in two consecutive blocks. Before each block, an instruction cue was displayed for 9 s (e.g. ‘self’), followed by a central fixation crosshair for 3 s. Each block contained 10 trials: five negative and five positive words for ‘self’ and ‘general’ conditions or five color and five non-color words for the control condition. Each trial consisted of a word displayed for 1 s, followed by a fixation crosshair for 9 s. Words were randomly intermixed such that words of each type (i.e. positive and negative or color and non-color) followed each other equally often.

Self-referential processing is associated with a mnemonic superiority, relative to other encoding strategies (Symons and Johnson, 1997). To ensure that the subjects adequately performed the task, we looked for this behavioural effect. The scanning runs were immediately followed by an unexpected recognition task, performed outside the scanner, in which the subjects had to discriminate between studied and unstudied personality traits.

fMRI scanning
Stimuli were generated by the Paradigm software (http://www.eye-brain.com) and projected on the centre of a screen mounted outside the scanner. Subjects viewed the screen through mirror glasses.

Four functional runs of 170 contiguous volumes (14 axial slices of 5 mm thickness obtained every 2 s) were acquired on a 1.5-T whole-body scanner (SIGNA, GE, Milwaukee, WS, USA), using T2-weighted gradient echo, echo-planar imaging (EPI) sequence, sensitive to blood oxygen level-dependent contrast (repetition time: 2 s, echo time: 40 ms, flip angle: 90°, matrix: 64 × 64, field of view: 220 × 220 mm). Each run lasted 340 s. The first three volumes of each run were discarded to reach signal equilibrium. High-resolution three-dimensional T1-weighted images (3D fast gradient echo inversion recovery sequence, inversion time: 400 ms, repetition time: 1600 ms, echo time: 5 ms, matrix: 256 × 256 × 128, field of view: 220 × 220 mm, slice thickness: 1.5 mm) were acquired for anatomical localization.

fMRI data analysis
We used SPM5 software (http://www.fil.ion.ucl.ac.uk/spm/software/spm5).

Preprocessing
EPI volumes were corrected for slice timing, realigned to the first image, co-registered with the high-resolution T1-weighted image and normalized into a standard stereotactic space. The normalization used the Montreal Neurological Institute (MNI) template and the rigid transformations computed during the segmentation of the high-resolution T1-weighted image. Finally, the normalized EPI volumes were smoothed using an isotropic Gaussian kernel filter of 8 mm full-width half-maximum.

Activation analysis
For each subject, we computed an individual statistical parametric map using the general linear model and an event-related approach (Friston et al., 1998). Each trial onset was convolved with the canonical hemodynamic response function (HRF) and its temporal derivative (TD) to create regressors of interest. Because patients and controls slightly differed regarding reaction times (see ‘Results’ section), we used the TD to deal with potential shifts in
the timing of the HRF (Friston et al., 1998). A high-pass filter was applied and the motion realignment parameters were included as regressors of non-interest (Friston et al., 1996). The following individual contrast images were obtained for the HRF estimates: self/positive vs control, self/negative vs control, general/positive vs control and general/negative vs control.

We performed a random effects second-level ANOVA with one group factor and two within-groups factors (i.e. condition and valence) using the first-level individual contrast images. First, we looked for a group main effect (patients vs controls), in both conditions (self and general). Second, we performed two conjunction analyses with a conjunction null hypothesis to look for regions displaying a condition main effect (‘self’ vs ‘general’) or a valence main effect (positive vs negative) in both groups (patients and controls) (Nichols et al., 2005). We used conjunction analysis to identify only regions that were significantly activated by both patients and controls. Third, we looked for significant interactions between factors (i.e. group, condition and valence). In each case, we first computed a F-contrast. Post hoc t-contrasts followed whenever a significant effect was found.

To ensure that these functional results were not reflecting a structural impairment, we subsequently used optimized Voxel-based morphometry (Ashburner and Friston, 2000). Briefly, native T1-weighted volumes were normalized and segmented in grey and white matter partitions. Prior probability maps that were relevant to tissue segmentation were warped to the individual brains, making the creation of a customized template unnecessary. We used a smoothing kernel with a full width at half maximum of 12 mm. Groups were compared using an ANOVA looking for clusters of five contiguous voxels whose global maxima meet an uncorrected threshold of \( P < 0.001 \) for regions identified during the fMRI activation analysis and a False Discovery Rate (FDR)-corrected threshold of \( P < 0.05 \), for any other region.

**Functional connectivity analysis**

Activation analysis identified a single region that was activated in ‘self’ vs ‘general’ condition in both groups (see ‘Results’ section). To ensure that this region, located in the MFG, was activated to the same extent by patients and controls, we used the group \( \times \) condition F-contrast as an exclusive mask with a liberal threshold of \( P < 0.05 \), uncorrected.

