

A brain model of disturbed self-appraisal in depression

Authors:

Christopher G. Davey, M.D., Ph.D.^{1,2,3}

Michael Breakspear, M.D., Ph.D.^{4,5}

Jesus Pujol, M.D., Ph.D.⁶

Ben J. Harrison, Ph.D.^{1,2,3}

Author affiliations:

¹Orygen, The National Centre of Excellence in Youth Mental Health, Melbourne, Australia

²Centre for Youth Mental Health, The University of Melbourne, Melbourne, Australia

³Melbourne Neuropsychiatry Centre, Department of Psychiatry, The University of Melbourne, Melbourne, Australia

⁴QIMR Berghofer Medical Research Institute, Brisbane, Australia

⁵Metro North Mental Health Service, Brisbane, Australia

⁶MRI Research Unit, Department of Radiology, Hospital del Mar, CIBERSAM G21, Barcelona, Spain

Corresponding author:

Christopher Davey

Orygen, The National Centre of Excellence in Youth Mental Health

35 Poplar Rd

Parkville, Victoria 3052, Australia

e-mail: c.davey@unimelb.edu.au

Keywords: depression, self, youth, fMRI, neural network modeling, dynamic causal modeling

Word count, abstract: 240

Word count, manuscript: 3,910

Abstract

Objective: A disturbed sense of self is a core feature of depression. The medial prefrontal cortex, which has a central role in self-appraisal processes, is often implicated in the illness, although it remains unclear how functional alterations of the region contribute to the observed disturbances. The aim of the study was to clarify the role of the medial prefrontal cortex in self-appraisal processes in depression.

Method: We applied a recently developed dynamic network model of self-directed cognition to functional magnetic resonance imaging data from 71 youth with moderate-to-severe major depressive disorder, none of whom were being treated with medication, and 88 healthy control participants. Bayesian model averaging was used to determine parameter estimates for the dynamic causal models, which were compared between groups.

Results: While self-directed cognitive processes in depression were shown to rely on the same dynamic network as in healthy controls, the medial prefrontal cortex had a ‘hyper-regulatory’ effect on posterior cingulate cortex in the depressed participants, with self-appraisal causing significantly more negative modulation of connectivity between medial prefrontal cortex and posterior cingulate cortex than in the controls (odds ratio=0.54, 95% CI 0.38 to 0.77; $p=0.0008$). This parameter was inversely related with a depression factor related to poor concentration and inner tension ($r=-0.32$; 95% CI -0.57 to -0.09; $p=0.01$).

Conclusions: The exaggerated influence of the medial prefrontal cortex on posterior cingulate cortex in depression is a neural correlate of the disturbed self-appraisal that is characteristic of the illness.

Introduction

People with depression view themselves through a critical lens. They see themselves, says Freud, as ‘worthless, incapable of any achievement and morally despicable’ (1); and to Beck and colleagues, as ‘defective, inadequate, diseased or deprived’ (2). Such disturbances in self-appraisal are fundamental features of depression, and are specifically addressed by effective psychotherapies for the illness (2, 3).

Neurobiological research into self-appraisal processes has consistently implicated activity of midline cortical regions: in particular, medial prefrontal cortex and posterior cingulate cortex (4, 5). These regions are core components of the so-called default mode network, which shows greater activity and connectivity when a person is at rest than when engaged with external tasks. The substantial overlap between self-related and resting-state brain activity can be explained – at least in part – by the fact that people spend more time introspecting and day-dreaming when they are not engaged with the external environment (6, 7).

We have recently characterized a dynamic network model of self-directed cognition, composed of medial prefrontal cortex, posterior cingulate cortex, and inferior parietal lobule (8). Self-related processes – at rest and when prompted to explicitly think about one’s attributes – are supported by activity in posterior regions and regulated by feedback from the medial prefrontal cortex; with the regulatory influence of medial prefrontal cortex further enhanced by the more effortful self-directed cognition required by explicit self-appraisal.

Alterations in functioning of the medial prefrontal cortex have been frequently implicated in major depressive disorder. Early nuclear imaging studies of depression reported increased activity in medial prefrontal cortex during rest (9–11). More recent functional magnetic resonance imaging studies have reported increased resting-state connectivity between the medial prefrontal cortex and other default mode regions (12–14). While these resting-state abnormalities suggest a possible brain substrate for the alterations in self-appraisal observed in depression, they do not demonstrate it directly. Studies that *have* examined the response to self-referential stimuli in depression have been inconsistent, reporting that activity in the medial prefrontal cortex is both increased (15, 16) and

decreased (17, 18), leaving the mechanisms linking alterations in medial prefrontal cortical functioning to disturbed self-appraisal in depression uncertain.

There is a clear need for more sophisticated brain models, including the use of computational models that allow for a more concise mapping between brain activity and psychological states (19). Dynamic causal modeling provides a principled means of characterizing the role of the medial prefrontal cortex in depression by measuring its dynamic interactions – or *effective connectivity* – with other regions, inferring the strength and direction of regional interactions (20).

The aim of this study was to examine disturbed self-appraisal processes in depression, with an emphasis on the medial prefrontal cortex. Our study included adolescents and young adults, allowing us to focus on a developmental period when self-related concepts have become especially salient (21, 22), and when vulnerability to depression has risen substantially from childhood (23). The increased salience of the self, and greater vulnerability to depression, are likely related to neurodevelopmental processes that continue throughout adolescence and early adulthood (24), not ending until a young person reaches their mid-twenties (25). We hypothesized that disturbed self-appraisal – we use the term ‘disturbed’ in its everyday sense, as referring to a disruption of normal functioning – would be reflected in depressed youth appraising themselves in more negative terms; and that this would be associated with altered interactions between medial prefrontal cortex and posterior cingulate cortex. We additionally sought to investigate how variability in the clinical presentation of depression, including anxiety co-morbidity, co-varied with alterations in network parameters related to self-appraisal.

Methods

Participants

Eighty-six unmedicated depressed participants were recruited from mental health clinics in the western and northern suburbs of Melbourne, Australia. Patients were between 15 and 25 years of age, and had major depressive disorder determined by Structured Clinical Interview for DSM-IV (SCID) (26). The depression was of at least moderate severity, defined as a score on the Montgomery-Åsberg Depression Rating Scale (MADRS) of 20 or greater. These young patients were typically in the early

phases of illness, when the cumulative effects of treatments, recurrence, and chronicity have yet to accrue. Exclusion criteria were: diagnoses of psychotic or bipolar disorders, full-scale intelligence quotient – estimated by the Weschler Test of Adult Reading (27) – of less than 85, and current treatment with medication. Most of the depressed participants (66%) had a co-morbid anxiety disorder, reflecting the relatively severe clinical presentation of these young people, all of whom were attending mental health services. The most common co-morbid anxiety disorder was social anxiety disorder, which was present in 35% of patients (Table 1).

This clinical cohort was compared to control participants matched for age and sex. Ninety-five healthy control participants were recruited by advertisements placed in online media. Exclusion criteria were current or past diagnoses of mental illness according to structured assessment with the SCID, and estimated IQ of less than 85. These participants were included in our previous report of a brain model of self-directed cognition (8). All participants (and their parents if they were under 18 years of age) provided their informed consent to participate in the study, which was approved by the Melbourne Health Human Research and Ethics Committee.

