Psychological Medicine (2017), **47**, 2640–2652. © Cambridge University Press 2017
doi:10.1017/S003329171700112X

ORIGINAL ARTICLE

Functional neuroimaging of sex differences in autobiographical memory recall in depression

K. D. Young[1,2]*, J. Bodurka[1,3,4] and W. C. Drevets[5]

[1] *Laureate Institute for Brain Research, Tulsa, OK, USA*
[2] *Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA*
[3] *Biomedical Engineering Center, Norman, OK, USA*
[4] *Stephenson School of Biomedical Engineering University of Oklahoma College of Engineering, Norman, OK, USA*
[5] *Janssen Research and Development, LLC, of Johnson & Johnson, Inc., New Brunswick, NJ, USA*

Background. Females are more likely than males to develop major depressive disorder (MDD). The current study used fMRI to compare the neural correlates of autobiographical memory (AM) recall between males and females diagnosed with MDD. AM overgenerality is a persistent cognitive deficit in MDD, the magnitude of which is correlated with depressive severity only in females. Delineating the neurobiological correlates of this deficit may elucidate the nature of sex-differences in the diathesis for developing MDD.

Methods. Participants included unmedicated males and females diagnosed with MDD ($n = 20$/group), and an age and sex matched healthy control group. AM recall in response to positive, negative, and neutral cue words was compared with a semantic memory task.

Results. The behavioral properties of AMs did not differ between MDD males and females. In contrast, main effects of sex on cerebral hemodynamic activity were observed in left dorsolateral prefrontal cortex and parahippocampal gyrus during recall of positive specific memories, and middle prefrontal cortex (mPFC), and precuneus during recall of negative specific memories. Moreover, main effects of diagnosis on regional hemodynamic activity were observed in left ventrolateral prefrontal cortex and mPFC during positive specific memory recall, and dorsal anterior cingulate cortex during negative specific memory recall. Sex × diagnosis interactions were evident in the dorsomedial prefrontal cortex, caudate, and precuneus during positive memory recall, and in the posterior cingulate cortex, insula, precuneus and thalamus during negative specific memory recall.

Conclusions. The differential hemodynamic changes conceivably may reflect sex-specific cognitive strategies during recall of AMs irrespective of the phenomenological properties of those memories.

Received 28 July 2016; Revised 30 March 2017; Accepted 31 March 2017; First published online 27 April 2017

Key words: Autobiographical memory, depression, emotion, fMRI, sex differences.

Introduction

Females are twice as likely as males to develop major depressive disorder (MDD) (Kessler, 2003). Examining the neurocognitive variables underlying this sex difference and determining their neural bases may prove informative for developing interventions to prevent or treat MDD. Considerable research has been devoted to elucidating factors that may explain this difference, which has investigated variables in childhood environment, social roles, cultural norms, hormones, and genetic factors (reviewed in Piccinelli & Wilkinson, 2000). The strongest associations have involved traumatic childhood experiences, competing social roles, and maladaptive coping skills (e.g. rumination; Nolen-Hoeksema, 1987; Piccinelli & Wilkinson, 2000). Some of these variables, including traumatic childhood events (Stokes *et al.* 2004; Williams *et al.* 2007) and ruminative coping styles (Watkins & Teasdale, 2001; Williams *et al.* 2007), have been associated with an overgeneral autobiographical memory (AM) retrieval style.

Overgeneral AM is characterized by recalling an increased number of categorical memories (recollection of recurring events without reference to a single event) and a reduced number of specific memories (recollection of a single event that occurred at an identified time and place) relative to healthy controls (HCs) (Williams & Broadbent, 1986; Williams *et al.* 2007). Overgeneral AM recall is a persistent cognitive deficit observed in MDD (Williams *et al.* 2007), as well as in those at high-risk for developing MDD (Young *et al.* 2013*a*, 2015). Extant data indicate that in non-clinical samples, an overgeneral AM retrieval style is associated with depressive symptoms, *but only in females*

* Address for correspondence: K. D. Young, Western Psychiatric Institute and Clinic, 3811 O'Hara Street, Pittsburgh, PA, 15213, USA. (Email: youngk@pitt.edu)

(Ros *et al.* 2014). In contrast, in males repressive coping has been the only variable associated with depressive symptoms (Ros *et al.* 2014). The authors hypothesize that this sex difference is attributable to emotional elements of AMs being more integral and salient to females than males, and therefore more relevant to the onset of depression.