To further explore the functional connectivity of this MFG region, we extracted the individual time-series data within a sphere of 10 mm radius around the main peak (MNI coordinates: 6, 48, 18) for each run for every subject. To avoid extracting data from voxels not uniquely activated by the ‘self’ condition, any voxel not activated in ‘self’ vs ‘general’ at a liberal threshold of \( P < 0.05 \) was removed from the data before extraction. Consequently, three controls and three patients were excluded from the functional connectivity analysis as they did not activate any voxel in ‘self’ vs ‘general’ in this region at a liberal threshold of \( P < 0.05 \). We then computed an individual statistical parametric map for each subject, with the time-series as a regressor of interest and the motion realignment parameters as regressors of non-interest. A high-pass filter was applied. We obtained one first-level contrast image for each subject. We used these images in a second-level t-test to examine the extent to which the functional connectivity of this MFG region was significantly different across groups.

**RESULTS**

**Subjects**

Fifteen depressed inpatients (1 man, 14 women) and 15 healthy subjects (4 men, 11 women) were recruited (Table 1). All patients were taking antidepressant and were tested within the first week of receiving their treatment. Seven patients were taking a selective serotonin reuptake inhibitor, five a serotonin-norepinephrine reuptake inhibitor and three a tricyclic antidepressant. Sedative drugs were not allowed on the experiment day.

**Behavioural results**

We used SPSS software to perform ANOVAs with one group factor (i.e. patients vs controls) and two within-groups factors (i.e. ‘self’ vs ‘general’ and positive vs negative). Due to technical problems, judgment and recognition data were lost for one control and recognition data were lost for one patient.

Regarding responses during ‘self’ and ‘general’ conditions, there was a valence main effect \( (P < 0.001), \) with more ‘yes’ for positive words overall, a group \( \times \) valence interaction \( (P < 0.001), \) with less ‘yes’ for positive words and more ‘yes’ for negative words, in patients vs controls, and a condition \( \times \) valence interaction \( (P < 0.001), \) with more ‘yes’ for positive words and less ‘yes’ for negative words, in ‘self’ vs ‘general’ condition.

Regarding the reaction times during ‘self’ and ‘general’ conditions, there was a condition main effect \( (P < 0.001), \)
with slower responses in ‘self’ condition and a condition × valence interaction ($P=0.016$), with slower responses for positive words than negative words in ‘self’ condition. There was also a trend for a main effect of group ($P=0.065$), patients tending to be slower than controls overall.

Regarding the recognition task, a condition main effect ($P<0.001$) was explained by a better recognition of the words processed in ‘self’ vs ‘general’ condition, in both patients ($P=0.008$) and controls ($P=0.001$). There was also a valence main effect ($P<0.020$) that was explained by a better recognition of the positive words overall.

There was neither a group × condition interaction nor a group × condition × valence interaction for any of the behavioural variables.

**fMRI results**

Regarding the MFG, which was our *a priori* region, clusters of five contiguous voxels whose global maxima meet a threshold of $P<0.001$, uncorrected, are reported. Regarding any other region, clusters of five contiguous voxels whose global maxima meet a FDR-corrected threshold of $P<0.05$ are reported. We used the Wake Forest University School of Medicine PickAtlas software toolbox to generate a MFG mask based on the Talairach Daemon database (Lancaster et al., 2000; Maldjian et al., 2003).

**Activation analysis**

Regarding the main effects, there was a condition main effect and a valence main effect but no significant group main effect.

Regarding the condition main effect, the conjunction analysis identified a single cluster in the MFG (594 mm$^3$; MNI coordinates: 6, 48, 18; $F_{(2,112)}=15.32$, $P<0.001$, uncorrected). This condition main effect was explained by a MFG activation (1296 mm$^3$; MNI coordinates: 6, 48, 18; $t_{(2,112)}=3.91$, $P<0.001$, uncorrected) in ‘self’ vs ‘general’ condition in both patients and controls (Figure 1). This cluster survived a contrast masking procedure using the group × condition F-contrast as an exclusive mask with a threshold of $P<0.05$, uncorrected. We subsequently used this cluster (MNI coordinates: 6, 48, 18; Brodmann Area 9) to look for differences in MFG functional connectivity between patients and controls.

Regarding the valence main effect, the conjunction analysis identified two cluster of activation that were restricted to right (3051 mm$^3$; MNI coordinates: 45, −18, 42; $F_{(2,112)}=38.94$, $P<0.05$, FDR-corrected) and left (1269 mm$^3$; MNI coordinates: −36, −24, 48; $F_{(2,112)}=29.19$, $P<0.05$, FDR-corrected) sensorimotor regions, for negative vs positive and positive vs negative contrasts, respectively. This was expected due to the requirement to use the right and the left thumb to answer ‘yes’ and ‘no’, respectively.

