BRAIN
A JOURNAL OF NEUROLOGY

Shared dimensions of performance and activation dysfunction in cognitive control in females with mood disorders

Kelly A. Ryan, Erica L. Dawson,[‡] Michelle T. Kassel, Anne L. Weldon, David F. Marshall, Kortni K. Meyers, Laura B. Gabriel, Aaron C. Vederman,[δ] Sara L. Weisenbach,[†] Melvin G. McInnis, Jon-Kar Zubieta and Scott A. Langenecker[‡]

Major depressive disorder and bipolar disorder share symptoms that may reflect core mood disorder features. This has led to the pursuit of intermediate phenotypes and a dimensional approach to understand neurobiological disruptions in mood disorders. Executive dysfunction, including cognitive control, may represent a promising intermediate phenotype across major depressive disorder and bipolar disorder. This study examined dimensions of cognitive control in women with major depressive disorder or bipolar disorder in comparison to healthy control subjects using two separate, consecutive experiments. For Experiment 1, participants completed a behavioural cognitive control task (healthy controls = 150, major depressive disorder = 260, bipolar disorder = 202; age range 17–84 years). A sample of those participants (healthy controls = 17, major depressive disorder = 19, and bipolar disorder = 16) completed a similar cognitive control task in an event-related design functional magnetic resonance imaging protocol for Experiment 2. Results for Experiment 1 showed greater impairments on the cognitive control task in patients with mood disorders relative to healthy controls ($P < 0.001$), with more of those in the mood disorder group falling into the 'impaired' range when using clinical cut-offs ($<$5th percentile). Experiment 2 revealed only a few areas of shared activation differences in mood disorder greater than healthy controls. Activation analyses using performance as a regressor, irrespective of diagnosis, revealed within and extra-network areas that were more active in poor performers. In summary, performance and activation during cognitive control tasks may represent an intermediate phenotype for mood disorders. However, cognitive control dysfunction is not uniform across women with mood disorders, and activation is linked to performance more so than disease. These findings support subtype and dimensional approaches to understanding risk and expression of mood disorders and are a promising area of inquiry, in line with the Research Domain Criteria initiative of NIMH.

University of Michigan Medical School, Department of Psychiatry, 2101 Commonwealth Blvd, Suite C, Ann Arbor, MI 48105, USA
‡ Present address: Ohio State University, Department of Psychiatry, Columbus, OH, USA
δ Present address: Now in private practice in Portland, OR, USA
† Present address: University of Illinois at Chicago, Department of Psychiatry, Chicago, IL, USA

Correspondence to: Kelly A Ryan,
Ph.D., Department of Psychiatry,
University of Michigan,
2101 Commonwealth Blvd,
Suite C, Ann Arbor,
MI 49105
E-mail: karyan@umich.edu

Introduction

Using Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-V) diagnostic criteria, mood disorders form discrete categories despite heterogeneous and sometimes overlapping symptoms. Much of the resulting research in mood disorders uses this conventional clinical approach. A parallel and potentially complimentary approach to the classification model of the DSM-V is the dimensional approach promulgated by the NIMH Research Domain Criteria (RDoC), which highlights the potential utility of dimensional approaches to better reflect the underlying neurobiological disruptions in illnesses (e.g. cognitive control dysfunction evident in mood disorders). Indeed, there are many shared symptoms and diagnostic features in mood disorders reflecting the potential for a core neurobiological intermediate phenotype. Some intermediate phenotypes may be shared whereas other intermediate phenotypes may distinguish between these disorders.

Diagnostic categories have been the basis for neuroimaging and genome-wide association studies (Garvey et al., 1986; Clayton, 1990; Moore and Williams, 2009; Savitz and Drevets, 2009), but disease-specific risk factors identified by extant investigations have been equivocal based on the current categorical system of mood disorder diagnoses (Dolan et al., 1993; Sheline et al., 1999; Ravnkilde et al., 2003; Lange et al., 2004; Drevets, 2007; Hamilton et al., 2008; Surguladze et al., 2008). This has led to the pursuit of intermediate phenotypes, sometimes called endophenotypes. Intermediate phenotypes are defined as 'mediating neural mechanisms that bridge the biological gap from DNA sequence to pathological behaviour' (Meyer-Lindenberg and Weinberger, 2006). An intermediate phenotype can represent one of possibly many risk characteristics for development of a mood disorder. The value of identifying intermediate phenotypes lies in the specificity of biological substrates predicting intermediate phenotypes and the enhanced opportunity to identify individuals at risk for developing a specific disorder based on positive intermediate phenotype status.

Among the clinical commonalities between patients with major depressive disorder (MDD) and bipolar disorder (BD) (see Cuellar et al., 2005 for a review; Hirschfeld et al., 2003; Goodwin, 2007), cognitive impairments have been identified in both populations relative to healthy comparisons, particularly in the areas of cognitive flexibility, processing speed, and divided attention (Kerr et al., 2005; Langenecker et al., 2007a; Godard et al., 2011; Ryan et al., 2012). However, performances on measures of sustained or divided attention are not universally equivalent between psychiatric groups. For example, Maalouf and colleagues (2010) reported poorer sustained attention performance among adults with BD type I compared to those with MDD, whereas Taylor Tavares and colleagues (2007) reported poorer performance among adults with MDD than those with BD type II and another showed that those with 'bipolar spectrum disorder' mildly under-performed compared to adults with MDD on a sustained attention measure (Smith et al., 2006). Overall, disruptions in executive functioning, particularly for areas like cognitive control, may confer risk for mood disorders, but many previous studies are limited by small sample size, ill-defined mood disorder samples, or atypically-homogeneous samples that do not accurately reflect the overall population of adults with these disorders.

Focal and circuit-based neurobiological abnormalities are inconsistently documented in both MDD and BD variability exists in structural, functional, and connectivity results across studies. Convergent findings indicate abnormally elevated subcortical activity and reduced prefrontal cortical activity during emotion processing paradigms in both MDD and BD when compared to healthy controls (Phillips et al., 2008; Almeida and Phillips, 2013). Previous research findings suggest that neuropsychological deficits exhibited by subjects with mood disorder relate to abnormalities in the prefrontal cortex and other functionally and structurally connected regions, which negatively impact emotional and behavioural regulation (Beblo et al., 2011). However, few neuroimaging studies have directly compared patterns of abnormal function in BD and MDD (Taylor Tavares et al., 2008; Almeida et al., 2009, 2010; Bertocci et al., 2012). For example, Bertocci and colleagues (2012) showed that individuals with unipolar depression had elevated dorsal anterior mid-cingulate cortical activity compared to the BD and healthy control (HC) groups, suggesting abnormal recruitment of attentional control. Putamen activation, on the other hand, was higher in both patient groups relative to healthy controls in the context of equivalent performance, which may represent a task-independent, shared disease biomarker. With small samples inherent in most imaging studies, it is often challenging to estimate the true significance of any performance or activation differences between groups. Larger imaging studies, or imaging studies embedded within larger behavioural studies, may be better able to dissociate disease-specific from performance-variant parameters, and those parameters that might be driven by both disease and performance. This may be particularly true when neural circuits are disrupted and result in diminished cognitive control, a potential shared risk among individuals with different mood disorders.