Paradigm design

Participants completed a functional magnetic resonance imaging task composed of three experimental conditions: self-appraisal, external attention, and rest-fixation. In the self-appraisal condition, participants were presented with a personality adjective and asked whether or not the word described them. Words were drawn from a frequently used list of personality adjectives (28), and included words such as ‘skeptical’, ‘perfectionistic’, and ‘lucky’. We selected words that were distributed around the median rating for ‘likeableness’ reported in the original dataset, from the subset of words rated as most ‘meaningful’ (see Davey and colleagues (8)). Participants viewed eight blocks of six words, presented for 5 s each, and responded to the question, ‘Does this word describe you?’, by pressing the left or right button on a fiber-optic response pad. In the external attention condition, participants viewed eight blocks of six words, also presented for 5 s each, and responded to the question, ‘Does this word have four or more vowels?’.

The two lists of 48 words that formed the self-appraisal and external attention conditions were matched on likeableness ratings and number of vowels and were counterbalanced across participants.

Each 32 s block (2 s instruction followed by six words presented for 5 s each) was interspersed with a 10 s rest-fixation block in which participants were asked to fixate on a centrally presented cross-hair. Behavioral data (accuracy and response-times) were analyzed with Stata 14.1 (StataCorp, College Station, USA).

Image acquisition

Functional neuroimaging data were acquired using a 3T General Electric Signa Excite system equipped with an eight-channel phased-array head coil in combination with ASSET parallel imaging. The functional sequence consisted of a single shot gradient-recalled EPI sequence in the steady state (repetition time, 2 s; echo time, 35 ms; and pulse angle, 90°) in a 23-cm field-of-view, with a 64 x 64-pixel matrix and a slice thickness of 3.5 mm (no gap). Thirty-six interleaved slices were acquired parallel to the anterior-posterior commissure line with a 20° anterior tilt to better cover ventral prefrontal cortical brain regions. The total sequence time was 11 min 22 sec, corresponding to 341 whole brain echo-planar imaging volumes. A T1-weighted high-resolution anatomical image was acquired for each participant to assist with functional time-series co-registration (140 contiguous slices; repetition time, 7.9 s; echo time, 3 s; flip angle, 13°; in a 25.6 cm field-of-view, with a 256 x 256 pixel matrix and a slice thickness of 1 mm).

Image preprocessing

Imaging data were processed with Statistical Parametric Mapping software (SPM12; Wellcome Trust Centre for Neuroimaging, UK) using MATLAB version 8.2 (The MathWorks Inc, Natick, USA). Motion correction was performed by aligning each participant's time series to the first image using least-squares minimization and a six-parameter rigid-body spatial transformation. Participants' data were excluded if movement in the translational or rotational planes exceeded 2 mm or 2°, respectively. These realigned functional images were then co-registered to each participants' respective T1 anatomical scans, which were segmented and spatially normalized to the International Consortium for Brain Mapping template using the unified segmentation approach. The functional images were interpolated to 2 mm isotropic resolution, and were smoothed with a 5 mm full-width-at-half-maximum gaussian filter.

Of the 181 participants who completed the scanning session, 22 were subsequently excluded: 10 due to excessive head movement during scanning (7 depressed and 3 control participants), and 12 due to poor performance on the external attention task – defined as less than 85% accuracy (8 depressed and 4 control participants). The final composition of the imaging sample thus included 71 depressed and 88 control participants (Table 1). The groups showed a modest but significant difference in full-scale intelligence quotient: IQ was added as a covariate to all group comparisons.

Defining the nodes

The nodes for our dynamic causal modeling analyses were identified using a general linear model approach in SPM12. We defined the nodes as regions that were (i) more active during self-appraisal and rest-fixation than during external attention, and (ii) additionally activated by self-appraisal over-and-above rest-fixation. We used the external attention condition as the implicit baseline: it was an appropriate baseline condition for our objectives, being matched with self-appraisal on stimulus features, but mandating specific attentional demands that suppressed the likelihood of task-independent activity. For each participant, primary regressors for the self-appraisal and rest-fixation conditions were defined by specifying the onset and duration of each block. The regressors were convolved with a canonical hemodynamic response function, and a high-pass filter set at 128 s was used to remove low-frequency drifts. Parameter estimates were calculated at each voxel using the general linear model with local auto-correlation correction.

As in our previous study (8), the second-level analysis identified regions that showed significant conjunction between activation to rest-fixation, and activation to self-appraisal minus rest-fixation. In other words, the nodes for our dynamic causal modeling analysis were composed of brain regions that were broadly activated by self-appraisal (across rest-fixation and self-appraisal conditions), but further activated by explicit self-appraisal. Whole-brain, voxel-level, family-wise-error corrected statistical thresholds were applied to define these regions ($p_{\text{FWE}} < 0.05$).

As in our previous work, these contrasts identified regions in medial prefrontal cortex, posterior cingulate cortex, and left inferior parietal lobule. Within these broader regions, we then defined subject-specific volumes-of-interest following the principles established by Stephan and colleagues

(29), which outline the requirement to define subject-specific time-series whose precise anatomical location varies. The group coordinates for the volumes-of-interest were the maxima for the main effect of self-appraisal and rest-fixation for clusters in the medial prefrontal cortex, posterior cingulate cortex, and inferior parietal lobule. For each participant we extracted regional time-series, summarized as the first eigenvariate of voxels that showed significant activation to self-appraisal and rest-fixation ($p < 0.01$) within 5 mm of the subject-specific maxima, which were themselves no more than 10 mm from the group maxima. The dynamic causal modeling analysis only included participants who had volumes-of-interest identified in each of the three regions according to these criteria. We extracted time-series for all volumes-of-interest from 68 of the 71 depressed participants and from 82 of the 88 control participants.

Dynamic causal modeling

The aim of dynamic causal modeling is to infer the causal architecture of a network of neural regions. Using a Bayesian approach, a model is selected from a set of predefined models that is the most likely to generate the observed imaging data, while also penalizing for model complexity. The dynamic causal models delineate how dynamics in one brain region influence dynamics in others, incorporating both the core set of inter-regional connections and the modulation of those influences by the experimental manipulations (30). Effective connectivity between brain regions – the directed influence one region has on another – can be either positive or negative, such that an increase in activity in one region causes an increase or decrease, respectively, in the rate of change in another (for more detailed descriptions, see articles by Stephan and colleagues (30) and Kahan & Foltynie (31)).

For this study a three-region deterministic dynamic causal model was specified for each participant with bidirectional effective connections between medial prefrontal cortex, posterior cingulate cortex, and inferior parietal lobule. We defined a space of 32 candidate models where ‘broad self’ (self-appraisal and rest-fixation) drove either medial prefrontal cortex (models 1–16) or posterior cingulate cortex (models 17–32). Self-appraisal modulated efferent connections from medial prefrontal cortex to posterior cingulate cortex, inferior parietal lobule, both, or neither; and/or connections from posterior cingulate cortex to medial prefrontal cortex, inferior parietal lobule, both, or neither (i.e., a $2 \times 4 \times 4 = 32$ candidate model-space; see Supplementary Fig. 1).

Within the depressed and control groups, the models were compared using random-effects Bayesian model selection in DCM12. We computed posterior probabilities and protected exceedance probabilities within each group (32, 33). The protected exceedance probability represents the probability that a given model is more frequent than the others (above and beyond chance) and was our primary measure for model selection (32). The strength of effective connectivity and modulatory effects were summarized using random-effects Bayesian model averaging. Bayesian model averaging is the recommended approach for comparing parameter estimates from dynamic causal models between groups, and involves computing the average model parameters, weighted by the posterior probability of each model (30). The model space was partitioned according to whether broad self drove medial prefrontal cortex (models 1–16) or posterior cingulate cortex (models 17–32), and average connectivity estimates (weighted by their posterior model probabilities) were obtained from models in the winning family for each group (34). Between-group differences for the model parameters were examined using logistic regression in Stata, adjusting for estimated IQ, reaction times, age, and sex. Analyses of between-group differences for the model parameters were adjusted for multiple comparisons using the false discovery rate (35) to determine significance ($q < 0.05$).

