



ARTICLE

Aberrant working memory processing in major depression: evidence from multivoxel pattern classification

Matti Gärtner^{1,2}, M. Elisabetta Ghisu^{3,4}, Milan Scheidegger⁵, Luisa Bönke¹, Yan Fan¹, Anna Stippl¹, Ana-Lucia Herrera-Melendez¹, Sophie Metz¹, Emilia Winnebeck⁶, Maria Fissler⁶, Anke Henning^{7,8}, Malek Bajbouj¹, Karsten Borgwardt^{3,4}, Thorsten Barnhofer^{6,9} and Simone Grimm^{1,2,5}

Major depressive disorder (MDD) is often accompanied by severe impairments in working memory (WM). Neuroimaging studies investigating the mechanisms underlying these impairments have produced conflicting results. It remains unclear whether MDD patients show hyper- or hypoactivity in WM-related brain regions and how potential aberrations in WM processing may contribute to the characteristic dysregulation of cognition–emotion interactions implicated in the maintenance of the disorder. In order to shed light on these questions and to overcome limitations of previous studies, we applied a multivoxel pattern classification approach to investigate brain activity in large samples of MDD patients ($N = 57$) and matched healthy controls ($N = 61$) during a WM task that incorporated positive, negative, and neutral stimuli. Results showed that patients can be distinguished from healthy controls with good classification accuracy based on functional activation patterns. ROI analyses based on the classification weight maps showed that during WM, patients had higher activity in the left DLPFC and the dorsal ACC. Furthermore, regions of the default-mode network (DMN) were less deactivated in patients. As no performance differences were observed, we conclude that patients required more effort, indexed by more activity in WM-related regions, to successfully perform the task. This increased effort might be related to difficulties in suppressing task-irrelevant information reflected by reduced deactivation of regions within the DMN. Effects were most pronounced for negative and neutral stimuli, thus pointing toward important implications of aberrations in WM processes in cognition–emotion interactions in MDD.

Neuropsychopharmacology (2018) 43:1972–1979; <https://doi.org/10.1038/s41386-018-0081-1>

INTRODUCTION

It is well established that major depressive disorder (MDD) is accompanied by disruptions in different cognitive domains [1]. Working memory (WM) is fundamental to the performance of many cognitive tasks and day-to-day activities [2], and deficits in WM are top-ranked endophenotype candidates for recurrent MDD [3]. Neuroimaging studies that have tried to shed light on mechanisms associated with WM performance in MDD have reported conflicting results, with regard to whether MDD patients show hyper- or hypofrontality during WM. Studies in healthy subjects clearly demonstrated that the dorsolateral prefrontal cortex (DLPFC) is implicated in numerous cognitive functions relevant to WM. Activity in the anterior cingulate cortex (ACC) during WM tasks is often described in relation to increased effort, complexity, or attention [4, 5]. Since increased frontal task-related activation has been described when MDD patients show no behavioral WM impairments, it has been hypothesized that they need greater activation within the same neural network to maintain a similar level of performance as healthy control subjects

(HC) [6–8]. In this theoretical model, hyperfrontality in MDD occurs to compensate for a lack of deactivation in regions of the default-mode network (DMN [9]). By contrast, other studies reported hypoactivation in the ACC [10], and parietal cortex [11], while observing impaired performance during WM tasks in patients. Overall, these findings might suggest that intact performance in MDD is associated with increased cortical activity, while impaired performance is associated with reduced cortical activation. In line with this view, a recent meta-analysis by Wang et al. [12] matched task performance for predominantly verbal WM demands and reported hyperactivation exclusively in the left DLPFC and hypoactivation in precuneus and insula. However, even though this meta-analysis included 11 studies and 160 patients, data were derived from 13 different WM experiments utilizing a wide range of stimuli, thereby making an interpretation of patterns of hyper- and hypoactivation in distinct brain regions rather difficult. Such difficulties are further underlined by another recent meta-analysis of 34 individual neuroimaging experiments testing cognitive processing in MDD patients that reported no significant results

¹Department of Psychiatry and Psychotherapy, Charité—Universitätsmedizin Berlin, Campus Benjamin Franklin, Hindenburgdamm 30, 12203 Berlin, Germany; ²MSB—Medical School Berlin, Calandrellistraße 1-9, 12247 Berlin, Germany; ³Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 26, 4058 Basel, Switzerland; ⁴SIB Swiss Institute of Bioinformatics, Basel, Switzerland; ⁵Department of Psychiatry, Psychotherapy and Psychosomatics, Psychiatric Hospital, University of Zurich, 8032 Zurich, Switzerland; ⁶Dahlem Institute for Neuroimaging of Emotion, Freie Universität, Habelschwerdter Allee 45, 14195 Berlin, Germany; ⁷Max Planck Institute for Biological Cybernetics, Tübingen, Germany; ⁸Institute of Physics, Ernst-Moritz Arndt University Greifswald, Greifswald, Germany and ⁹Department of Psychology, University of Exeter, Sir Henry Wellcome Building for Mood Disorders Research, Perry Road, Exeter EX4 4QG, UK

Correspondence: Matti Gärtner (matti.gaertner@charite.de)

These authors contributed equally: Matti Gärtner, M. Elisabetta Ghisu

Received: 30 November 2017 Revised: 20 April 2018 Accepted: 23 April 2018

Published online: 2 May 2018

across studies. The authors of this meta-analysis also suggested that this finding is most likely due to differences in methods, i.e., utilized stimuli, experimental design, small and heterogeneous patient groups, and inappropriate statistical inference procedures [13].

Since impairments at the cognition–emotion interface rather than of cognitive functions per se may be most characteristic for MDD [1], it seems particularly worthwhile to investigate cognition–emotion interactions in the brain in order to broaden our understanding of WM deficits in MDD. Studies in HC demonstrated that emotional state and mood influence WM [14, 15], presumably via the activation of mood-congruent representations in WM [16], as negative mood is related to more frequent negative thoughts and to selective attention to negative stimuli [17, 18]. Consequently, negative biases in MDD, i.e., a pattern of cognition biased toward negative information and the resulting inability to reallocate attention away from negative emotional information, might prevent MDD patients from effectively completing ongoing cognitive tasks and result in WM deficits [19]. Furthermore, impairments in the ability to flexibly and efficiently update WM might prevent the removal of negative, no longer goal-relevant content from WM and thus facilitate perseverative thinking, such as rumination [20]. Accordingly, higher levels of maladaptive, depressive rumination have been associated with increasing levels of DMN dominance, i.e., increased DMN activation at rest [21]. Increased resting state activity in DMN regions is accompanied by a corresponding lack of task-induced deactivation, which has been related to cognitive deficits in MDD [9, 22].