In the context of the NIMH RDoC initiative, the present study examines shared and specific dimensions of sustained attention across two mood disorders groups, MDD and BD, using two separate experiments. The study focuses on females, to limit sex-specific heterogeneity in performance and functional abnormalities. For Experiment 1, which included a demographically equivalent sample of females who were tested by our group in the last decade, we hypothesized that women with MDD and BD would perform worse than healthy counterparts on a behavioural measure of cognitive control. We further hypothesized that a greater portion of women with mood disorders would perform in the impaired range (using a dimensional, multiple cut-score framework for determining 'impairment') relative to healthy controls. Experiment 2 included a subset of the Experiment 1 sample. We again hypothesized that women with MDD and BD would perform worse than healthy controls on a similar cognitive control task using an event-related functional MRI protocol. Additionally, we predicted that women with MDD and BD would show disruption of neural circuitry that typically supports sustained attention relative to healthy controls. Lastly, we predict that greater activation within the cognitive control network would be related to better performance irrespective of group.

Materials and methods

Participants

Only females were included in this study to reduce sex-specific heterogeneity in performance and activation. For Experiment 1, sample size included 150 healthy controls (HC), 266 MDD (Major Depressive Disorder), and 202 BD (Bipolar Disorder) type I, type II, and NOS who were recruited from separate research protocols at the University of Michigan. Healthy controls were recruited using flyers and were determined to be free of any past or current psychiatric or neurological disorder, including current substance use disorders. The participants with a mood disorder were recruited for multiple research protocols being conducted through our laboratories or at intake psychiatry appointments at the University of Michigan. For patients recruited from the psychiatry clinic (MDD = 189, BD = 16), diagnosis was obtained through clinic interview using DSM-IV criteria and confirmed by review of medical records. Participants were also recruited from a large longitudinal study of bipolar disorder [Heinz C. Prechter Longitudinal Study of Bipolar Disorder (Langenecker *et al.*, 2010; Ryan *et al.*, 2012); HC = 70, MDD = 4, BD = 186]; diagnoses were determined using the Diagnostic Interview for Genetic Studies (DIGS; Nurnberger *et al.*, 1994). The other participants (HC = 80, MDD = 73) were recruited following completion of other ongoing studies in our laboratory in the Department of Psychiatry at the University of Michigan. These latter participants were diagnosed using the Structured Clinical Interview for the DSM-IV (SCID-IV; First *et al.*, 1995). All MDD participants were screened for current or past mania before entering their

respective studies. Participants with BD were in the euthymic ($n = 99$), depressed ($n = 91$), or mixed state ($n = 12$), but not a manic state, to better equilibrate the impact of symptoms across groups. Participants with MDD were in the depressed ($n = 179$) and euthymic ($n = 64$) states.

Table 1 provides the demographic, clinical data, and results of group comparison analyses for participants in Experiment 1. The participants with BD were older than the healthy controls but similar in age to the participants with MDD. Random removal of younger healthy control participants to minimize group differences did not affect primary findings. There were no group differences for education or estimated IQ. Among the patient groups, the bipolar disorder participants had a younger age of psychiatric illness onset and were more likely to be on medications than the participants with MDD. The MDD participants were more likely to be depressed than those with bipolar disorder. A repeated measures ANOVA indicated no main effect for mood disorder diagnosis, $F(1) = 0.01$, $P = 0.92$, medication, $F(1) = 1.85$, $P = 0.17$, or the interaction term, $F(1) = 0.003$, $P = 0.96$.

We hypothesized in Experiment 1 that there would be a general pattern of disruption in cognitive control systems in MDD and BD. However, several clinical characteristics of the patients (e.g. number of prior episodes, psychosis, and severity) were not uniform in collection across studies. Current depressive symptoms were available for 95% ($n = 587$) of the participants using one or more measures: Patient Health Questionnaire-9 (Kroenke *et al.*, 2001) ($n = 188$), Hamilton Depression Rating Scale (Hamilton, 1967) ($n = 399$), and/or Beck Depression Inventory, 2nd edition (Beck *et al.*, 1988) ($n = 131$); cut-off scores to signify clinical depression were different for each measure so participants were categorized as Depressed/Not Depressed based on each measures' respective cut-off using clinical convention (Table 1). Age of illness onset for any mood/anxiety disorder and presence/absence of psychiatric medication were also collected for each sample, although information beyond this level of data was not uniformly available. Each research project was approved by the Institutional Review Board at the University of Michigan and consent was obtained or waived consistent with the Declaration of Helsinki.

For Experiment 2, 52 females were recruited from ongoing studies to undergo functional MRI (HC = 17, MDD = 19, BD = 16). The neuroimaging sample overlapped with the original sample described above, with the exception of five additional participants (HC = 2, MD = 1, BD = 2) whose Experiment 1 performance on the cognitive control task prior to scanning was lost or corrupted. Neuroimaging occurred after collection of other performance data (Experiment 1 data). For Experiment 2, there were no differences among the three groups in age or education. MDD and BD groups did not differ in first age of onset or depression symptoms (Table 1).