We further analyzed associations between the relevant connectivity parameter and clinical and behavioral characteristics of the depressed participants. This included depression severity, as assessed by the MADRS. We performed a factor analysis of the MADRS to determine associations with symptom dimensions, using similar methods to previous analyses of the instrument (36, 37). This involved calculating the principal-component factors, retaining those with an eigenvalue greater than 1, and subjecting them to oblique rotation. We then calculated partial correlations between the factors and our model parameter of interest, adjusted for total MADRS score, reaction times, age, and sex. Analyses of associations with clinical and behavioral variables were adjusted for multiple comparisons using the false discovery rate (35) to determine significance ($q < 0.05$). We also examined the effects of anxiety co-morbidity in the depressed participants, focusing especially on social anxiety disorder, which was the most prevalent disorder.

Results

Behavioral results

The overall task performance in the depressed and control participants did not significantly differ (Table 1). There was no difference in accuracy in the external attention condition ($p=0.30$), and both groups demonstrated similar reaction times for the self-appraisal and external attention tasks ($p>0.15$). The depressed participants more often answered ‘yes’ to the question ‘Does this word describe you?’ ($p=0.0001$); and the mean likeableness ratings for words that were affirmed as describing the self were significantly lower for the depressed compared with the control participants ($p<0.0001$).

Identifying a common model

The conjunction analyses of the rest-fixation (versus external attention baseline) and the self-appraisal minus rest-fixation contrasts identified, independently, three regions in each group – in medial prefrontal cortex, posterior cingulate cortex, and left inferior parietal lobule – showing a high degree of anatomical consistency (Fig. 1, Supplementary Table 1). There was evidence of increased posterior cingulate cortex activation to the self-appraisal condition in the depressed compared to the control participants ($p=0.01$; Fig. 1).

The three regions identified for each group were included in a dynamic network, and dynamic causal modeling was used to determine the network’s most likely structure, given the data. Importantly, we validated our previously described model (8), showing that for both the control and depressed groups the same model outperformed the other models. This common winning model was composed of a driving influence of broad self (self-appraisal and rest-fixation) on posterior cingulate cortex, with self-appraisal modulating both of the efferent connections from each of posterior cingulate cortex and medial prefrontal cortex (Fig. 2). In the control participants, the winning model had a protected exceedance probability of 0.90; and in the depressed participants, had a protected exceedance probability of 0.92 (Supplementary Fig. 2).

Estimating and comparing the model parameters

Parameter estimates were calculated using Bayesian model averaging over a model-space partitioned according to whether broad self drove posterior cingulate cortex or medial prefrontal cortex. We

found clear evidence in both groups to suggest that broad self drove posterior cingulate cortex activity (in control participants, exceedance probability > 0.99 ; in depressed participants, exceedance probability > 0.99). In both participant groups, Bayesian model averaging over the 16 candidate-models in the winning family demonstrated significant positive endogenous influences from posterior cingulate cortex to medial prefrontal cortex – directly and indirectly via the inferior parietal lobule – and significant negative endogenous influences from medial prefrontal cortex to posterior cingulate cortex. The modulatory effects of self-appraisal were such that posterior cingulate cortex had a greater positive influence on medial prefrontal cortex and inferior parietal lobule, and medial prefrontal cortex had a greater negative influence on posterior cingulate cortex. For the depressed, but not control, participants, medial prefrontal cortex also had a significant negative influence on inferior parietal lobule (Fig. 2, Table 2).

Between group-comparison showed that the modulatory influence of self-appraisal on effective connectivity between the medial prefrontal cortex and posterior cingulate cortex (henceforth referred to as ‘MPFC–PCC connectivity’) was significantly more negative for the depressed compared to the control participants (odds ratio=0.54, 95% CI 0.38 to 0.77; $p=0.0008$; Table 2). The influence of sex and age on the connectivity parameters are presented in Supplementary Tables 2 and 3.

Association with clinical and behavioral variables

We examined associations between the strength of modulation of connectivity between medial prefrontal cortex and posterior cingulate cortex and clinical and behavioral characteristics of the depressed participants. There was no association between MPFC–PCC connectivity and variables related to self-appraisal (proportion of personality adjectives affirmed or their valence; $p>0.25$), nor was there significant association between connectivity and the total MADRS score ($p>0.75$). Our factor analysis of the MADRS revealed the presence of four underlying factors, similar to the analysis by Craighead and Evans (37); see Supplementary Table 4. As in their analysis, we observed a factor that consisted of the MADRS items *concentration difficulties* (‘difficulties in collecting one’s thoughts mounting to incapacitating lack of concentration’) and *inner tension* (‘feelings of ill-defined discomfort, edginess, inner turmoil, mental tension’) (38). The strength of modulation of MPFC–PCC connectivity was significantly negatively correlated with this factor ($r=-0.32$; 95% CI -0.57 to -0.09;

$p=0.01$; Figure 3; Supplementary Table 5), such that depressed patients who had greater difficulties with concentration and greater inner tension showed even more negative MPFC–PCC connectivity. Additional tests for associations between MPFC–PCC connectivity and clinical and behavioral variables, and between the concentration/tension factor and clinical and behavioral variables, are presented in Supplementary Table 5.

Comparison of depressed participants with co-morbid social anxiety disorder ($n=24$) with depressed participants without an anxiety disorder ($n=23$) showed significantly more negative modulation of MPFC–PCC connectivity in the former group (odds ratio=0.44; 95% CI 0.22 to 0.87; $p=0.02$). Details of the subgroups with and without social anxiety disorder in Supplementary Table 6.

Discussion

We studied brain network functioning in a large sample of unmedicated help-seeking depressed participants, finding specific and novel disturbances in effective connectivity associated with self-appraisal processes. We first validated the same dynamic network model in depressed participants as in the control participants; one in which the medial prefrontal cortex acts to direct broad self-related processes by regulating posterior cingulate cortex activity, which is enhanced during explicit self-appraisal. The magnitude of this negative influence is even more pronounced in depression, suggesting that the medial prefrontal cortex has a ‘hyper-regulatory’ influence in the illness.

The medial prefrontal cortex, particularly the most rostral portion (Brodmann Area 10) implicated in our model, is activated by a diverse range of tasks (39). They have in common that they require integration of stimuli from different sources, such as in analogical reasoning, where one set of concepts is mapped to another (40). In integrating stimuli from different sources, the rostral medial prefrontal cortex has a role in selectively attending to and switching between them (41), thereby supporting cognitive flexibility (42).

These integrative processes are important in the self-appraisal task, which requires mapping abstract semantic concepts to internal representations. In our model, the medial prefrontal cortex appears to modify self-representations generated by posterior cortical regions to integrate them with the meaning of the personality adjectives (8).

The behavioral results of our study confirm that the self is appraised more negatively in depression, and show that this is associated with more negative MPFC–PCC connectivity in the depressed compared to the control participants. The self-concept shows complex changes in depression: it is not only more negatively appraised, but as a result of the changing sense of self, is also less stable (43). This more aversive and unstable self is likely to take more cognitive effort to appraise, consistent with our understanding of rostral medial prefrontal cortical function.

The particular aspects of self-appraisal that are associated with the MPFC–PCC connectivity parameter are difficult to confirm with our data, although we have provisional evidence. The strength of MPFC–PCC connectivity was associated with a depression factor composed of difficulty with concentration and inner tension. Depressed participants with high levels of the factor may have had more difficulty with the conceptual integration and cognitive flexibility demanded by the self-appraisal task. Of note, these participants were biased towards affirming the personality adjectives (see Supplementary Table 5), likely exacerbating the appraisal demands. The demands of the task were conceivably even greater in participants with co-morbid depression and social anxiety disorder. In both disorders there is evidence of an increase in self-focused attention that is accompanied by difficulties shifting attention between self-related and external stimuli (44).