The aim of the present study was to gain a better understanding of WM deficits in MDD, by overcoming some of the difficulties that might have caused inconsistent results in previous studies. By investigating a large sample of MDD patients and HC during a WM task that incorporated positive, negative, and neutral stimuli we aimed to probe WM processes in the context of cognition–emotion interactions in the brain. Moreover, rather than applying state-of-the-art statistical inference procedures, we applied a multivoxel pattern classification (MVPC) approach, to circumvent the main limitations of classical general linear model analysis and allow for the detection of distributed patterns of activity, while at the same time providing a solution to the multiple comparisons problem. We hypothesized that this approach would enable distinction between patients and HC based on WM-related brain activity. Our goal in this context was to design a model, which would allow to identify those brain regions, which are most informative for group discrimination. Specifically, we expected to find MDD-associated hyperactivation in WM-related regions such as the DLPFC and ACC, and a lack of deactivation in regions of the DMN. We expected these differences to be most pronounced for negative emotional stimuli and to be meaningfully associated with ruminative tendencies.

MATERIALS AND METHODS

Subjects

Male and female subjects with an acute MDD episode ($N=57$) and matched healthy control subjects ($N=61$) were recruited at the Free University of Berlin (FUB) and at the University of Zuerich (UZH) from responses to advertising in local newspapers and mailing lists. Additionally, patients were recruited at the Affective Disorders Unit at the Department of Psychiatry (UZH). Healthy subjects were screened for psychiatric disorders using the short version of the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders. Exclusion criteria for all subjects were major medical illness, history of seizures, head trauma with loss of consciousness, and pregnancy. Exclusion criteria for healthy controls furthermore included absence of present and past diagnosis of psychiatric or neurologic disease.

Specific psychiatric exclusion criteria for patients consisted of atypical forms of depression, suicidal ideation, any additional psychiatric disorder, history of substance abuse or dependence, and electroconvulsive therapy in the previous 6 months. We allowed patients who were currently taking antidepressants into the study provided that the medication had not been changed during the last 4 weeks before entry into the study. The study was carried out in accordance with the latest version of the Declaration of Helsinki. Participants entered into the study after a full explanation of the purpose of the study and the study procedures and after written consent was obtained as approved by the institutional review boards.

Task and procedure

Stimuli were German nouns taken from the Berlin Affective Word List (BAWL [23]). The stimuli were classified as positive, negative, and neutral according to the BAWL norms and matched according to arousal levels, imageability, and number of letters. The stimuli were consecutively presented within a 2-back WM task, which provides an established means of both studying the interface between WM and emotion and eliciting BOLD responses in cognition- and emotion-related regions [24, 25]. Each block consisted of 15 words of either positive, negative, or neutral valence presented for 500 ms with an interstimulus interval of 1500 ms and was followed by a fixation trial (10–14 s). In total, the task consisted of 15 blocks (5 of each valence category), and a total of 225 stimuli were presented. Participants responded to the stimuli by pushing a fiber-optic light-sensitive key press.

Psychometric measures

The Beck Depression Inventory [26] was used to determine depression severity. The Response Style Questionnaire was used to measure trait-like coping styles that are not associated with state effects of depressed mood [27].

Functional magnetic resonance imaging data acquisition and analysis

Functional data were acquired on a Siemens Trio 3T (FUB), and a Philips Achieva 3T scanner (UZH) using standard echo planar imaging sequences [24, 25], and preprocessed in SPM12 (Wellcome Trust Centre for Neuroimaging, London, UK) using standard parameters (for details see Supplementary Materials and Methods). A fixed-effect model at a single-subject level was performed to create images of parameter estimates. For each subject, the following contrast images of parameter estimates were calculated: (1) all WM conditions versus fixation condition (WM > fixation); (2) positive WM condition versus fixation condition (Pos > fixation); (3) negative WM condition versus fixation condition (Neg > fixation); (4) neutral WM condition versus fixation condition (Neu > fixation); (5) emotional WM conditions versus neutral WM condition (Emo > neutral).

Multivoxel pattern classification

MVPC was used to discriminate patients from controls and to identify group associated patterns of activity. The above-mentioned contrast images were taken as input for the classifier. Support vector machines (SVM) with a linear kernel were employed for classification [28]. In order to determine the optimal model, SVM with three different feature selection strategies were tested and compared: SVM without feature selection, SVM with recursive feature elimination [29], and SVM with feature selection based on ranking using F -score values (SVM-fScore [30]).

The out-of-sample performance of the classifier was evaluated via a leave one out cross validation (LOOCV). The choice of LOOCV over larger folds was motivated by the relatively small number of samples available. Indeed, LOOCV is a common approach in neuroimaging studies, because of the limited sample sizes compared to machine learning applications in other domains.