Measures

Experiment 1: Performance outside the scanner

Estimated premorbid verbal IQ was collected using the Synonym Knowledge Task (SKT; based on Shipley, 1946). Participants were presented with a word and then asked to

Table 1 Demographic and clinical data for healthy control, MDD, and BD groups in Experiments 1 and 2

Experiment 1	HC $n = 150$	MDD $n = 266$	BD $n = 202$	F/ X^2	P-value	Post hoc
Age	34.3 (16.1)	36.8 (12.9)	38.3(12.4)	3.86	0.022	BD > HC
Education	15.7 (2.3)	15.5 (2.6)	15.4 (2.1)	0.78	0.461	NS
First age at onset	NA	23.5 (11.9)	16.5 (7.9)	56.99	< 0.001	
Synonym Knowledge[a] (per cent correct)	75.4 (18.0)	75.2 (17.1)	77.5 (12.5)	1.18	0.308	NS
% Depressed[b,c]	NA	73.7	51.5	29.58	< 0.001	
% on Medication[d]	NA	63.3	84.9	29. 56	< 0.001	
Experiment 2	**HC $n = 17$**	**MDD $n = 19$**	**BD $n = 16$**	**F/ X^2**	**P**	**Post hoc**
Age	37.1 (18.9)	44.1 (16.3)	44.1 (11.5)	1.09	0.344	NS
Education	16.0 (1.6)	15.9 (2.3)	15.8 (1.7)	0.07	0.934	NS
Hamilton Depression	NA	15.2 (7.7)	16.8 (5.4)	0.49	0.488	NS

Data are presented as mean (SD).
[a]Estimated IQ using Synonym Knowledge Test, per cent correct.
[b]Chi-Square (Fisher's Exact).
[c]Cut-off for depression: PHQ-9 > 9, HDRS > 9, BDI > 12.
[d]Per cent taking any psychiatric medication.
NA = not available; NS = not significant.

choose one of four additional words that was most similar. There was no penalty for guessing and no time limit for responding.

Cognitive control was assessed using the Parametric Go/No-go Task (PGNG; Langenecker et al., 2005); based upon several works (Garavan et al., 1999; Nielson et al., 2002; Langenecker and Nielson, 2003). The PGNG task (Fig. 1A) is an 11-min test that consists of three separate levels, but only the first level was used for this study. The first level (Go Task) measures attention and response time, resulting in two dependent measures of cognitive control, and takes 3 min to complete. A serial stream of letters is presented with each letter appearing for 500 ms with a 0-ms interstimulus interval. Responses are made by pressing a computer keyboard key (the letter 'n', or index finger in functional MRI) as quickly as possible using the index finger of the preferred hand. Percentage of correct Go target trials (accuracy—measuring sustained attention and set maintenance) is computed by dividing the number of correct target responses by the number of total possible target responses. Response time to Go targets (measuring processing speed in a multiple target search) is the average response time for all correct targets.

Experiment 2: Performance inside the scanner

To assess cognitive control in the scanner, the Go Task of the PGNG was embedded within a list learning task, the Semantic List Learning Task (SLLT; Langenecker et al., 2004). The results from this list learning memory task are outside the purview of the present study. The Go task of the PGNG (functional MRI-PGNG) was used for 14-s blocks as a brief distracter for the memory task, embedded between encoding and silent rehearsal blocks (Fig. 1B). Go Task of the PGNG is described above, except during functional MRI it is split into 14-s blocks that occur at ~67.75 s intervals throughout the SLLT memory encoding task (see Fig. 1 for an illustration of the functional MRI-PGNG design.). There were 15 blocks spread across five runs. Accuracy and response time were recorded.

Functional MRI acquisition

Whole-brain imaging was performed using a GE Signa 3T scanner (release VH3). Functional MRI series consisted of 36 contiguous oblique-axial sections acquired using a forward-reverse spiral sequence, which provides excellent functional MRI sensitivity (Glover and Thomason, 2004). The image matrix was 64×64 over a 24 cm field of view for a $3.75 \times 3.75 \times 4$ mm voxel. The 36-slice volume was acquired serially at 1750 ms temporal resolution for a total of 154 time points for the combined SLLT/PGNG task in each of five runs for a total of 770 time points. Anatomical images were also collected using between 104–124 high-resolution fast SPGR IR axial images (echo time = 3.4 ms, repetition time = 10.5 ms, 27° flip angle, number of excitations = 1, slice thickness = 1–1.5 mm, field of view = 24 cm, matrix size = 256×256) for each participant, and these were used for co-registration and normalization purposes.

Functional MRI processing

Image preprocessing was conducted using SPM8, including realignment, slice timing correction, co-registration, normalization to the MNI world space, and smoothing with a five full-width at half-maximum filter. Contrast images were derived based upon activation for correct events of the functional MRI-PGNG. These were computed by using the blood oxygen level-dependent signal for all correct responses relative to the implicit baseline for each individual in a first level analysis. The SPM8 haemodynamic response function (HRF) model was used to model the blood oxygen level-dependent response in the first level models. Random effects analyses were conducted using whole brain analyses in SPM8.

Statistical analyses

For Experiment 1, a multivariate analysis of variance (MANOVA) was used, with participant group (BD, MDD, and HC) as the independent variable and PGNG accuracy and response time as the dependent variables. *Post hoc*
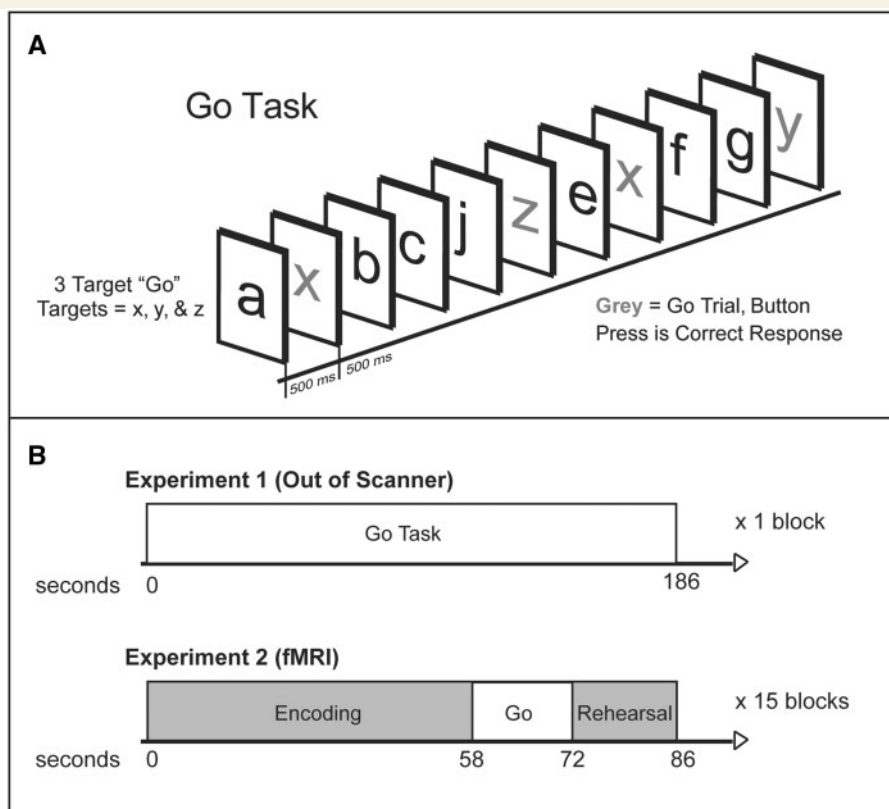
Figure 1 Go Task and Experiment 1 and 2 task presentation. (A) PGNG. The first level (Go Task) was used for this study and measures attention and response times and has two dependent measures of cognitive control. **(B)** Experiments 1 and 2 task presentation. During Experiment 1, only the PGNG was administered. During Experiment 2, the functional MRI protocol included administration of the SLLT and the PGNG was embedded within this task.