Our study has examined one aspect of self-related processing – the cognitive appraisal of the self’s personality attributes. Other aspects of the self are also disturbed in depression – for example, in its relationship to negative memory biases (45) and interoceptive awareness (46). Future studies could extend our model to study the importance of medial prefrontal cortex in regulating the relevant networks. Similarly, we observed effects related to posterior cingulate cortex function that suggest broader imbalance in the default mode network’s contributions to self-disturbance, which could also be elaborated in future studies.

Other limitations of our study should be noted. It is not clear that the connectivity changes we have demonstrated are specific to depression. We showed that MPFC–PCC connectivity was even more disturbed in patients with comorbid major depressive disorder and social anxiety disorder, but we did not have a group with social anxiety disorder alone to clarify to what extent the changes might be attributable to the disorder. Second, while there are clear advantages to studying young patients

early in the course of illness – during a period in which self-related concepts are particularly salient and depression is less chronic and influenced by treatments – the focus limits the extent to which the results might be generalized to the broader population of depressed adults. It is also a period in which neurodevelopmental processes are ongoing – a factor likely related to increased depression vulnerability – and while we accounted for linear age-related changes, it is possible that brain connectivity differences will be less prominent in later periods. It will be of interest to examine our model in older depressed populations. Last, factor analysis requires a degree of subjectivity in decisions related to the procedure, with no objective criteria for an optimal solution (47). Our examination of connectivity associations with the MADRS concentration/tension factor should therefore be considered exploratory.

Our study of a large group of unmedicated participants with moderate-to-severe depression allowed us to apply a neural modeling approach that was able to more fully characterize the connectivity parameters that underlie disturbed self-appraisal processes in depression, and map them to aspects of the clinical presentation. Our model demonstrates that the medial prefrontal cortex has a critical role in coordinating an integrated and dynamic neural representation of the self. Many studies have implicated abnormal medial prefrontal cortical function in depression (9–18), which these findings suggest might be a reflection, in part, of disturbances in the way medial prefrontal cortex directs self-appraisal processes. The connectivity disturbances suggest themselves as targets for cognitive behavioral therapy and other therapies that aim to affect change in the way the self is perceived, and might aid a mechanistic understanding of therapy response.

Competing interests

The authors have no competing interests to declare.

Acknowledgements

This study was funded by National Health and Medical Research Council of Australia (NHMRC)

Project Grants 1064643 (Principal Investigator BJH) and 1024570 (Principal Investigator CGD).

CGD was supported by an NHMRC Career Development Fellowship (1061757). We thank R.

Kerestes and K. Stephanou for their contributions to data collection, as well as staff from the Sunshine

Hospital Medical Imaging Department (Western Health, Melbourne).

References

1. Freud S: Mourning and melancholia. London, Hogarth Press, 1953,
2. Beck AT: Cognitive Therapy of Depression, Guilford Press, 1979
3. Klerman GL, Weissman MM, Rounsaville BJ, Chevron ES: Interpersonal Psychotherapy of Depression, New York, Basic Books, 1984
4. Davey CG, Allen NB, Harrison BJ, Dwyer DB, Yücel M: Being liked activates primary reward and midline self-related brain regions. *Hum Brain Mapp* 2010;31:660–668
5. Northoff G, Bermpohl F: Cortical midline structures and the self. *Trends Cogn Sci* 2004;8:102–107
6. Baird B, Smallwood J, Schooler JW: Back to the future: autobiographical planning and the functionality of mind-wandering. *Conscious Cogn* 2011;20:1604–1611
7. Schilbach L, Eickhoff SB, Rotarska-Jagiela A, Fink GR, Vogeley K: Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the “default system” of the brain. *Conscious Cogn* 2008;17:457–467
8. Davey CG, Pujol J, Harrison BJ: Mapping the self in the brain’s default mode network. *NeuroImage* 2016;132:390–397
9. Drevets WC, Videen TO, Price JL, Preskorn SH, Carmichael ST, Raichle ME: A functional anatomical study of unipolar depression. *J Neurosci* 1992;12:3628–3641
10. Kennedy SH, Evans KR, Krüger S, Mayberg HS, Meyer JH, McCann S, Arifuzzman AI, Houle S, Vaccarino FJ: Changes in regional brain glucose metabolism measured with positron emission tomography after paroxetine treatment of major depression. *Am J Psychiatry* 2001;158:899–905
11. Mayberg HS, Lozano AM, Voon V, McNeely HE, Seminowicz D, Hamani C, Schwalb JM, Kennedy SH: Deep brain stimulation for treatment-resistant depression. *Neuron* 2005;45:651–660
12. Davey CG, Harrison BJ, Yücel M, Allen NB: Regionally specific alterations in functional connectivity of the anterior cingulate cortex in major depressive disorder. *Psychol Med* 2012;42:2071–2082
13. Sheline YI, Price JL, Yan Z, Mintun MA: Resting-state functional MRI in depression unmasks increased connectivity between networks via the dorsal nexus. *Proc Natl Acad Sci U S A* 2010;107:11020–11025
14. Greicius MD, Flores BH, Menon V, Glover GH, Solvason HB, Kenna H, Reiss AL, Schatzberg AF: Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biol Psychiatry* 2007;62:429–437
15. Lemogne C, le Bastard G, Mayberg H, Volle E, Bergouignan L, Lehericy S, Allilaire JF, Fossati P: In search of the depressive self: extended medial prefrontal network during self-referential processing in major depression. *Soc Cogn Affect Neurosci* 2009;4:305–312
16. Yoshimura S, Okamoto Y, Onoda K, Matsunaga M, Okada G, Kunisato Y, Yoshino A, Ueda K, Suzuki S, Yamawaki S: Cognitive behavioral therapy for depression changes medial prefrontal and ventral anterior cingulate cortex activity associated with self-referential processing. *Soc Cogn Affect Neurosci* 2014;9:487–493
17. Grimm S, Ernst J, Boesiger P, Schuepbach D, Hell D, Boeker H, Northoff G: Increased self-focus in major depressive disorder is related to neural abnormalities in subcortical-cortical midline structures. *Hum Brain Mapp* 2009;30:2617–2627
18. Johnson MK, Nolen-Hoeksema S, Mitchell KJ, Levin Y: Medial cortex activity, self-reflection and depression. *Soc Cogn Affect Neurosci* 2009
19. Corlett PR, Fletcher PC: Computational psychiatry: a Rosetta Stone linking the brain to mental illness. *The Lancet Psychiatry* 2014;1:399–402
20. Friston KJ, Harrison L, Penny W: Dynamic causal modelling. *NeuroImage* 2003;19:1273–1302
21. Pfeifer JH, Peake SJ: Self-development: integrating cognitive, socioemotional, and neuroimaging perspectives. *Dev Cogn Neurosci* 2012;2:55–69
22. Sebastian C, Burnett S, Blakemore SJ: Development of the self-concept during adolescence. *Trends Cogn Sci* 2008;12:441–446