Table 1. Demographic, clinical, and behavioral data

	MDD (N = 57)	HC (N = 61)	Group statistics
Age (M, SD)	40.5 ± 12.7	38.3 ± 10.1	$t(116) = 1.04, p > 0.1$
Sex (m/f)	25/32	35/26	$\chi^2(1, N = 118) = 2.15, p > 0.1$
BDI (M, SD)	27.42 ± 8.28	na	na
No. of episodes (M, SD)	7.51 ± 5.43	na	na
Medication status ^a (med. free/on med.)	35/22	na	na
WM accuracy (in %; M, SD)	73.6 ± 16.5	70.3 ± 30.0	$t(115) = 0.73, p > 0.1$
WM reaction time (in ms; M, SD)	646 ± 137	590 ± 179	$t(115) = 1.90, p > 0.05$

^aNumber of patients that took antidepressant medication during the study
MDD major depressive disorder patients, HC healthy control subjects

The classification weight maps for subsequent analyses were constructed by averaging the weights over all folds of the cross validation. The use of linear separation boundaries allowed a straightforward interpretation of the feature weights, implying that higher absolute weights corresponded to the most discriminative features. Due to the applied label convention, a positive weight sign indicated higher values in patients, and a negative weight sign indicated higher activation in controls. The performance of the classifier was evaluated in terms of accuracy, sensitivity, and specificity. All classification analyses were performed in Python (Python 2.7.10) and the Nilearn library (v0.2.6). Permutation tests with 1000 repetitions were used to assess the statistical significance of the obtained classification accuracies [31]. For control purposes, a 10-fold cross-validation approach to evaluate the out-of-sample performance of the classifier was implemented. Moreover, additional classification analyses on subject subgroups from the two scanning sites was performed to assess the robustness of the results. More details of the classification procedures are described in the Supplementary Materials and Methods.

Region of interest analyses

To further investigate potential differences between valence conditions of the WM task, post hoc region of interest (ROI) analyses were conducted on the SVM-fScore weight map of the WM > fixation contrast. To set the focus on the most relevant regions, 20% of the highest classification weights were kept, while the remaining 80% were masked. Additionally, a cluster threshold of 50 voxels was set to focus on the physiologically meaningful clusters.

Two types of mixed model analyses of variance (ANOVAs), both with group (MDD versus HC) as between-subject factor, were calculated on selected ROI data. In the first ANOVA, the within-subject factor condition was based on the contrast of parameter estimates (COPE) of the three WM conditions (Pos > fixation, Neg > fixation, and Neu > fixation). Since aberrant brain activity during rest has been reported in MDD [32], a second ANOVA including the fixation condition was conducted on the raw parameter estimates (RPE) of all conditions (fixation, positive, negative, and neutral). To account for multiple testing due to several parallel ANOVAs and post hoc t-tests, alpha levels were adjusted using Bonferroni correction.

RESULTS

Sample and WM performance

Patients and controls did not differ with regard to age and sex. Clinical, demographic, and behavioral data are summarized in Table 1. WM accuracy and reaction times were analyzed using two-way repeated measures ANOVAs with the factors condition (positive, negative, and neutral) and group (MDD and HC). There

Table 2. Results of the SVM-fScore classification on the contrasts of interest

Contrast	Accuracy (%)	Sensitivity (%)	Specificity (%)
WM > fixation	66.10***	68.42	63.93
Pos > fixation	63.56***	68.42	59.02
Neg > fixation	71.18***	71.93	70.49
Neu > fixation	73.73***	71.93	75.41
Emo > Neu	49.15 (ns)	64.91	34.43

Asterisks depict significant classification results based on permutation tests: *** $p < 0.001$; ns— $p < 0.1$ (not significant)

were no significant effects of group, condition and no significant interaction effects of these factors on WM accuracy and reaction times. However, results showed a trend for a significant effect of group on reaction times ($F_{1,115} = 3.544, p = 0.062$), with Bonferroni adjusted post hoc t-tests revealing marginally significantly slower reaction times in depressive patients for negative words ($t_{115} = -2.27, p = 0.025$). Behavioral data of one subject could not be analyzed because of a corrupted log-file. A negative correlation was found between WM accuracy and rumination (Pearson's coefficient, $r_{46} = -0.35, p = 0.017$).

Functional magnetic resonance imaging: classification results

The MVPC method revealed significant classification accuracies for all WM versus fixation contrasts as indicated by the results of the permutation tests. Highest classification accuracy was observed for the Neu > fixation contrast (73.7%) and for the Neg > fixation contrast (71.2%). Lower but still significant accuracies were obtained for the WM > fixation contrast (66.1%) and for the Pos > fixation contrast (63.6%), while no significant predictions were found for the Emo > Neu contrast (49.2%). Highest classification accuracies were obtained for the SVM-fScore method. All classification results for this classifier are provided in Table 2, and a comparison of the results of all three feature selection methods is shown in Supplementary Table S1. The 10-fold cross-validation results showed comparable performances to LOO, as reported in Supplementary Table S2. The classification results investigating the site effect showed that combining samples from the two sites provided either comparable or improved performances than single-site approaches. More information is given in the Supplementary Materials and Methods.

Functional magnetic resonance imaging: ROI results

The cluster extraction from the weight image of the WM > fixation contrast revealed 14 ROIs that met the cluster inclusion criteria. Seven of the ROIs were based on positive weights depicting higher activity in MDD compared to HC, and seven ROIs were

Table 3. ROIs extracted from weight map of the WM > fixation contrast

Region	Direction	Size (voxels)	Coordinates (center, mm)
Left DLPFC	MDD > HC	61	-38 34 29
Dorsal ACC, SMA	MDD > HC	170	-13 -3 54
PCC	HC > MDD	51	-2 -45 20
Right IPL	HC > MDD	111	53 -33 37
Left STG, insula	HC > MDD	164	-45 7 -6
Right MTG	MDD > HC	80	62 -36 -3
Precuneus	MDD > HC	69	7 -55 42
Left IPL	MDD > HC	81	-40 -37 49
LG, cuneus	MDD > HC	125	-3 -81 2
Thalamus	MDD > HC	127	1 -2 7
Precuneus/MOG	HC > MDD	59	-13 -66 28
Left anterior PFC	HC > MDD	123	-32 52 6
Cerebellum, declive	HC > MDD	133	0 -78 -25
Right MOG	HC > MDD	208	22 -87 8

MDD major depressive disorder patients, *HC* healthy control subjects, *DLPFC* dorsolateral prefrontal cortex, *MTG* middle temporal gyrus, *IPL* inferior parietal lobe, *LG* lingual gyrus, *ACC* anterior cingulate cortex, *SMA* supplementary motor area, *PCC* posterior cingulate cortex, *PFC* prefrontal cortex, *STG* superior temporal gyrus, *MOG* middle occipital gyrus

based on negative weights depicting higher activity in HC compared to MDD (Table 3 and Fig. 1). The contribution of the extracted ROIs based on average cluster weights is shown in Supplementary Figure S1.