Regarding the interactions, there was a significant group × condition interaction within a more dorsal region of the MFG (85 mm$^3$; MNI coordinates: 9, 42, 33; $F_{(2,112)}=13.35$, $P<0.001$, uncorrected), henceforth referred to as dorsomedial frontal gyrus, but neither a group × valence, nor a condition × valence, nor a group × condition × valence interaction. A *post hoc* $t$-contrast showed that the group × condition interaction was explained by an activation of the dorsomedial frontal gyrus (216 mm$^3$; MNI coordinates: 9, 42, 33; $t_{(1,112)}=3.65$, $P<0.001$, uncorrected) in ‘self’ vs ‘general’ condition, which was unique to patients (Figure 2A). Additionally, there was a similar group × condition interaction in the left inferior frontal gyrus (1080 mm$^3$; MNI coordinates: −42, 9, 36; $t_{(1,112)}=4.70$, $P<0.05$, FDR-corrected) (Figure 2A). The activation–deactivation pattern underlying this interaction was similar across the two regions (Figure 2B).

Removing the patient with comorbid panic attacks or introducing gender, age, educational level as covariates yielded similar results. The activation of these two regions in ‘self’ vs ‘general’ condition was neither explained by a difference of grey matter volume between patients and controls nor correlated with the severity of depression or the number of days under antidepressant in patients (even at a liberal threshold of $P<0.05$, uncorrected).

**Functional connectivity**

In controls and patients, the activity of the posterior cingulate cortex was positively correlated with the MFG cluster that was equally activated in ‘self’ vs ‘general’ condition in both groups (MNI coordinates: 6, 48, 18) (Table 2, Figure 3). No region was more positively correlated with this cluster in controls than patients, whereas some regions were more positively correlated with this cluster in patients than controls, including the dorsomedial frontal gyrus, the dorsal anterior cingulate cortex (ACC) and the right and left
DLPFC (Table 2, Figure 3). Contrast masking showed that these regions were positively correlated in patients and not negatively in controls, thus accounting for the differences between groups.

**DISCUSSION**

This study aimed to explore the neural bases of the excessive self-focus in major depression. Because the main effects were not relevant to our hypotheses, we will focus on the group \times condition interaction. Although both patients and healthy subjects activated the MFG (MNI coordinates: 6, 48, 18) during self-referential processing, the activation of a more dorsal part of the MFG (MNI coordinates: 9, 42, 33), henceforth referred to as dorsomedial frontal gyrus, was unique to patients, as was the activation of the dorsolateral prefrontal cortex (DLPFC). The more ventral part of the MFG, which was equally activated by self-referential processing in both groups, was part of a larger functional network that was common to patients and controls and included the posterior cingulate cortex. These cortical midline structures are known to be active during both resting state and self-referential processing (Gusnard et al., 2001; Lou et al., 2004; Northoff et al., 2006). There is a reciprocal modulation between these regions and the DLPFC in healthy subjects (Greicius et al., 2004), whereas depressed patients actually displayed an increased functional connectivity between the MFG, the DLPFC and a region overlapping the dorsomedial frontal gyrus and the dorsal ACC. These results suggest that there is an abnormal reciprocal cortico-cortical modulation in major depression (Mayberg, 2003; Seminowicz et al., 2004).

The activation of the left DLPFC during self-referential processing, but only in patients, suggests that they require cognitive control during self-referential processing (Lieberman, 2007). Although neuroimaging studies in depression have generally found decreased activity of the
DLPFC at rest, activation studies have shown either increased or decreased activation of the DLPFC (Davidson et al., 2002; Fitzgerald et al., 2006, 2008). For instance, depressed patients need greater left DLPFC and dorsal ACC activation to maintain a level of performance similar to controls during a working memory task (Harvey et al., 2005; Matsuo et al., 2007).

The construct of self-focus was introduced by Duval and Wicklund (1972) to account for the links between attention to the self and affect. Building on this early model, Carver and Scheier (1998) proposed that self-focus leads one to consider the discrepancy between his or her current state and a salient standard. Positive affect is experienced if the current state surpasses the standard, whereas negative affect is experienced if the current standing falls short of the standard. Experiencing negative affect generates attempts either to decrease this discrepancy or to avoid self-focus. Depression may then occur when one is unable either to fix the discrepancy or to avoid self-focus (Mor and Winquist, 2002).