ANOVAs were performed to clarify the nature of the associations between group membership and PGNG performance.

A set of planned analyses to evaluate cut scores for dimensional markers were conducted using the distribution of the healthy control group as the comparison group. Use of cut-scores is complementary to a dimensional approach, as it enables an understanding of what extreme scores along a dimension mean in reference to the rest of the population. In practical terms, cut scores would be necessary to define 'illness severity' and determine need for services. Frequency cut-offs are computed at the 5th, 9th, and 16th percentiles to denote impairment, borderline impairment, and low average performance, which is consistent with standard clinical practice in use of dimensional scales. Strategies of this type can be effective in determining how dimensional measures can be used in clinical practice.

For Experiment 2, multi-voxel MANOVAs were compared in SPM 8 with the same independent variable, and voxel-by-voxel activation as the dependent variable, with whole brain AlphaSim correction of $P < 0.05$ using combined height and extent thresholds ($P < 0.005$, $k > 55$). ANOVAs were also used to compare participant group by PGNG performance for participants using both in- and out-of-scanner (performance data from Experiment 1) data. These imaging analyses were conducted to evaluate shared (HC versus all mood disorders) versus specific (MDD versus HC, BD versus HC) aspects of these mood disorders. We also conducted a

dimensional analysis of activation based upon accuracy as a regression onto activation for all events. As the bipolar disorder group was over-represented at the lower end of the performance regression, we carefully evaluated known potential confounds of performance by disease and also factors that are related to course and treatment (e.g. medications, number of manic episodes) that might explain any potential or observed differences between MDD and BD. *Post hoc* data reduction was accomplished with principal axis factor analysis.

Results

Experiment 1

Performance differences between the control group and mood disorder groups: shared dysfunction

When comparing the mood disorder groups (MDD and BD) to healthy controls (Fig. 2A), the patient groups performed worse [$F(4, 1232) = 9.25$, $P < 0.001$ for both accuracy, $F(2, 618) = 4.30$, $P = 0.01$ and response time, $F(2, 622) = 18.83$, $P < 0.001$]. *Post hoc* analyses confirmed that the healthy control group performed better than both mood disorder groups for accuracy (MDD, $P = 0.02$; BD, $P = 0.03$) and response time (MDD, $P < 0.001$ and BD,
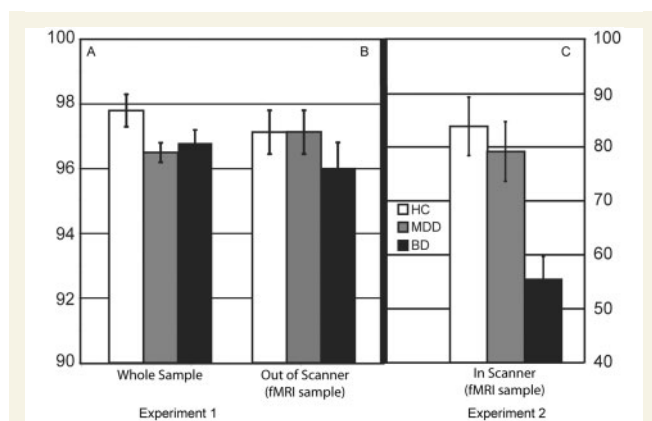
Figure 2 Behavioural performance. (A) Experiment 1 behavioural performance. MDD and BD groups performed significantly worse for accuracy compared to the healthy controls. Accuracy is represented as per cent correct on the Go trial of the PGNG. **(B)** Experiment 2 behavioural performance out of the scanner. There were no significant differences between groups in accuracy performance for the Go Task of the PGNG task. **(C)** Experiment 2 behavioural performance in the scanner. The bipolar disorder group performed worse than the healthy controls and MDD groups for accuracy. Accuracy is defined as per cent correct on the Go trial of the PGNG. fMRI = functional MRI.

Figure 3 Dimensional cut-offs for response time and accuracy. (A) Dimensional cut-offs for response time. There were significantly more patients below the 16th, 9th, and 5th percentile cut-offs relative the healthy control group. **(B)** Dimensional cut-offs for accuracy. There were significantly more patients below the 16th, 9th, and 5th percentile cut-offs relative the healthy control group.

$P < 0.001$). There were no significant differences for accuracy (t = 0.10, $P = 0.92$) or response time (t = −0.10, $P = 0.92$) between the MDD and BD groups. Results were identical after matching groups on age, and also in separate analyses matching patient groups on depression severity.

To examine dimensional markers using the cut-offs for response time (percentages are illustrated graphically in Fig. 3A), there were significantly more patients below the 16th percentile [$\chi^2(1) = 20.13$, $P < 0.001$], the 9th percentile [$\chi^2(1) = 10.39$, $P = 0.006$], and the 5th percentile cut-offs [$\chi^2(1) = 11.23$, $P = 0.004$] relative to the healthy control group. For accuracy (Fig. 3B), there were significantly more patients below the 16th percentile [$\chi^2(1) = 17.03$, $P < 0.001$], the 9th percentile [$\chi^2(1) = 17.07$, $P < 0.001$], and the 5th percentile cut-offs [$\chi^2(1) = 16.82$, $P < 0.001$] relative to the healthy control group.