23. Costello EJ, Copeland W, Angold A: Trends in psychopathology across the adolescent years: what changes when children become adolescents, and when adolescents become adults. *J Child Psychol Psychiatry* 2011;52:1015–1025
24. Davey CG, Yucel M, Allen NB: The emergence of depression in adolescence: Development of the prefrontal cortex and the representation of reward. *Neurosci Biobehav Rev* 2008;32:1–19
25. Gogtay N, Thompson PM: Mapping gray matter development: implications for typical development and vulnerability to psychopathology. *Brain Cogn* 2010;72:6–15
26. First MB, Spitzer RL, Gibbon M, Williams JBW: Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I), Washington, American Psychiatric Publishing, 1997
27. Wechsler D: Manual for the Wechsler Test of Adult Reading (WTAR), San Antonio, Texas, The Psychological Corporation, 2001
28. Anderson NH: Likableness ratings of 555 personality-trait words. *J Pers Soc Psychol* 1968;9:272–279
29. Stephan KE, Marshall JC, Penny WD, Friston KJ, Fink GR: Interhemispheric integration of visual processing during task-driven lateralization. *J Neurosci* 2007;27:3512–3522
30. Stephan KE, Penny WD, Moran RJ, den Ouden HE, Daunizeau J, Friston KJ: Ten simple rules for dynamic causal modeling. *NeuroImage* 2010;49:3099–3109
31. Kahan J, Foltynie T: Understanding DCM: ten simple rules for the clinician. *NeuroImage* 2013;83:542–549
32. Rigoux L, Stephan KE, Friston KJ, Daunizeau J: Bayesian model selection for group studies - revisited. *NeuroImage* 2014;84:971–985
33. Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ: Bayesian model selection for group studies. *NeuroImage* 2009;46:1004–1017
34. Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ, Schofield TM, Leff AP: Comparing families of dynamic causal models. *PLoS Comput Biol* 2010;6:e1000709
35. Benjamini Y, Yekutieli D: The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* 2001;11:165–1188
36. Suzuki A, Aoshima T, Fukasawa T, Yoshida K, Higuchi H, Shimizu T, Otani K: A three-factor model of the MADRS in major depressive disorder. *Depress Anxiety* 2005;21:95–97
37. Craighead WE, Evans DD: Factor analysis of the Montgomery-Asberg Depression Rating Scale. *Depression* 1996;4:31–33
38. Montgomery SA, Asberg M: A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979;134:382–389
39. Burgess PW, Wu H-C: Rostral prefrontal cortex (Brodmann area 10): Metacognition in the brain. Edited by Stuss DT, Knight RT. Oxford, UK, Oxford University Press, 2013, pp 524–534
40. Volle E, Gilbert SJ, Benoit RG, Burgess PW: Specialization of the rostral prefrontal cortex for distinct analogy processes. *Cereb Cortex* 2010;20:2647–2659
41. Burgess PW, Dumontheil I, Gilbert SJ: The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends Cogn Sci* 2007;11:290–298
42. Kim C, Johnson NF, Cilles SE, Gold BT: Common and distinct mechanisms of cognitive flexibility in prefrontal cortex. *J Neurosci* 2011;31:4771–4779
43. Dobson KS, Shaw BF: Specificity and stability of self-referent encoding in clinical depression. *J Abnorm Psychol* 1987;96:34–40
44. Mor N, Winquist J: Self-focused attention and negative affect: a meta-analysis. *Psychol Bull* 2002;128:638–662
45. Dalgleish T, Watts FN: Biases of attention and memory in disorders of anxiety and depression. *Clin Psychol Rev* 1990;10:589–604
46. Paulus MP, Stein MB: Interoception in anxiety and depression. *Brain Struct Funct* 2010;214:451–463
47. Henson RK: Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement* 2006;66:393–416

Tables

Table 1: Demographic and clinical characteristics of the participants, with their behavioral responses

	Controls		MDD		Controls vs. MDD	
	N	%	N	%	<i>p</i> -value	
Participants	88		71			
Females	46	52.3	37	52.2		0.99
Co-morbid anxiety disorder – social anxiety disorder			47	66.2		
			25	35.2		
	Mean	S.D.	Mean	S.D.	<i>d</i>	<i>p</i> -value
Age, y	20.1	2.9	20.2	2.7	-0.05	0.75
FSIQ	110.3	8.0	107.1	8.3	0.39	0.01
MADRS	1.8	2.5	32.8	4.7	-8.47	<0.0001
Accuracy on external attention task, %	96.4	6.7	95.4	4.5	0.17	0.30
Mean reaction time for external attention task, s	1.92	0.40	2.01	0.40	-0.22	0.18
Mean reaction time for self- appraisal task, s	1.68	0.35	1.70	0.29	-0.05	0.78
Proportion of personality adjectives affirmed, %	50.5	8.1	56.9	11.3	-0.66	0.0001
Likeableness of adjectives affirmed*	3.4	0.2	3.0	0.2	1.81	<0.0001

*The likeableness ratings of the personality adjectives (rated from 0–6) were provided in the original article by Anderson (28).

Table 2: Parameter estimates for each group and their comparison

Model parameters	Controls		MDD		Controls vs. MDD		
	Mean	95% CI	Mean	95% CI	Odds ratio	95% CI	<i>p</i> -value
Fixed							
MPFC→PCC	-0.18	-0.23 to -0.13***	-0.25	-0.31 to -0.20***	0.29	0.06 to 1.33	0.11
MPFC→IPL	-0.04	-0.09 to 0.01	0.01	-0.04 to 0.05	2.85	0.53 to 15.34	0.22
PCC→MPFC	0.20	0.16 to 0.25***	0.25	0.20 to 0.31***	2.89	0.52 to 16.10	0.23
PCC→IPL	0.33	0.28 to 0.38***	0.34	0.29 to 0.38***	0.87	0.16 to 4.84	0.88
IPL→MPFC	0.11	0.06 to 0.15***	0.12	0.06 to 0.17***	1.01	0.21 to 4.97	0.99
IPL→PCC	0.01	-0.06 to 0.08	-0.06	-0.13 to 0.02	0.58	0.20 to 1.71	0.32
Modulatory							
MPFC→PCC	-0.52	-0.73 to -0.32***	-1.13	-1.40 to -0.86***	0.54	0.38 to 0.77	0.0008†
MPFC→IPL	-0.07	-0.18 to 0.04	-0.17	-0.28 to -0.05*	0.72	0.35 to 1.46	0.36
PCC→MPFC	0.62	0.42 to 0.82***	0.30	0.18 to 0.41***	0.54	0.32 to 0.90	0.02
PCC→IPL	0.56	0.32 to 0.80***	0.37	0.17 to 0.57**	0.80	0.55 to 1.15	0.23

MPFC = medial prefrontal cortex; PCC = posterior cingulate cortex; IPL = inferior parietal lobule

*** One-sample t-test, *p*<0.0001

** One-sample t-test, *p*<0.001

* One-sample t-test, *p*<0.01

† Logistic regression, *q*<0.05 (false discovery rate correction)

Figure legends

Fig. 1. Identification of the network nodes. Regions that showed activation to the rest-fixation and self-appraisal conditions (compared to the external attention baseline), and even greater response to self-appraisal, were identified in the medial prefrontal cortex (MPFC), posterior cingulate cortex (PCC), and left inferior parietal lobule (IPL) in the control and depressed participants. We extracted time-series for each participant from these regions. The self-appraisal condition caused greater posterior cingulate cortex activation in the depressed compared to the control participants ($p=0.01$), with no differences demonstrated in the other regions. MDD = major depressive disorder.

Fig. 2. Parameter estimates determined by Bayesian model averaging for the optimal model in the control and depressed participants. In both groups, fixed connectivity comprised posterior cingulate cortex (PCC) having positive influence on medial prefrontal cortex (MPFC) – both directly, and indirectly via its influence on inferior parietal lobule (IPL) – and medial prefrontal cortex having negative influence on posterior cingulate cortex. These effects were enhanced by the modulatory effects of self-appraisal: posterior cingulate cortex had a greater positive influence on medial prefrontal cortex and inferior parietal lobule, and medial prefrontal cortex had a greater negative influence on posterior cingulate cortex. In the depressed, but not the control participants, medial prefrontal cortex had a negative modulatory influence on inferior parietal lobule. MDD = major depressive disorder.