The COPE-based ANOVA in the left DLPFC showed a main effect of group that was driven by stronger activation in MDD patients compared to HC ($F_{1,116} = 9.38, p < 0.01$). This finding was confirmed by the ANOVA based on the RPE, which showed a main effect of group ($F_{1,116} = 12.32, p < 0.001$), condition ($F_{3,116} = 44.121, p < 0.001$), and an interaction effect ($F_{3,116} = 7.49, p < 0.001$). Paired comparisons showed that effects were driven by stronger differences (MDD > HC) in the three WM conditions compared to the fixation condition (Fig. 2a, b).

The ANOVA based on the COPE in the dorsal ACC (dACC) ROI revealed a main effect of group driven by stronger activation in MDD patients compared to HC ($F_{1,116} = 8.77, p < 0.01$). This finding was confirmed by the ANOVA based on the RPE, which showed a main effect of group ($F_{1,116} = 10.11, p < 0.01$), condition ($F_{3,116} = 123.69, p < 0.001$), and an interaction effect ($F_{3,116} = 6.18, p < 0.001$). Paired comparisons showed that effects were driven by stronger differences (MDD > HC) in the three WM conditions compared to the fixation condition (Fig. 2c, d).

In the posterior cingulate cortex (PCC) the ANOVA based on the COPE revealed a marginally significant main effect of group that was driven by stronger deactivation in patients compared to controls ($F_{1,116} = 5.35, p < 0.05$). Notably, this finding was not confirmed by the ANOVA based on RPE. Here a main effect of condition ($F_{3,116} = 144.47, p < 0.001$), and an interaction effect ($F_{1,116} = 4.31, p < 0.01$) were observed. Post hoc comparisons showed that effects were driven by stronger PCC activity in patients during the fixation condition (Fig. 2e, f).

In the right inferior parietal lobe (IPL), the COPE-based ANOVA showed a main effect of group ($F_{1,116} = 9.95, p < 0.01$) that was driven by relative deactivation in patients and activation in controls. The ANOVA based on RPE showed a significant interaction effect ($F_{3,116} = 6.25, p < 0.001$). Although paired comparison did not reveal any significant differences, descriptive

statistics suggest larger differences in the fixation condition compared to the WM conditions (Fig. 2g, h).

In the left STG/insula the COPE-based ANOVA showed a main effect of group ($F_{1,116} = 8.76, p < 0.01$) that was driven by relative deactivation in patients and activation in controls. The ANOVA based on RPE showed a significant interaction effect ($F_{1,116} = 5.51, p < 0.01$). Post hoc comparisons showed that patients had marginally higher activity during the fixation condition, while no difference was observed during the WM conditions (Fig. 2i, j).

Using Pearson's correlation coefficient a significant negative correlation was found between dACC activation and reaction times for negative stimuli ($r_{118} = -0.22, p = 0.019$). Furthermore, significant positive correlations were found between rumination (reflection subscale) and activation of the right IPL ($r_{46} = 0.4, p = 0.006$) and dACC ($r_{46} = 0.33, p = 0.027$).

DISCUSSION

Based on multivariate pattern classification, this study demonstrated that depressed patients can be distinguished from healthy controls with good classification accuracy and sensitivity based on functional activation patterns during an emotional WM task. In particular, our prediction results outperform both a majority class and random chance prediction. Thus, our results are in line with a recent meta-analysis showing that MDD patients can be distinguished from healthy controls using different magnetic resonance imaging-based modalities [33]. Moreover, the majority of functionally aberrant regions with discriminative power were located in the DMN, regions involved in cognitive control, and the DLPFC. Highest classification accuracies were achieved for neutral and negative stimuli. However, differential activations between neutral and emotional stimuli did not reveal significant classification results.

In left DLPFC, MDD patients showed higher BOLD responses during task conditions compared to HC. This finding is in accordance with results of a recently published meta-analysis of WM-related brain activity in MDD, which reported hyperactivation exclusively in the left DLPFC [12]. Importantly, even though this meta-analysis included data from 13 different WM experiments and therefore a wide range of stimuli, left DLPFC hyperactivation in MDD remained significant when task performance for predominantly verbal WM demands was matched to that of HC. Since our data do not indicate impaired task performance in MDD, our finding of left DLPFC hyperactivation would rather support the hypothesis that hyperfrontality during WM tasks reflects the need for greater activation to maintain a similar level of performance as HC [6–8]. With regard to cognition–emotion interaction, the DLPFC plays a major role in executive control processes, i.e., directing attention away from task-irrelevant emotional distractors during WM [34], which is a key sub-process implicated in effortful voluntary emotion regulation [35]. Greater recruitment of the DLPFC might therefore reflect increased resources to perform the WM task, while inhibiting the allocation of attention toward the processing of emotional stimuli. Interestingly, our data revealed that only left, but not right DLPFC was hyperactive in MDD patients during the WM task. However, the WM task utilized in our study included only verbal stimuli, which are predominantly processed in the left hemisphere [2]. While numerous behavioral studies have shown biases toward negative emotional stimuli in MDD [19, 36] and accordingly reduced activity in left DLPFC for positive stimuli [37], others reported no differences in task performance and functional activity [6, 7]. The MVPC and ROI results presented here also argue against valence-specific effects with regard to left DLPFC activation. Rather, hyperfrontality in MDD might occur to compensate for a lack of deactivation in regions of the DMN [9, 22], in order to keep an effective functional loop between the respective regions and to maintain behavioral performance. Consistently, our data show stronger PCC activation