Experiment 2

Performance and functional MRI results

There were no significant differences between groups in accuracy performance outside the scanner (Experiment 1 PGNG) for Level 1 of the PGNG task, $F(2, 43) = 0.61$, $P = 0.55$ (Fig. 2B) and the means were equivalent to the group means observed in the Experiment 1 sample. Healthy control participants produced shorter response times in comparison to both patient groups, $F(2, 50) = 3.32$, $P < 0.05$, similar to Experiment 1. During the functional MRI-PGNG (in scanner), the bipolar disorder group performed worse than the healthy control and MDD groups for accuracy and response time, $F(2, 50) = 8.18$, $P < 0.01$
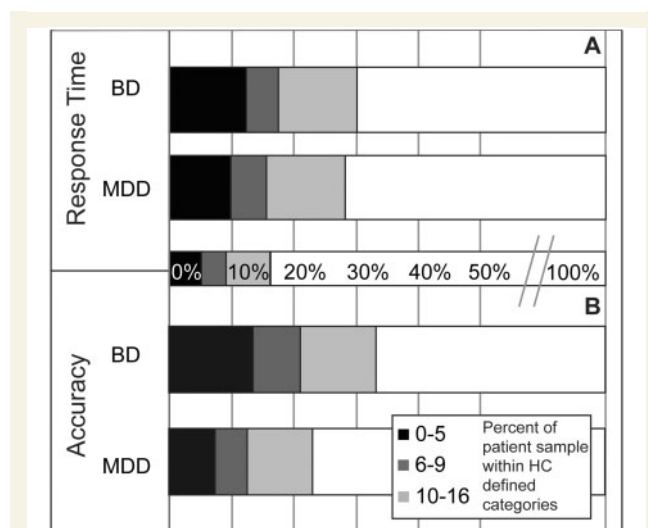
and $F(2, 50) = 6.20$, $P < 0.01$, respectively (Fig. 2C). When comparing this sample's Experiment 1 out-of-scanner performance (Fig. 2B) to Experiment 2 inside-scanner performance (Fig. 2C), each diagnostic group showed significantly slower response time [HC, t = −6.9, $P < 0.001$, mean (M) = 428.3, standard deviation (SD) = 57.1; MDD, t = −9.1, $P < 0.001$, M = 459.4, SD = 49.7; BD, t = −6.9, $P < 0.001$, M = 459.2, SD = 55.4] and better accuracy during Experiment 1 (HC, t = 2.6, $P = 0.02$, M = 0.97, SD = 0.05; MDD, t = 3.2, $P = 0.006$, M = 0.97, SD = 0.04; BD, t = 6.1, $P < 0.001$, M = 0.96, SD = 0.09) as compared to Experiment 2 (response time: HC, M = 499.8, SD = 72.6; MDD, M = 523.3, SD = 53.7, BD, M = 570.8, SD = 59.7; accuracy: HC, M = 0.83, SD = 0.23; MDD, M = 0.78, SD = 0.24; BD, M = 0.57, SD = 0.19). Test retest correlations for response time between out-of-scanner and in-scanner were strong (r = 0.70, $P < 0.001$), but weak for accuracy (r = 0.134, $P = 0.370$), likely due to the greater variability in performance across participants.

Go task activation

Task-specific regions active in all groups included the cognitive control network—bilateral dorsal cingulate, inferior parietal lobule, medial dorsal thalamus, and dorso-lateral prefrontal cortex (Supplementary Fig. 1 and Supplementary Table 1).

Imaging analysis by diagnosis, shared and distinctive illness features

Shared mechanisms of illness were evaluated in a two-way ANOVA using diagnosis (mood disorder versus healthy controls) as the independent variable and event-related activation for correct responses (hits) to Go events as the
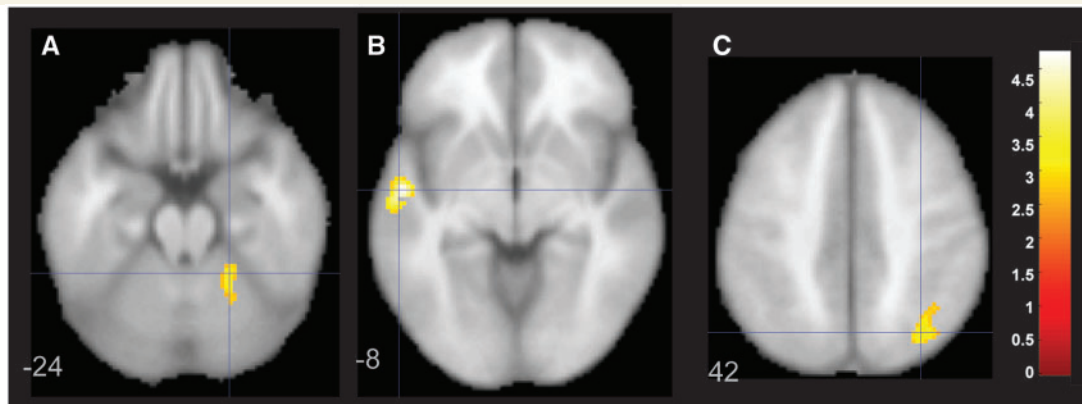
Figure 4 Figure demonstrating statistically greater activation in left superior temporal gyrus (B) and right superior parietal lobule (A) and cerebellum (C) for the bipolar disorder and MDD groups relative to the healthy control group. Panels are labelled at different *z* coordinates and MDD/BD greater than healthy controls.
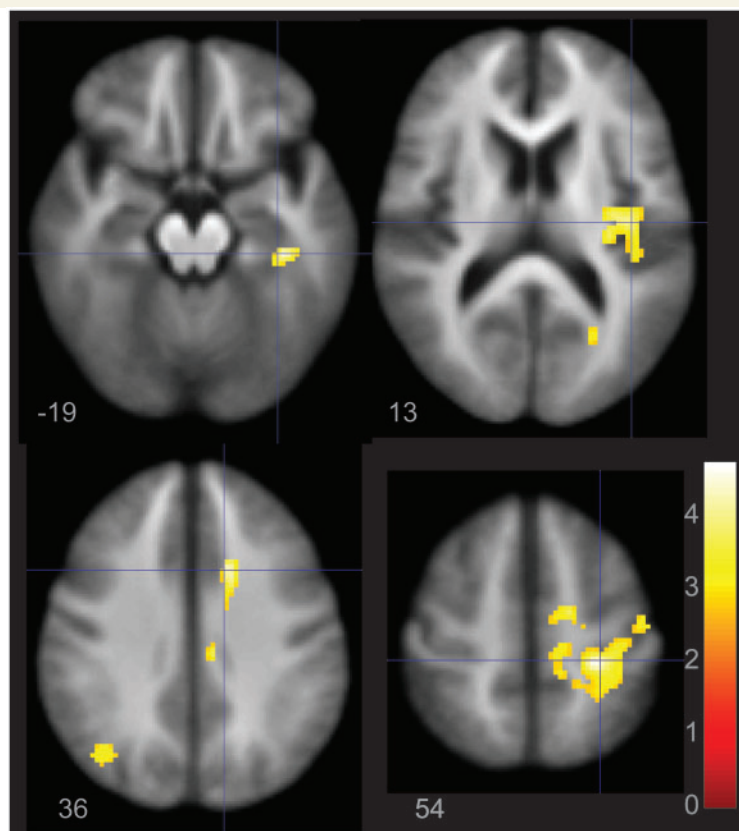

Figure 5 Regions of the brain that are significantly, positively correlated with performance, covarying for group membership.

dependent variable in whole brain analyses. There were no regions with greater activation in the healthy control group relative to the BD and MDD females. The BD and MDD groups had greater activation relative to the healthy control group in left superior temporal gyrus and right superior parietal lobule and cerebellum (Fig. 4).