Fig. 3. Pairwise examination of the dynamic causal modeling connectivity parameters in the depressed and control participants. The modulatory effects of self-appraisal were significantly more negative for MPFC–PCC connectivity in the depressed compared to the control participants ($p=0.0008$). There was also a trend for more negative modulatory effects of self-appraisal on MPFC–PCC connectivity ($p=0.02$). MPFC = medial prefrontal cortex; PCC = posterior cingulate cortex; IPL = inferior parietal lobule; MDD = major depressive disorder.

Fig. 4. Individual differences in MPFC–PCC connectivity among the depressed participants. The strength of modulation of MPFC–PCC connectivity was significantly negatively correlated with a MADRS factor composed of difficulties with concentration and inner tension.

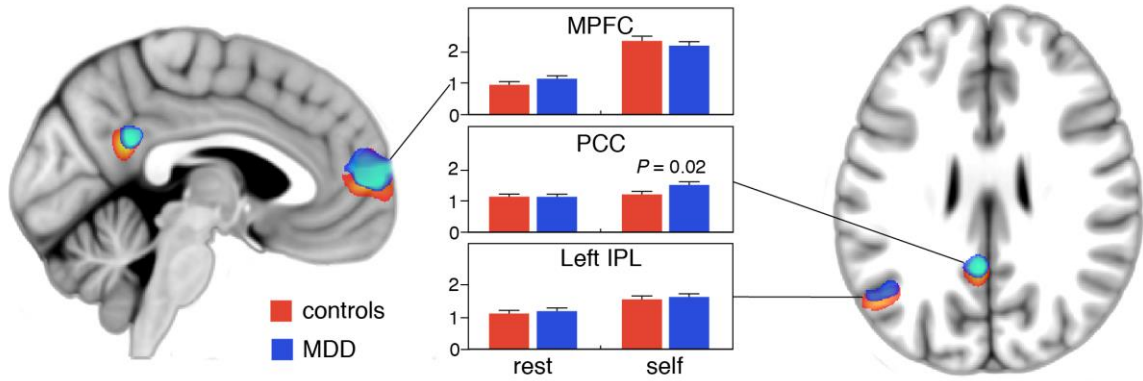


Fig 1

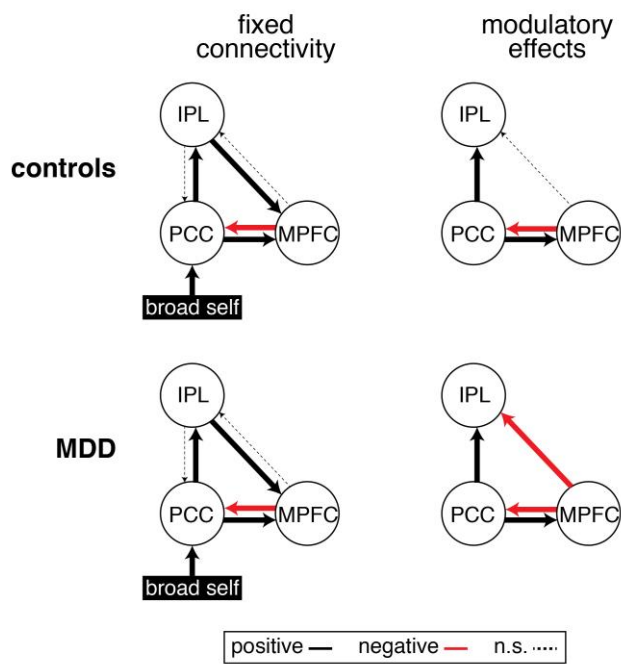


Fig 2

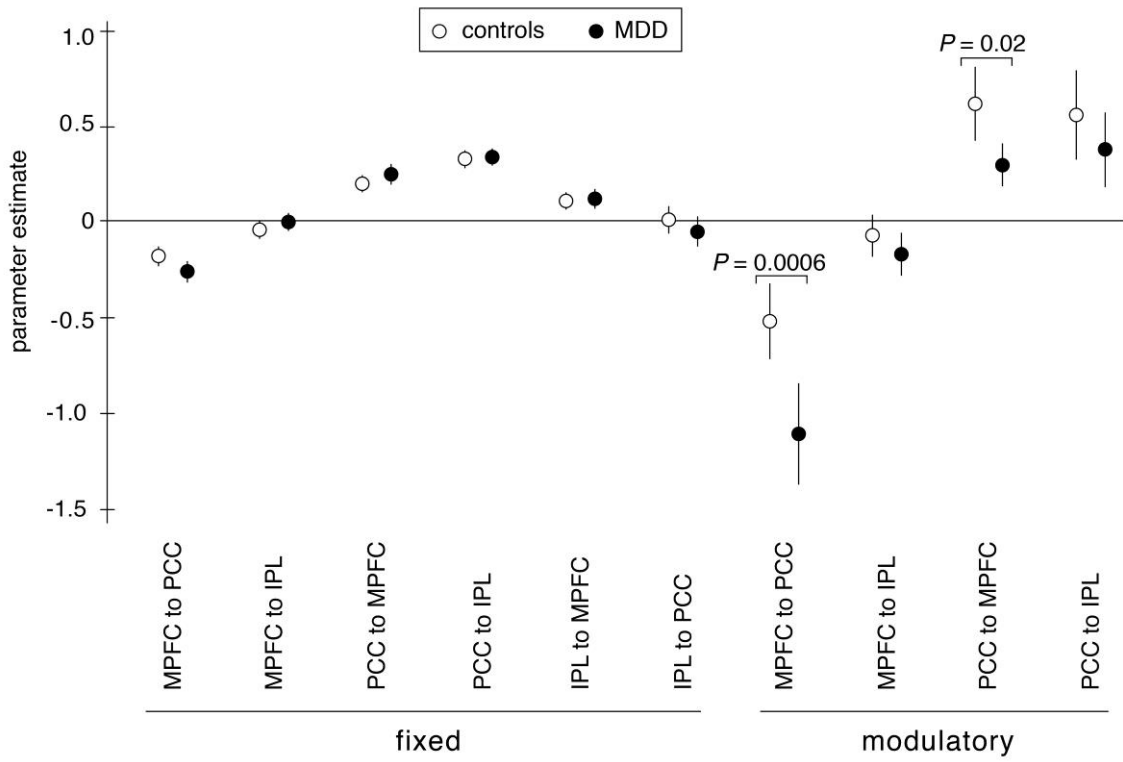


Fig 3

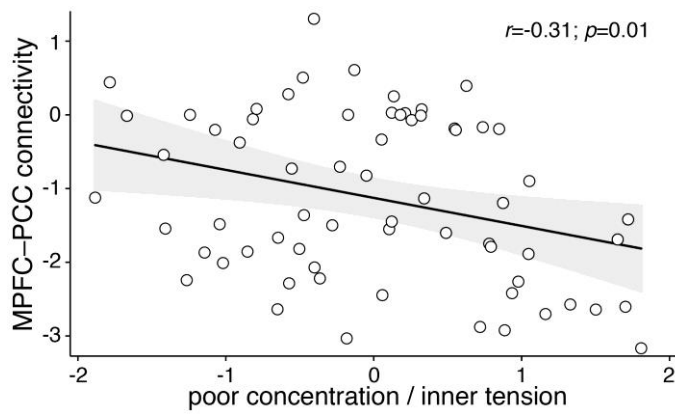


Fig 4

Supplementary Table 1. Activations for the conjunction of rest-fixation, and self-reference greater than rest-fixation.