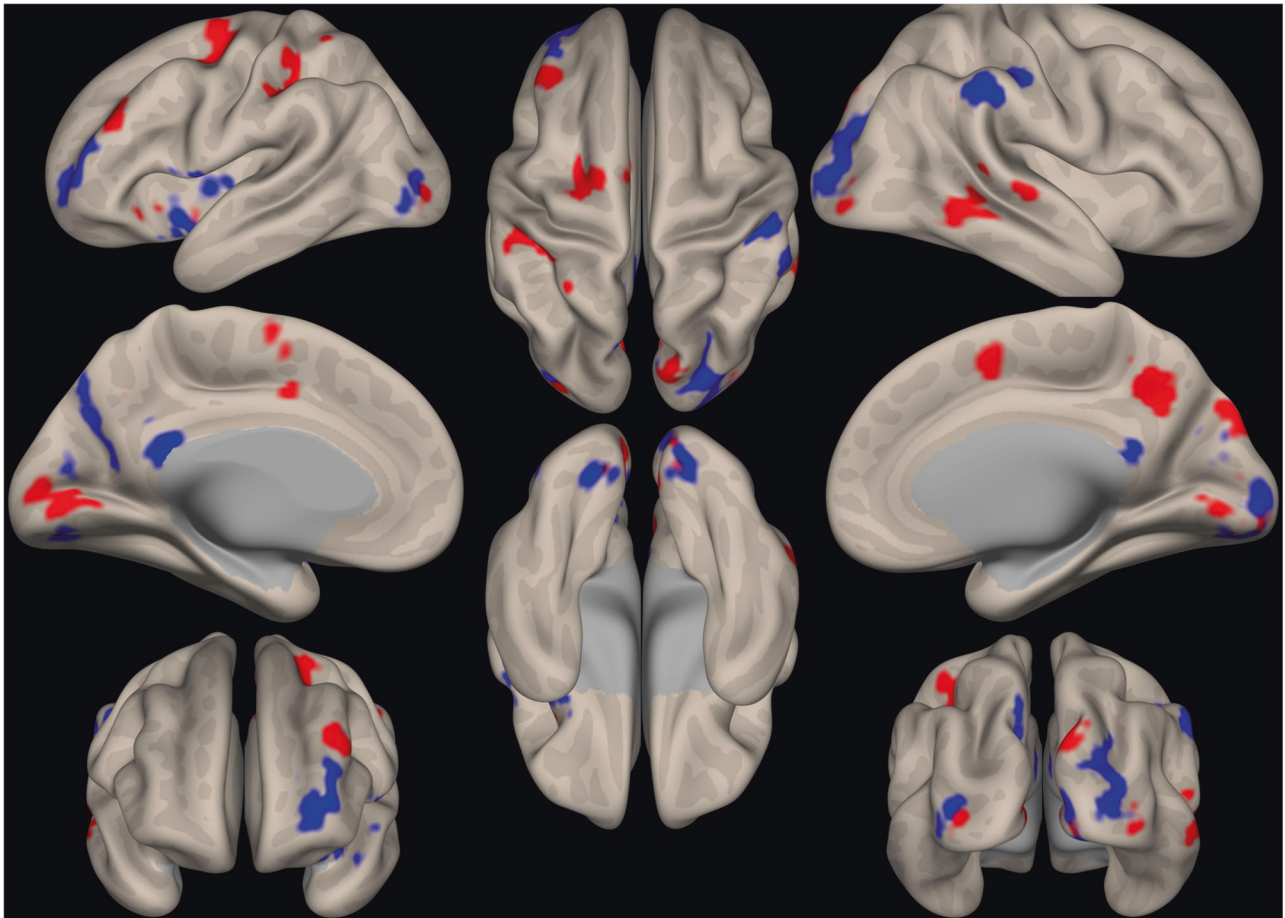


Fig. 1 SVM weight map. The location of the most relevant SVM classification weights from the WM > fixation contrast are shown (20% of the highest weights with a cluster threshold of 50 voxels). Red regions depict more activation in MDD patients. Blue regions depict more activation in healthy controls

in MDD during rest and diminished deactivation during task conditions. PCC is an important hub in the DMN and characterized by increased activity at rest and deactivations during various emotional–cognitive tasks [9, 38]. It might also play a direct role in regulating the focus of attention by controlling the balance between internally and externally focused thought [39]. In the healthy brain, a failure of appropriate deactivation is associated with inefficient cognitive function [40]. It has been suggested that a failure to suppress PCC activity might reflect the intrusion of internal mentation into task performance [41]. Accordingly, increased PCC activation at rest and decreased deactivation during emotional and cognitive tasks have been reported in MDD [9, 42] and might indicate a generally limited potential for adaptive adjustment of this region in MDD patients [43]. It has to be noted that based on the COPE (WM > fixation) patients showed stronger deactivation in the PCC. This result was reversed when looking at the RPE. This apparent contradiction is based on large resting state PCC activation in patients that leads to relatively stronger deactivation during the task conditions.

MDD patients showed higher BOLD responses than HC during the WM task in dACC extending to the supplementary motor area (SMA). The cognitive subdivision of dACC shows strong connections with DLPFC regions, SMA, and parietal cortex, and has been implicated in response selection and processing of cognitively demanding information. Activity in this region during WM tasks is often described in relation to increased effort, complexity, or attention [4, 5]. Higher activation in MDD patients regardless of valence as observed here might assure intact WM performance.

On the other hand, abnormal dACC functioning has also been associated with biased attention to negative stimuli and rumination [44]. Accordingly, our association analysis revealed a significant correlation between rumination and dACC activation.

In right IPL, MDD patients showed higher activation at rest, but diminished BOLD responses during task conditions compared to HC. This region is relevant for visual-spatial processing, usually recruited during *n*-back tasks and an intermediate node between cognitive control and default-mode networks [45]. Decreased BOLD responses in IPL may be reflective of inadequate communication between these networks, such that larger areas of local cortex need to be recruited in order to shift internal resources from internal (i.e., DMN-related) to external (i.e., cognitive control) functions during WM [46]. While results showed decreased BOLD responses regardless of emotional valence, our data nevertheless suggest that these might be associated with rumination and result in higher reaction times for negative and neutral stimuli, possibly because rumination disrupts allocation of cognitive resources and increases recall of negative life events [47].