Distinct mechanisms of illness were evaluated in a comparison between MDD versus HC and BD versus HC. To diminish multiple comparison problems (and type 1 error), the main effect of group was used to define regions of interest, and signals from these group effect regions of interest were extracted for subsequent analyses. There

Table 2 Distinct features ANOVA showing independent effects of BD and MDD illness

Contrast/lobe	BA	x	y	z	Z	mm³	Factor
		Talairach coordinates					
Diagnostic group effect							
Frontal							
Middle frontal	8	32	16	41	4.16	2840	Group 1
Medial frontal	9	15	41	16	3.46	1144	Group 1
Dorsal cingulate*	24	8	−19	39	4.38	10120	Group 3
Dorsal anterior cingulate	32	−14	20	41	3.76	2472	Group 1
Parietal							
Precuneus*	7	10	−62	51	4.11	1064	Group 2
*	7	−15	−63	39	3.2	472	Group 3
*	7	23	−65	30	3.14	640	Group 2
Posterior cingulate*	31	31	−63	16	3.73	1472	Group 1
Temporal							
Middle temporal	21	−38	−13	−8	4.08	1608	Group 3
Insula	13	−42	−7	15	3.72	1056	Group 3
Occipital							
Cuneus*	18	−18	−81	25	4.16	14712	Group 1
	18	13	−93	18	3.31	552	Group 1
Subcortical							
Declive*		13	−58	−14	3.75	600	Group 3

Group 1 = MDD hyper > HC > BD hypo.
Group 2 = BD hyper > MDD > HC hypo.
Group 3 = MDD hyper, HC and BD Hypo.
*Overlaps significantly with the task positive 'Go' network (Supplementary Fig. 1).
BA = Brodmann area.

Table 3 Brain regions involved with performance

Contrast/lobe	BA	x	y	z	Z	mm³	Factor
		Talairach coordinates					
Positive performance							
Frontal							
Dorsal cingulate*	32	10	10	34	3.85	4480	Performance 1
Medial frontal*	6	8	−19	60	3.7	664	Performance 1
Parietal							
Precuneus	7	−27	−67	30	3.09	472	Performance 3
Precuneus/ posterior *Cingulate	7/31	20	−34	45	4.17	7768	Performance 1
Occipital							
Cuneus	18	13	−79	20	3.18	1112	Performance 3
Lingual	18	18	−73	3	3.14	520	Performance 3
Temporal							
Parahippocampal	36	28	−30	−18	4.22	936	Performance 3
Subcortical							
Putamen/insula	13	30	−13	3	4.17	4648	Performance 1
Claustrum/insula	13	−38	−11	−3	3.39	448	Performance 1
Negative performance							
Parietal							
Posterior cingulate*	31	−21	−29	46	4.16	1856	Performance 2
Subcortical							
Dentate*		21	−50	−24	4.04	1112	Performance2

*Overlaps significantly with the task positive 'Go' network (Supplementary Fig. 1).
BA = Brodmann area.

were 13 regions with significant group differences (Fig. 5 and Table 2). Initial review suggested that some regions were defined by bipolar disorder hypoactivation and either healthy control hypoactivation or MDD hyperactivation. Post hoc data reduction was accomplished with principal axis factor analysis (Supplementary Fig. 2), and three factors were derived. The first factor was significantly different between groups [$F(2, 49) = 19.3$, $P < 0.0001$], with MDD hyperactivation and BD hypoactivation relative to healthy controls. The second factor was significantly different between groups [$F(2, 49) = 10.0$, $P < 0.0001$], with bipolar disorder activation greater than healthy controls and MDD, where bipolar disorder showed hyperactivation and healthy controls showed hypoactivation. The third factor was significantly different between groups [$F(2, 49) = 23.3$, $P < 0.0001$], with MDD displaying hyperactivation and healthy controls and bipolar disorder displaying hypoactivation (see Table 2 for factors, loadings, and factor clusters). None of these three factors were significantly correlated with accuracy (P-values > 0.11).

Imaging analysis by accuracy performance

Based on the dimensional perspective that disruption in fronto-parietal cognitive control networks is present in more participants with mood disorders relative to healthy controls but may not be present in all mood disordered

participants, we pursued a regression performance analysis for accuracy. This is based upon the premise that an intermediate phenotype need not be present in all of those in the patient groups, and may be present in some of the healthy controls. These risk intermediate phenotypes could be further evaluated for whether they are distinct to group, clinical or demographic features, or performance markers.

Performance-based regression analysis using accuracy as a predictor of brain activation identified nine regions that positively correlated with performance, and two regions negatively correlated with performance. Activation within these performance-related regions was extracted and initial review suggested that there were group differences in these regions. Post hoc data reduction was used with principal axis factor analysis, and three factors were derived (Supplementary Fig. 2). The first factor was correlated with performance (covarying for group, $r = 0.55$, $P < 0.0001$) and was significantly different between groups [$F(2, 49) = 13.1$, $P < 0.0001$], with MDD and healthy control activation greater than bipolar disorder. The correlation of this activation factor with accuracy was still significant after controlling for diagnostic group ($r = 0.44$, $P < 0.0001$). In a similar fashion, the third factor was significantly correlated with performance ($r = 0.59$, $P < 0.0001$) and was significantly different between groups [$F(2, 49) = 8.7$, $P = 0.001$], with MDD and healthy control activation greater than bipolar disorder (see Table 3 for four clusters). The correlation of this activation factor with accuracy was still significant after controlling for diagnostic group ($r = 0.49$, $P < 0.0001$). The second factor was significantly correlated with performance ($r = 0.49$,

P < 0.0001) but was significantly different between groups [*F*(2, 49) = 3.7, *P* = 0.03], with MDD and bipolar disorder activation greater than healthy controls (see Table 3 for factors, loadings and factor clusters). Images are displayed in Supplementary Fig. 3.

Potential impact of clinical, medication, and demographic variables

To be certain that disease-specific and performance-specific regions were not driven by clinical and demographic features, correlations between group and performance activation factors and several clinical, demographic, and medication loadings were analysed. These were conducted at the individual group level to avoid covariance confounding issues. There were no significant correlations with any of the demographic characteristics (age, education), medication loading for bipolar disorder group (according to the method outlined by Sackeim and colleagues, 2001), nor illness severity characteristics (depression score, number of years of illness, number of mood episodes separately, number of all clinical episode types combined) (all *P*-values > 0.07).