Brain region	Controls					MDD				
	Cluster size	Peak t-value	Peak co-ordinate			Cluster size	Peak t-value	Peak co-ordinate		
			x	y	z			x	y	z
Medial prefrontal cortex	771	8.1	2	60	10	683	8.7	0	60	10
Posterior cingulate cortex	164	9.1	-4	-50	26	110	7.8	-4	-48	28
Left inferior parietal lobule	226	7.3	-50	-66	26	97	6.2	-46	-62	28
Left middle temporal gyrus	52	6.2	-58	-16	-10	–	–	–	–	–

Only clusters of 50 or more voxels are reported.

Supplementary Table 2: Effects of sex on parameter estimates for each group and for the groups combined.

Model parameters	Controls			MDD			Combined		
	Odds ratio	95% CI	<i>p</i> -value	Odds ratio	95% CI	<i>p</i> -value	Odds ratio	95% CI	<i>p</i> -value
Fixed									
MPFC→PCC	0.81	0.11 to 5.88	0.84	3.64	0.89 to 14.79	0.07	2.85	0.62 to 12.96	0.18
MPFC→IPL	1.25	0.14 to 11.04	0.84	4.25	0.24 to 76.33	0.33	2.32	0.42 to 12.80	0.33
PCC→MPFC	0.44	0.04 to 4.75	0.50	0.11	0.01 to 1.79	0.12	0.33	0.06 to 1.84	0.20
PCC→IPL	1.26	0.15 to 10.73	0.84	0.31	0.02 to 5.85	0.43	0.73	0.13 to 3.99	0.72
IPL→MPFC	2.68	0.28 to 25.38	0.86	0.90	0.07 to 10.73	0.93	1.60	0.31 to 8.16	0.58
IPL→PCC	0.45	0.11 to 1.95	0.29	0.56	0.10 to 3.19	0.52	0.47	0.16 to 1.42	0.18
Modulatory									
MPFC→PCC	0.86	0.52 to 1.45	0.58	1.30	0.79 to 2.13	0.30	1.00	0.72 to 1.40	0.98
MPFC→IPL	0.80	0.31 to 2.08	0.65	0.41	0.11 to 1.46	0.17	0.62	0.29 to 1.32	0.22
PCC→MPFC	0.80	0.46 to 1.37	0.42	0.59	0.19 to 1.83	0.36	0.70	0.43 to 1.14	0.15
PCC→IPL	1.02	0.66 to 1.58	0.93	1.82	0.87 to 3.78	0.11	1.15	0.80 to 1.65	0.46

MPFC = medial prefrontal cortex; PCC = posterior cingulate cortex; IPL = inferior parietal lobule
 An odds ratio > 1 indicates that parameter values are greater in males compared to females.

Supplementary Table 3: Associations between age and parameter estimates for each group and for the groups combined.

Model parameters	Controls			MDD			Combined		
	Correl. coeff.	95% CI	<i>p</i> -value	Correl. coeff.	95% CI	<i>p</i> -value	Correl. coeff.	95% CI	<i>p</i> -value
Fixed									
MPFC→PCC	-0.20	-0.42 to 0.02	0.07	-0.01	-0.25 to 0.23	0.97	-0.11	-0.27 to 0.05	0.20
MPFC→IPL	0.00	-0.22 to 0.22	0.99	-0.22	-0.44 to 0.02	0.08	-0.05	-0.21 to 0.11	0.57
PCC→MPFC	0.19	-0.03 to 0.39	0.09	-0.02	-0.26 to 0.22	0.85	0.10	-0.06 to 0.26	0.21
PCC→IPL	0.06	-0.16 to 0.27	0.57	0.18	-0.06 to 0.40	0.16	0.09	-0.07 to 0.25	0.28
IPL→MPFC	0.02	-0.20 to 0.24	0.86	0.07	-0.17 to 0.30	0.60	0.03	-0.13 to 0.19	0.71
IPL→PCC	0.09	-0.13 to 0.30	0.41	0.14	-0.10 to 0.37	0.27	0.07	-0.09 to 0.23	0.42
Modulatory									
MPFC→PCC	-0.08	-0.29 to 0.14	0.50	0.06	-0.18 to 0.29	0.63	-0.05	-0.21 to 0.11	0.53
MPFC→IPL	-0.04	-0.25 to 0.18	0.70	-0.04	-0.28 to 0.20	0.76	-0.04	-0.20 to 0.12	0.60
PCC→MPFC	0.15	-0.07 to 0.36	0.19	0.11	-0.13 to 0.34	0.41	0.04	-0.12 to 0.20	0.64
PCC→IPL	0.02	-0.20 to 0.24	0.81	0.09	-0.15 to 0.32	0.48	0.01	-0.15 to 0.17	0.86

MPFC = medial prefrontal cortex; PCC = posterior cingulate cortex; IPL = inferior parietal lobule

Supplementary Table 4. Factor loadings of the MADRS items.

MADRS item	Factor 1	Factor 2	Factor 3	Factor 4
Apparent sadness	0.86			
Reported sadness	0.75			
Inability to feel	0.48			
Reduced appetite		0.69		
Reduced sleep		0.67		
Pessimistic thoughts		0.67		
Lassitude			0.74	
Suicidal thoughts			0.68	
Concentration difficulties				0.81
Inner tension				0.47
% variance explained	17.9	16.4	14.5	13.1

Supplementary Table 5. Associations between parameters of interest and clinical and behavioral variables.

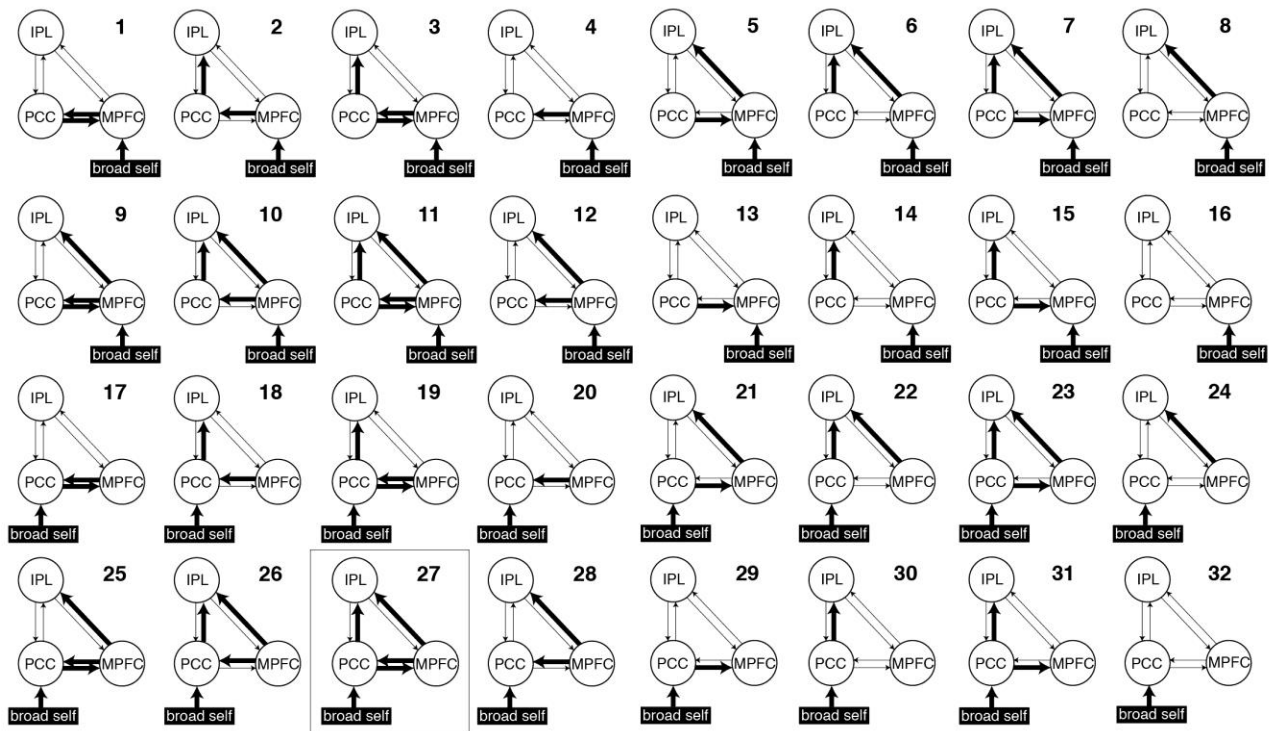
	Correlation coefficient	95% CI	<i>p</i>-value	<i>q</i>-value
Association with MPFC–PCC connectivity:				
Total MADRS score	0.03	-0.21 to 0.27	0.79	0.79
MADRS sadness factor	-0.08	-0.31 to 0.17	0.57	0.72
MADRS somatic/pessimism factor	0.17	-0.08 to 0.39	0.21	0.61
MADRS lassitude/suicidality factor	0.10	-0.14 to 0.38	0.45	0.67
MADRS concentration/tension factor	-0.32	-0.51 to -0.08	0.01	0.04*
Proportion of personality adjectives affirmed	0.15	-0.10 to 0.38	0.27	0.61
Likeableness of personality adjectives affirmed	-0.06	-0.30 to 0.18	0.64	0.72
Association with MADRS concentration/tension factor				
Proportion of personality adjectives affirmed	0.35	0.10 to 0.54	0.008	0.04*
Likeableness of personality adjectives affirmed	-0.10	-0.34 to 0.14	0.45	0.67