A cluster differentiating MDD from HC extended from STG to anterior insula, which is viewed as an interface of cognitive, affective, and homeostatic mechanisms, and is suggested to represent an integral structure for stimulus-driven processing and monitoring of the internal milieu [48]. Previous work by Gu et al. [49] suggested that anterior insula is incorporated in a network integrating cognition and emotion [15]. Within this network, anterior insula represents interoceptive changes of unique relevance to subjective experience, whereas control regions, such

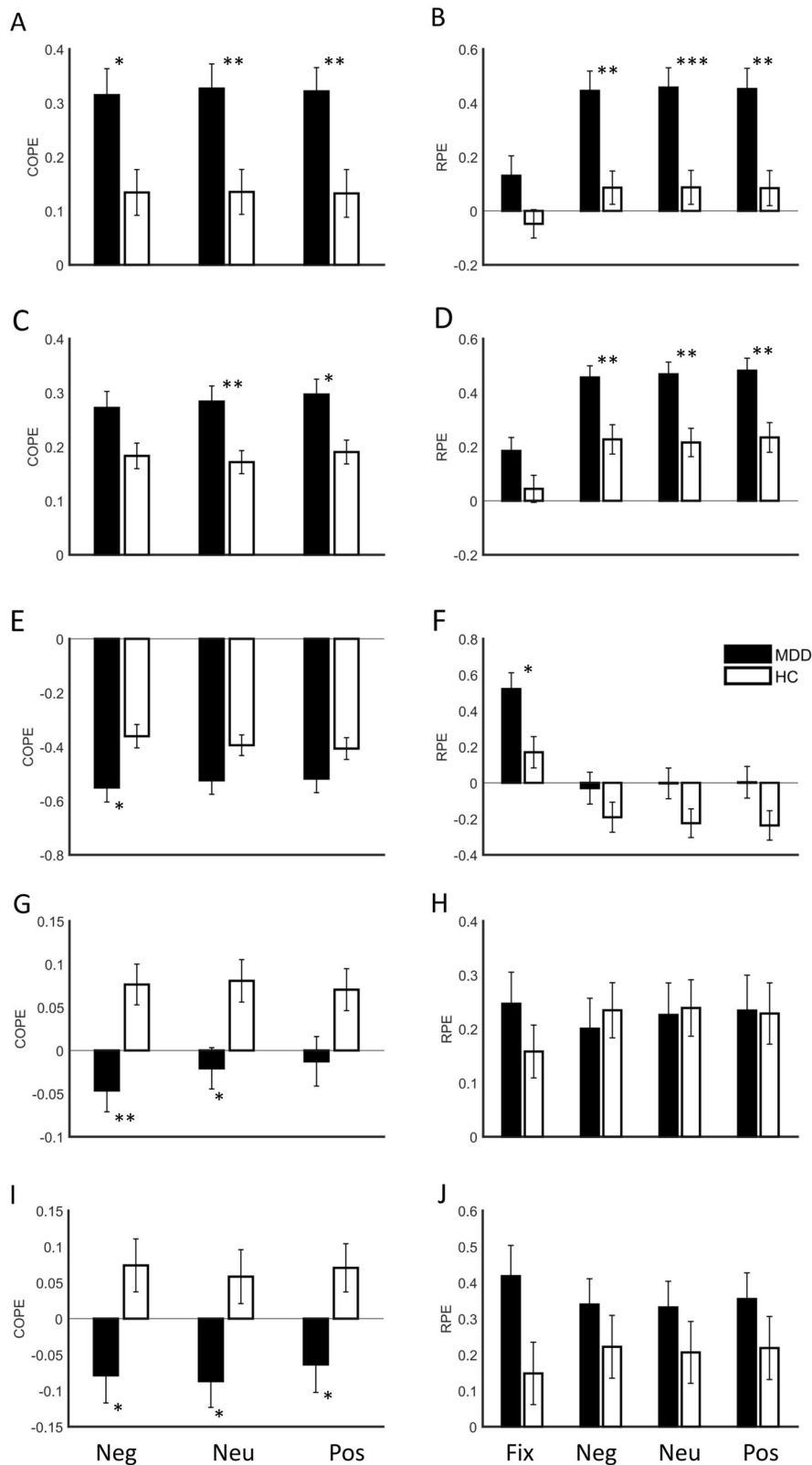


Fig. 2 ROI results. **a, b** Results from the left DLPFC ROI. **c, d** Results from the dorsal ACC ROI. **e, f** Results from the PCC ROI. **g, h** Results from the right IPL ROI. **i, j** Results from the left STG/insula ROI. The left column shows the results for the contrast of parameter estimates. The right column shows the results for the raw parameter estimates. Asterisks depict significant differences (*t*-statistic) between MDD patients and healthy controls (HC): ****p* < 0.001; ***p* < 0.01; **p* < 0.05

as DLPFC, maintain online representations of cognitive demand and stimulus features as well as goal-directed implementation [50], all of which are operations required in cognition–emotion integration. When comparing rest and task within the MDD group, our data show higher activation during rest than during task, which is consistent with previous findings [6, 12, 43] and might indicate increased interoceptive awareness or salience of internal stimuli, while salience of external stimuli is diminished, thereby impairing cognitive processing. This idea is supported by previous findings by Delaveau et al. [51], showing that symptom reduction induced by antidepressant medication increases insula activation during task-related conditions.

Our findings regarding the effect of emotional content on WM performance in HC are in accordance with previous results from several studies that found no impact of emotion on WM performance [24, 25, 52]. In MDD patients, however, we found slower reaction times for negative stimuli, while WM accuracy did not differ between MDD patients and HC. One could hypothesize that reaction times may be more sensitive to small modulations by emotional content than accuracy and therefore emotional content may be more likely to modulate the efficiency with which information is processed as compared with the accuracy with which it is held online. The recruitment of neural networks implicated in emotion processing might result in additional inputs to the WM system [53]. Therefore, it may be that many additional facets of information must be inhibited to allow for processing of only the task-relevant information in the context of the WM task. This increased demand on inhibition may slow the response times in MDD patients, which would be especially true for negative stimuli. The failure of MDD patients to inhibit or discard mood-congruent negative information might increase rumination, and thereby underlie cognitive slowness and attentional deficits [54]. This is also supported by our findings of increased reaction times for negative stimuli in MDD patients and the association between rumination and WM accuracy.