Discussion

To our knowledge, this is the first study to integrate a large sample of females with MDD and bipolar disorder to examine shared dimensions of cognitive control in a behavioural task and a parallel functional MRI task. We demonstrated that there are shared 'performance impairments' in cognitive control among patients with mood disorders (MDD and BD) relative to healthy controls, with a greater number of those in the mood disorder group falling in the 'impaired' range using dimensional, multiple cut-off scores applicable to the clinical setting. We also showed that the neurobiology underlying these seemingly shared dimensions of impairment may not be as clear cut as we had anticipated due to larger more diverse independent group performance features. We demonstrated that performance remains a significant feature to reconcile across behavioural and imaging paradigms and is a viable intermediate phenotype with a clear neuroimaging signature.

A key finding in the present study was the equivalent percentage of sustained attention deficits in a large sample of females with BD and MDD, using a psychometrically well-validated task (Votruba and Langenecker, 2013). These attention difficulties appear to be localized in a performance-specific and disease-shared way to the right posterior parietal cortex. This region had a more nuanced interpretation, as it was different between groups [MDD (hyperactivation) > HC > BD (hypoactivation)] (Supplementary Fig. 2) but also was positively correlated with performance. In this sample, patients who performed well (mainly MDD) engaged this region to a greater extent when they responded correctly, often referred to as compensation. This suggests that with larger samples, including

good and poor performers in each sample, dissociations of disease by performance interactions might be estimated. Not surprisingly, right posterior parietal cortex is an important node in the executive network and suggests a key area of foci for further study and intervention (Posner, 1992). It also solidifies the importance of attention as a domain within the RDoC system, whereby activation and performance can be integrated using dimensional strategies (Cuthbert, 2005).

There were intriguing patterns of distinct processes of disease dysfunction that were not related to performance despite a carefully controlled behavioural paradigm. The general pattern was of equivalent performance and hyperactivation of the MDD group relative to the healthy control group. Typically, patterns of this type have been observed as compensatory (Langenecker *et al.*, 2007*b*). In contrast, the pattern for bipolar disorder was generally poorer performance and hypoactivation. However, as the groups were numerically and statistically equivalent in performance both in and out of the scanner, larger samples are needed before firm conclusions can be made.

There are some important caveats in this study, in part related to the size and scope of the experiments. First, the convergent findings in Experiment 1, with a very large sample size, cannot completely overcome the weakness of having small samples for imaging experiments. Complex interactions of shared, specific, and performance derivatives in imaging studies require larger studies. Second, inherent differences in clinical characteristics of BD relative to MDD—including number and presence of medications, number and 'intensity' of episodes, and different patterns of comorbidity—make parallel comparison studies challenging, although these patterns can be better understood with the use of larger samples. We showed differences in performance among the sample that participated in Experiment 2 when comparing their behavioural performance from outside the scanner to inside the scanner performance, raising the possibility that focus while in the scanner (Experiment 2) can result in greater anxiety and distractibility. Further, characteristics specific to bipolar disorder, such as high rates of comorbid anxiety, could further explain the significant difference in their performance outside versus inside the scanner. Langenecker and colleagues (2007*a*) showed that those with anxiety disorders have an excessive defensive response style, which could account for lower accuracy rates during Level 1 of the PGNG task. These are most likely explanations for behavioural performance differences in and outside the scanner as the PGNG has been well validated in other studies and shows strong reliability (Langenecker *et al.*, 2007*a*, *b*; Votruba and Langenecker, 2013).

Future studies could use the multiple cut-offs in large behavioural samples to pre-select individuals for entry into an imaging study evaluating thresholds of performance impairment; our sample was too small to examine this based on the number of exclusions that need to be considered for imaging studies (e.g. medications, body mass index, comorbidities, willingness to participate). This initial study only used females, thus it is not possible to attribute

these results to both genders with mood disorders. In addition, the task in the imaging environment differed in set-up relative to the computer version administered outside the scanner. It could be that the brief, 14-s block format was much more difficult for the bipolar disorder participants, revealing a greater difficulty in quickly (dis)engaging a mental set. Modifications of this type could be valuable in the future in determining disease-specific intermediate phenotypes against a backdrop of shared sustained attention difficulties in MDD and BD.

Overall, performance and activation during cognitive control tasks may be a good candidate for intermediate phenotypes in mood disorders. Understanding shared elements of risk, including cognitive control, can lead to better differentiation of disease-specific markers. Further, creating tasks with strong behavioural performance markers could augment circuit-based studies and such tasks could be readily translated into clinical practice. Therefore, subtype and dimensional approaches to understanding risk and expression of mood disorders are a promising area of future inquiry, in line with the new RDoC initiative of NIMH.

Supplementary material

Supplementary material is available at *Brain* online.

References

Almeida JR, Versace A, Mechelli A, Hassel S, Quevedo K, Kupfer DJ, et al. Abnormal amygdala-prefrontal effective connectivity to happy faces differentiates bipolar from major depression. Biol Psychiatry 2009; 66: 451–9.

Almeida JR, Versace A, Hassel S, Kupfer DJ, Phillips ML. Elevated amygdala activity to sad facial expressions: a state marker of bipolar but not unipolar depression. Biol Psychiatry 2010; 67: 414–21.

Almeida JRC, Phillips ML. Distinguishing between unipolar depression and bipolar depression: current and future clinical and neuroimaging perspectives. Biol Psychiatry 2013; 73: 111–18.

Beblo T, Sinnamon G, Baune BT. Specifying the neuropsychology of affective disorders: clinical, demographic and neurobiological factors. Neuropsychol Rev 2011; 21: 337–59.

Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clin Psychol Rev 1988; 8: 77–100.

Bertocci MA, Bebko GM, Mullin BC, Langenecker SA, Ladouceur CD, Almeida JR, et al. Abnormal anterior cingulate cortical activity during emotional n-back task performance distinguishes bipolar from unipolar depressed females. Psychol Med 2012; 42: 1417–28.

Clayton PJ. The comorbidity factor: establishing the primary diagnosis in patients with mixed symptoms of anxiety and depression. J Clin Psychiatry 1990; 51: 35–9.

Cuellar AK, Johnson SL, Winters R. Distinctions between bipolar and unipolar depression. Clin Psychol Rev 2005; 25: 307–39.

Cuthbert BN. Dimensional models of psychopathology: Research agenda and clinical utility. J Abnormal Psychol 2005; 114: 565–9.