q-values represent *p*-values that have been adjusted using the false discovery rate.

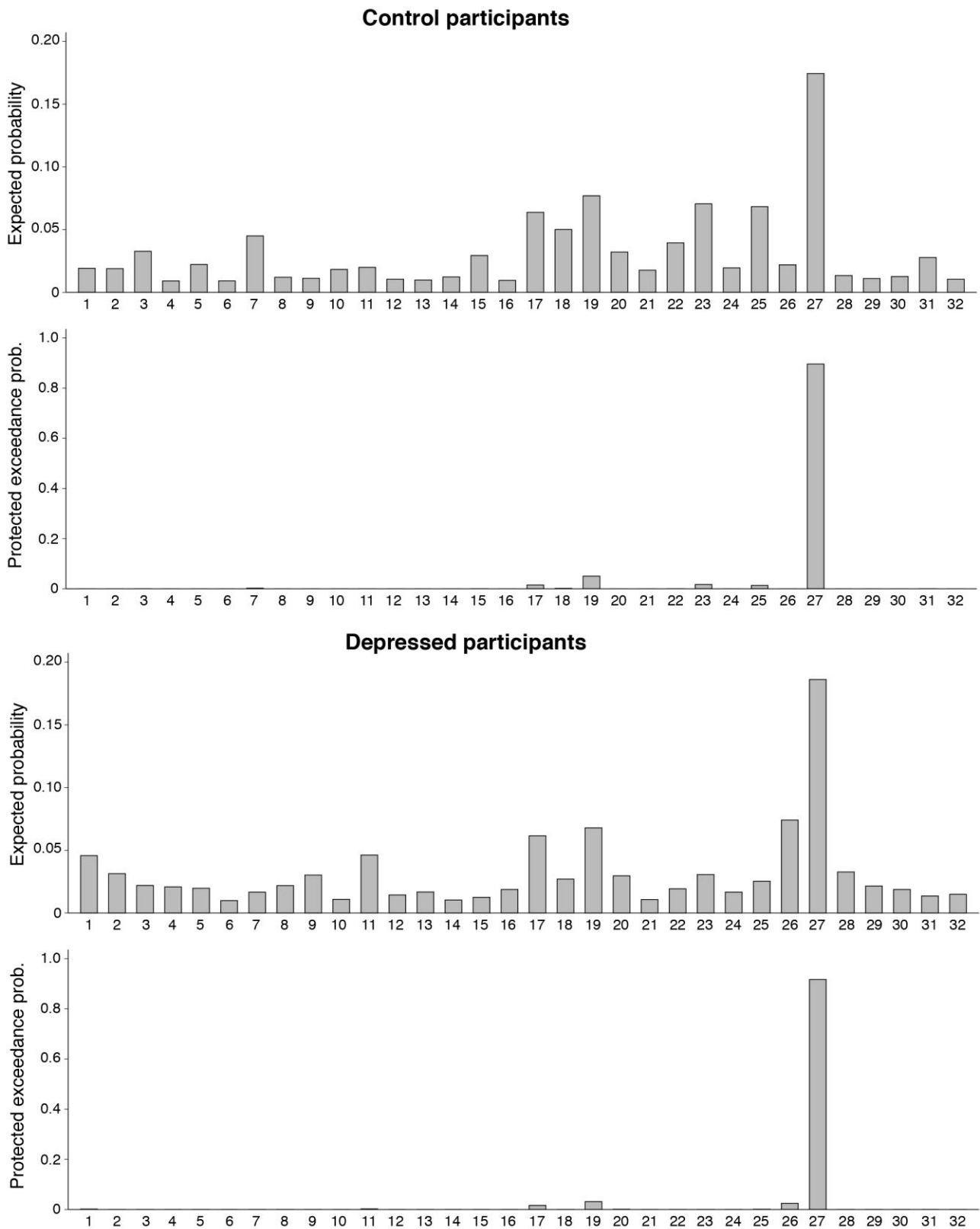
*significant at $q < 0.05$

Supplementary Table 6: Comparison of participants who have depression and social anxiety disorder with participants who have depression without an anxiety disorder

	MDD without anxiety disorder		MDD with social anxiety disorder		<i>p</i> -value
	N	%	N	%	
Participants	23		24		
Females	8	34.8	16	66.7	0.03
	Mean	S.D.	Mean	S.D.	
Age, <i>y</i>	19.3	2.6	20.3	2.8	0.16
FSIQ	106.7	9.2	107.0	6.8	
MADRS	34.3	4.8	32.0	4.3	0.10
Accuracy on external attention task, %	95.7	4.5	94.3	4.6	0.27
Mean reaction time for external attention task, <i>s</i>	2.1	0.4	2.0	0.4	0.44
Mean reaction time for self-appraisal task, <i>s</i>	1.8	0.3	1.6	0.3	0.08
Proportion of personality adjectives affirmed, %	56.7	10.8	55.6	11.4	0.99
Likeableness of adjectives affirmed	3.0	0.3	3.1	0.2	0.42
	Correl. coeff.	s.e.	Correl. coeff.	s.e.	
Correlation between conc./tension factor and MPFC–PCC connectivity, <i>r</i>	-0.40	0.20	-0.26	0.21	0.62



Supplementary Figure 1. The DCM candidate model-space. Network connections between the core-self regions were elaborated in 32 models, which varied according to whether broad self (rest-fixation and self-appraisal) drove network activity via MPFC or PCC, and whether self-appraisal modulated efferent connections from MPFC to PCC, IPL, both, or neither, and/or connections from PCC to MPFC, IPL, both, or neither (i.e., a $2 \times 4 \times 4 = 32$ candidate model-space). Bayesian model selection favored model 27 in both the depressed and control groups. MPFC = medial prefrontal cortex; PCC = posterior cingulate cortex; IPL = inferior parietal lobule.



Supplementary Figure 2. Bayesian model selection. BMS clearly favored model 27 in both the control participants (expected probability = 0.17, protected exceedance probability = 0.90) and the depressed participants (expected probability = 0.19, protected exceedance probability = 0.92).