Although the present study overcomes some of the crucial shortcomings of previous reports with respect to sample size and applied statistics, some limitations should be acknowledged. Probably, the most important limitation is that MDD is a very heterogeneous disease and different subtypes might result in different effects on cognitive processes. Our misclassification rate of ~30% may be due in part to this heterogeneity. Further research may build on the current classification results to investigate disease subtypes and relevance to treatment–response prediction. Furthermore, some of the MDD patients ($N = 22$) took different types of antidepressant medication during the study, which might have posed an additional source of variance. Although the applied MVPC approach inherently takes into account confounding factors and noise in the data, it has to be noted that this study did not focus on the effects of antidepressant medication. We performed control analyses (data not shown) between medicated and unmedicated patients that did not reveal significant group differences, but medication types were considered too diverse to draw definite conclusions from these analyses. Upcoming studies should investigate the effects of specific antidepressants on brain activations during WM. From a methodological point of view, it should be noted that the classification model was evaluated on a single dataset only. Therefore dataset-specific effects might have influenced the results. The validation of the current findings in a different sample is the next step, which is considered as future work.

To conclude, by applying MVPC, the present study demonstrates that functional activation patterns during an emotional WM task can be used to distinguish MDD patients from controls with good accuracy and sensitivity. While adequate WM performance in MDD is associated with frontal hyperactivation, patients show a lack of deactivation in regions of the DMN. This effect is most pronounced for negative and neutral stimuli and

associated with rumination, suggesting an important role of aberrations in WM processing for cognition–emotion interactions in MDD.

ACKNOWLEDGEMENTS

We thank the European Commission (grant number H2020-634541) and the German Research Foundation (grant number GR 4510/2-1) for funding this research. Furthermore, we thank our participants for giving time to take part in this study.

ADDITIONAL INFORMATION

The online version of this article (<https://doi.org/10.1038/s41386-018-0081-1>) contains supplementary material, which is available to authorized users.

Competing interests: KB is co-founder and co-owner of Computomics GmbH. The remaining authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Millan MJ, Agid Y, Brüne M, Bullmore ET, Carter CS, Clayton NS, et al. Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. *Nat Rev Drug Discov.* 2012;11:141–68.
2. Wager TD, Smith EE. Neuroimaging studies of working memory: a meta-analysis. *Cogn Affect Behav Neurosci.* 2003;3:255–74.
3. Glahn DC, Curran JE, Winkler AM, Carless MA, Kent JW, Charlesworth JC, et al. High dimensional endophenotype ranking in the search for major depression risk genes. *Biol Psychiatry.* 2012;71:6–14.
4. Duncan J, Owen AM. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* 2000;23:475–83.
5. Botvinick MM, Cohen JD, Carter CS. Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn Sci.* 2004;8:539–46.
6. Fitzgerald PB, Srithiran A, Benitez J, Daskalakis ZZ, Oxley TJ, Kulkarni J, et al. An fMRI study of prefrontal brain activation during multiple tasks in patients with major depressive disorder. *Human Brain Mapp.* 2008;29:490–501.
7. Harvey P-O, Fossati P, Pochon J-B, Levy R, LeBastard G, Lehericy S, et al. Cognitive control and brain resources in major depression: an fMRI study using the n-back task. *Neuroimage.* 2005;26:860–9.
8. Matsuo K, Glahn D, A M Peluso M, Hatch JP, Monkul ES, Najt P, et al. Prefrontal hyperactivation during working memory task in untreated individuals with major depressive disorder. *Mol Psychiatry.* 2007;12:158–66.
9. Grimm S, Boesiger P, Beck J, Schuepbach D, Bermühl F, Walter M, et al. Altered negative BOLD responses in the default-mode network during emotion processing in depressed subjects. *Neuropsychopharmacology.* 2009;34:932–43.
10. Pu S, Yamada T, Yokoyama K, Matsumura H, Kobayashi H, Sasaki N, et al. A multi-channel near-infrared spectroscopy study of prefrontal cortex activation during working memory task in major depressive disorder. *Neurosci Res.* 2011;70:91–7.
11. Hugdahl K, Rund BR, Lund A, Asbjørnsen A, Egeland J, Erland L, et al. Brain activation measured with fMRI during a mental arithmetic task in schizophrenia and major depression. *Am J Psychiatry.* 2004;161:286–93.
12. Wang XL, Du MY, Chen TL, Chen ZQ, Huang XQ, Luo Y, et al. Neural correlates during working memory processing in major depressive disorder. *Prog Neuropsychopharmacol Biol Psychiatry.* 2015;56:101–8.
13. Müller VI, Cieslik EC, Serbanescu I, Laird AR, Fox PT, Eickhoff SB. Altered brain activity in unipolar depression revisited: meta-analyses of neuroimaging studies. *JAMA Psychiatry.* 2017;74:47–55.
14. Aoki R, Sato H, Katura T, Utsugi K, Koizumi H, Matsuda R, et al. Relationship of negative mood with prefrontal cortex activity during working memory tasks: an optical topography study. *Neurosci Res.* 2011;70:189–96.
15. Gray JR, Braver TS, Raichle ME. Integration of emotion and cognition in the lateral prefrontal cortex. *Proc Natl Acad Sci USA.* 2002;99:4115–20.
16. Siemer M. Mood-congruent cognitions constitute mood experience. *Emotion.* 2005;5:296–308.
17. Mathews A, MacLeod C. Cognitive vulnerability to emotional disorders. *Annu Rev Clin Psychol.* 2005;1:167–95.
18. Rusting CL. Personality, mood, and cognitive processing of emotional information: three conceptual frameworks. *Psychol Bull.* 1998;124:165–96.
19. Gotlib IH, Krasnoperova E, Yue DN, Joormann J. Attentional biases for negative interpersonal stimuli in clinical depression. *J Abnorm Psychol.* 2004;113:121–35.
20. Joormann J, Tanovic E. Cognitive vulnerability to depression: examining cognitive control and emotion regulation. *Curr Opin Psychol.* 2015;4:86–92.