Dolan RJ, Bench CJ, Liddle PF, Friston KJ, Frith CD, Grasby PM, et al. Dorsolateral prefrontal cortex dysfunction in the major psychoses: Symptom or disease specificity? J Neurol Neurosurg Psychiatry 1993; 56: 1290–4.

Drevets WC. Orbitofrontal cortex function and structure in depression. Ann N Y Acad Sci 2007; 1121: 499–527.

First MB, Spitzer RL, Gibbon M. Structured clinical interview for DSM-IV Axis 1 disorder. New York: Biometrics Research Department, New York State Psychiatric Institute; 1995.

Garavan H, Ross T, Stein E. Right hemispheric dominance of inhibitory control: an event-related functional MRI study. Proc Natl Acad Sci USA 1999; 96: 8301–6.

Garvey MJ, Tollefson GD, Tuason VB. Is chronic primary major depression a distinct depression subtype? Compr Psychiatry 1986; 27: 446–8.

Glover GH, Thomason ME. Improved combination of spiral-in/out images for BOLD fMRI. Magn Reson Med 2004; 51: 863–8.

Godard J, Grondin S, Baruch P, Lafleur MF. Psychosocial and neurocognitive profiles in depressed patients with major depressive disorder and bipolar disorder. Psychiatry Res 2011; 190: 244–52.

Goodwin FK, Jamison KR. Manic-depressive illness: bipolar disorders and recurrent depression. New York, NY: Oxford University Press; 2007.

Hamilton JP, Siemer M, Gotlib IH. Amygdala volume in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. Mol Psychiatry 2008; 13: 993–1000.

Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol 1967; 6: 278–96.

Hirschfeld RM, Lewis L, Vornik LA. Perceptions and impact of bipolar disorder: how far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder. J Clin Psychiatry 2003; 64: 161–74.

Kerr N, Scott J, Phillips ML. Patterns of attentional deficits and emotional bias in bipolar and major depressive disorder. Br J Clin Psychol 2005; 44: 343–56.

Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001; 16: 606–13.

Lange N, Lake S, Sperling R, Brown J, Routledge C, Albert M, et al. Two macroscopic and microscopic brain imaging studies of human hippocampus in early Alzheimer's disease and schizophrenia research. Stat Med 2004; 23: 327–50.

Langenecker SA, Nielson KA. Frontal recruitment during response inhibition in older adults replicated with fMRI. Neuroimage 2003; 20: 1384–92.

Langenecker SA, Caveney A, Persad C, Giordani B. Semantic list learning test [Computer Program]. Ann Arbor; 2004.

Langenecker SA, Bieliauskas LA, Rapport LJ, Zubieta JK, Wilde EA, Berent S. Face emotion perception and executive functioning deficits in depression. J Clin Exp Neuropsychol 2005; 27: 320–33.

Langenecker SA, Caveney AF, Giordani B, Young EA, Nielson KA, Rapport LJ, et al. The sensitivity and psychometric properties of a brief computer-based cognitive screening battery in a depression clinic. Psychiatry Res 2007a; 152: 143–54.

Langenecker SA, Kennedy SE, Guidotti LM, Briceno EM, Own LS, Hooven T, et al. Frontal and limbic activation during inhibitory control predicts treatment response in major depressive disorder. Biol Psychiatry 2007b; 62: 1272–80.

Langenecker SA, Saunders EF, Kade AM, Ransom MT, McInnis MG. Intermediate cognitive phenotypes in bipolar disorder. J Affect Disord 2010; 122: 285–93.

Maalouf FT, Klein C, Clark L, Sahakian BJ, Labarbara EJ, Versace A, et al. Impaired sustained attention and executive dysfunction: bipolar disorder versus depression-specific markers of affective disorders. Neuropsychologia 2010; 48: 1862–8.

Meyer-Lindenberg A, Weinberger DR. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. Nat Rev Neurosci 2006; 7: 818–27.

Moore JH, Williams SM. Epistasis and its implications for personal genetics. Am J Hum Genet 2009; 85: 309–20.

Nielson KA, Langenecker SA, Garavan H. Differences in the functional neuroanatomy of inhibitory control across the adult life span. Psychol Aging 2002; 17: 56–71.

Nurnberger JI Jr Blehar MC, Kaufmann CA, York-Cooler C, Simpson SG, Harkavy-Friedman J. al. Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. Arch Gen Psychiatry 1994; 51: 849–59.

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, 833–57.

Posner M. Attention as a cognitive and neural system. Curr Direct Psychol Sci 1992; 1: 11–15.

Ravnkilde B, Videbech P, Clemmensen K, Egander A, Rasmussen NA, Gjedde A, et al. The Danish PET/depression project: cognitive function and regional cerebral blood flow. Acta PsychiatrScand 2003; 108: 32–40.

Ryan KA, Vederman AC, McFadden EM, Welsom AL, Kamali M, Langenecker SA, et al. Executive functioning components change across phases of bipolar disorder. Bipolar Disord 2012; 14: 527–36.

Sackeim HA. The definition and meaning of treatment-resistant depression. J Clin Psychiatry 2001; ; 62 ((Suppl 16): 10–7.

Savitz JB, Drevets WC. Imaging phenotypes of major depressive disorder: genetic correlates. Neuroscience 2009; 164: 300–30.

Sheline YI, Sanghavi M, Mintun MA, Gado MH. Depression duration but not age predicts hippocampal volume loss in medically healthy women with recurrent major depression. J Neurosci 1999; 19: 5034–43.

Shipley WC. Institute of living scale. Los Angeles: Western Psychological Services; 1946.

Smith DJ, Muir WJ, Blackwood DH. Neurocognitive impairment in euthymic young adults with bipolar spectrum disorder and recurrent major depressive disorder. Bipolar Disord 2006; 8: 40–6.

Surguladze SA, Elkin A, Ecker C, Kalidindi S, Corsico A, Giampietro V, et al. Genetic variation in the serotonin transporter modulates neural system-wide response to fearful faces. Genes Brain Behav 2008; 7: 543–51.

Taylor Tavares JV, Clark L, Cannon DM, Erickson K, Drevets WC, Sahakian BJ. Distinct profiles of neurocognitive function in unmedicated unipolar depression and bipolar II depression. Biol Psychiatry 2007; 62: 917–924.

Taylor Tavares JV, et al. Neural basis of abnormal response to negative feedback in unmedicated mood disorders. Neuroimage 2008; 42: 1118–26.

Votruba KL, Langenecker SA. Factor structure, construct validity and age- and education-based normative data for the parametric Go/No-go test. J Clin Exp Neuropsychol 2013; 35: 132–46.