Whereas the DLPFC is thought to implement cognitive control (Koechlin et al., 2003), the dorsal ACC and the dorsomedial frontal gyrus are thought to implement conflict monitoring (Kerns et al., 2004; Etkin et al., 2006). Self-focus may represent a particular instance of conflict monitoring (Duval and Wicklund, 1972; Carver and Scheier, 1998; Mor and Winquist, 2002), and the absence of any valence effect suggests that this conflict is between the current self and the stimulus itself. During the ‘self’ condition, positive and negative personality trait words may equally signal the absence of a personally meaningful state. For instance, the word ‘generous’ is just as likely to elicit self-focus (‘Am I not greedy?’) as the word ‘greedy’ itself. In depressed patients, the self-referential processing performed by the MFG may have required conflict monitoring by the dorsomedial frontal gyrus and the dorsal ACC, whatever the intrinsic valence of the presented word. The extended MFG activation and its increased connectivity with the dorsomedial frontal gyrus may then represent the neural correlates of a secondary compensatory mechanism rather than those of the excessive self-focus per se. This excessive self-focus may have required further cognitive control by the DLPFC for at least two reasons.

When one’s current self falls short of one’s own standard, subsequent negative affect may generates attempts either to decrease the discrepancy or to avoid self-focus (Duval and Wicklund, 1972; Carver and Scheier, 1998; Mor and Winquist, 2002). Thus, the dorsomedial frontal gyrus may have issued a call to the DLPFC to reduce the adverse emotional consequences of the discrepancy by either cognitive reappraisal or self-focus inhibition (Ochsner and Gross, 2005). These two mechanisms may account for the greater functional connectivity between the MFG and the DLPFC in depressed patients. Indeed, a recent fMRI study suggests that depressed patient may activate a larger prefrontal network than healthy subjects during emotional regulation (Johnstone et al., 2007). Further studies should investigate
more specifically the role of the DLPFC in depressive self-focus.

A serious limitation of our study is that patients were receiving medications whereas healthy subjects were not. Therefore, our findings need to be replicated in unmedicated patients, as medications may partially account for them. However, sedative drugs were not allowed on the experiment day, and the patients were tested within the first week of receiving their antidepressant treatment. Furthermore, if antidepressants were accounting for the present results, the number of days under antidepressant prior to the study should have been correlated, even slightly, with the dorsomedial frontal gyrus or the DLPFC activity during self-referential processing in depressed patients. We did not find such a correlation, even with a very liberal statistical threshold. It is therefore unlikely that the group × condition interaction, which reached a 50-fold more stringent statistical threshold, was only explained by antidepressants.

Some other methodological limitations should be considered. First, 14 slices of 5 mm thickness were insufficient to cover the whole brain for some subjects. Consequently, some regions were not included in second-level analyses, including the amygdala that was previously found to be activated during self-referential processing in depression (Siegle et al., 2002). Second, because the EPI volumes featured some signal drop in the most ventral part of the MFG, we did not fully explore this region. However, this risk of a type II error does not challenge the significant interaction found in the dorsomedial frontal gyrus. Regarding the risk of a type I error due to multiple comparisons, our findings in the MFG rely on hypotheses-driven analyses performed within an a priori region and the left DLPFC global maxima survived a FDR correction. Finally, gender was not perfectly matched.

Some theoretical limitations should also be considered. We operationalized self-focus as an increased tendency to engage oneself in self-referential processing. However, self-focus is unlikely to be a single psychological construct (Mor and Winquist, 2002). For instance, it remains unclear whether self-focus in major depression is a controlled or an automatic process or both. Controlled processes are associated with awareness, intention, effort and the capacity for interruption. Automatic processes lack one or more of these features (Lieberman, 2007). Here, the present task was weighted toward controlled processes and yielded results relevant for the controlled aspects of the excessive self-focus in depression. Future studies should also include a task weighted toward automatic processes to explore the automatic aspects of this excessive self-focus, as they may rely on different brain regions (e.g. ventral MFG, which was not fully explored in the present study). Further studies should also better characterize categorical (e.g. rumination) as well as dimensional (e.g. duration, intensity) aspects of self-focus. Finally, further studies should explore the neural bases of the depressive self in remitted patients to address trait vs state issues.

In summary, we used fMRI and two emotional conditions, either self-referential or not, to explore the neural bases of the excessive self-focus in major depression. Our results provided evidence for an extended cortico-cortical network during self-referential processing in depressed patients, suggesting the involvement of a greater cognitive control. To our knowledge, this is the first study that disambiguates the role of the MFG in the pathophysiology of depression.

REFERENCES


Johnstone, T., van Reekum, C.M., Urry, H.L., Kalin, N.H., Davidson, R.J. (2007). Failure to regulate: counterproductive recruitment of top-down...
prefrontal-subcortical circuitry in major depression. *Journal of Neuroscience*, 27, 8877–84.