21. Hamilton JP, Furman DJ, Chang C, Thomason ME, Dennis E, Gotlib IH. Default-mode and task-positive network activity in major depressive disorder: implications for adaptive and maladaptive rumination. *Biol Psychiatry*. 2011;70:327–33.
22. Rose EJ, Simonotto E, Ebmeier KP. Limbic over-activity in depression during preserved performance on the n-back task. *Neuroimage*. 2006;29:203–15.
23. Vö MLH, Conrad M, Kuchinke L, Urton K, Hofmann MJ, Jacobs AM. The Berlin Affective Word List Reloaded (BAWL-R). *Behav Res Methods*. 2009;41:534–8.
24. Grimm S, Weigand A, Kazzer P, Jacobs AM, Bajbouj M. Neural mechanisms underlying the integration of emotion and working memory. *Neuroimage*. 2012;61:1188–94.
25. Scheidegger M, Henning A, Walter M, Boeker H, Weigand A, Seifritz E, et al. Effects of ketamine on cognition–emotion interaction in the brain. *Neuroimage*. 2016;124:8–15.
26. Beck AT, Steer RA, Brown GK. BDI-II, Beck depression inventory: manual. San Antonio, TX, Boston: Psychological Corp., Harcourt Brace; 1996.
27. Treynor W, Gonzalez R, Nolen-Hoeksema S. Rumination reconsidered: a psychometric analysis. *Cogn Ther Res*. 2003;27:247–59.
28. Patel MJ, Khalaf A, Aizenstein HJ. Studying depression using imaging and machine learning methods. *NeuroImage Clin*. 2016;10:115–23.
29. Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. *Mach Learn*. 2002;46:389–422.
30. Chen Y-W, Lin C-J. Combining SVMs with various feature selection strategies. Feature extraction.. Berlin, Heidelberg: Springer; 2006. p. 315–24.
31. Ojala M, Garriga GC. Permutation tests for studying classifier performance. *J Mach Learn Res*. 2010;11:1833–63.
32. Broyd SJ, Demanuele C, Debener S, Helps SK, James CJ, Sonuga-Barke EJS. Default-mode brain dysfunction in mental disorders: a systematic review. *Neurosci Biobehav Rev*. 2009;33:279–96.
33. Kambeitz J, Cabral C, Sacchet MD, Gotlib IH, Zahn R, Serpa MH, et al. Detecting neuroimaging biomarkers for depression: a meta-analysis of multivariate pattern recognition studies. *Biol Psychiatry*. 2017;82:330–8.
34. Dolcos F, McCarthy G. Brain systems mediating cognitive interference by emotional distraction. *J Neurosci*. 2006;26:2072–9.
35. Phillips ML, Ladouceur CD, Drevets WC. A neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. *Mol Psychiatry*. 2008;13:829–57.
36. Gilboa-Schechtman E, Ben-Artzi E, Jeczemien P, Marom S, Hermesh H. Depression impairs the ability to ignore the emotional aspects of facial expressions: evidence from the Garner task. *Cogn Emot*. 2004;18:209–31.
37. Disner SG, Beevers CG, Haigh EAP, Beck AT. Neural mechanisms of the cognitive model of depression. *Nat Rev Neurosci*. 2011;12:467–77.
38. Fox MD, Snyder AZ, Vincent JL, Corbetta M, Essen DCV, Raichle ME. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA*. 2005;102:9673–8.
39. Leech R, Kamourieh S, Beckmann CF, Sharp DJ. Fractionating the default mode network: distinct contributions of the ventral and dorsal posterior cingulate cortex to cognitive control. *J Neurosci*. 2011;31:3217–24.
40. Bonnafe V, Leech R, Kinnunen KM, Ham TE, Beckmann CF, Boissezon XD, et al. Default mode network connectivity predicts sustained attention deficits after traumatic brain injury. *J Neurosci*. 2011;31:13442–51.
41. Weissman DH, Roberts KC, Visscher KM, Woldorff MG. The neural bases of momentary lapses in attention. *Nat Neurosci*. 2006;9:971–8.
42. Grimm S, Ernst J, Boesiger P, Schuepbach D, Boeker H, Northoff G. Reduced negative BOLD responses in the default-mode network and increased self-focus in depression. *World J Biol Psychiatry*. 2011;12:627–37.
43. Diener C, Kuehner C, Brusniak W, Ubl B, Wessa M, Flor H. A meta-analysis of neurofunctional imaging studies of emotion and cognition in major depression. *Neuroimage*. 2012;61:677–85.
44. Pizzagalli DA. Frontocingulate dysfunction in depression: toward biomarkers of treatment response. *Neuropsychopharmacology*. 2011;36:183–206.
45. Braga RM, Sharp DJ, Leeson C, Wise RJS, Leech R. Echoes of the brain within default mode, association, and heteromodal cortices. *J Neurosci*. 2013;33:14031–9.
46. Philip NS, Sweet LH, Tyrka AR, Carpenter SL, Albright SE, Price LH, et al. Exposure to childhood trauma is associated with altered n-back activation and performance in healthy adults: implications for a commonly used working memory task. *Brain Imaging Behav*. 2016;10:124–35.
47. Nolan SA, Roberts JE, Gotlib IH. Neuroticism and ruminative response style as predictors of change in depressive symptomatology. *Cogn Ther Res*. 1998;22:445–55.
48. Craig AD. Emotional moments across time: a possible neural basis for time perception in the anterior insula. *Philos Trans R Soc Lond B Biol Sci*. 2009;364:1933–42.
49. Gu X, Hof PR, Friston KJ, Fan J. Anterior insular cortex and emotional awareness. *J Comp Neurol*. 2013;521:3371–88.
50. Dosenbach NUF, Fair DA, Cohen AL, Schlaggar BL, Petersen SE. A dual-networks architecture of top-down control. *Trends Cogn Sci*. 2008;12:99–105.
51. Delaveau P, Jabourian M, Lemogne C, Guionnet S, Bergouignan L, Fossati P. Brain effects of antidepressants in major depression: a meta-analysis of emotional processing studies. *J Affect Disord*. 2011;130:66–74.
52. Döhl K, Sommer M, Ibach B, Rothmayr C, Meinhardt J, Hajak G. Neural correlates of emotional working memory in patients with mild cognitive impairment. *Neuropsychologia*. 2008;46:37–48.
53. LeDoux JE. Emotion circuits in the brain. *Annu Rev Neurosci*. 2000;23:155–84.
54. Lau MA, Christensen BK, Hawley LL, Gemar MS, Segal ZV. Inhibitory deficits for negative information in persons with major depressive disorder. *Psychol Med*. 2007;37:1249–59.