Studies examining sex differences in HCs in the phenomenological properties of AMs have provided inconsistent results, with some reporting females recalling more detailed and more emotionally intense AMs than males (Ross & Holmberg, 1992; Seidlitz & Diener, 1998), and others reporting no sex differences in the properties of AMs (Rubin *et al.* 1999; Piefke *et al.* 2005; St Jacques *et al.* 2011). A recent study by our laboratory also failed to find sex differences in the properties of recalled AMs in HCs, but nevertheless found sex differences in the brain regions underlying the recall of these AMs (Young *et al.* 2013*c*). Females showed increased activity relative to males in regions involved in emotional control, including the dorsolateral prefrontal cortex (DLPFC) (Levesque *et al.* 2003), as well as in regions involved in self-referential, emotional, and salience processing, including the insula, precuneus, posterior cingulate cortex (PCC), and amygdala (Davis & Whalen, 2001; Seeley *et al.* 2007; Buckner *et al.* 2008), particularly during negative AM recall. Interestingly, these results parallel neurophysiological differences seen when MDD and HCs engage in AM recall, including increased precuneus, insula, PCC, and amygdala activity during specific and negative AM recall in MDDs relative to HCs (Young *et al.* 2014, 2016).

Our previous results in HCs are also consistent with (Piefke *et al.* 2005) who reported females had greater hemodynamic activity in DLPFC than males while recalling AMs. Furthermore, other neuroimaging studies found males to have increased activity in the hippocampus/parahippocampus (Piefke *et al.* 2005; St Jacques *et al.* 2011). These studies all reported differences in functional activity during AM recall in the absence of behavioral differences. Collectively, these results support a hypothesis that males and females use different cognitive strategies during AM encoding and recall such that males use spatial contextual information to a greater degree, while females rely more on verbal and emotional information (Seidlitz & Diener, 1998).

While many studies have explored hormonal, genetic, social, and behavioral influences of sex differences in MDD (Piccinelli & Wilkinson, 2000), only one study used fMRI to examine the functional anatomic correlates of this difference. Briceno *et al.* (2015) examined differences in hemodynamic activity as males and females with and without MDD performed a facial-emotion identification task. Young females with MDD showed increased activity in the PCC and middle frontal gyrus, regions implicated in emotional processing, compared with young HC and MDD males. All emotions were compared with the control task, and potential sex differences during the different emotional conditions were not examined. Although numerous studies used fMRI to characterize differences in cerebral function between MDDs and HCs (reviewed in Nejad *et al.* 2013), few fMRI studies have investigated sex differences in MDD. Such studies hold the potential to elucidate the neurobiological basis underlying differences in the diathesis of developing MDD between males and females.

While we previously examined neural correlates of AM recall in depressed and at-risk individuals, our samples were too small to investigate sex differences. Therefore, the current study assessed sex differences in the phenomenological properties of AMs and the neurophysiological correlates of AM recall. Because overgeneral AM recall appears to be a persistent cognitive deficit in MDD (Williams *et al.* 2007), and because research suggests that AM impairment relates to depressive symptoms only in *females* (Ros *et al.* 2014), understanding how males and females differ in the brain regions recruited during AM recall may elucidate the neural basis for sex differences in the prevalence of MDD, and may clarify how sex differences in AM recall influence the development of MDD.

Based on previous studies examining sex differences in HCs during AM recall, and our previous studies of AM recall in MDD, we predicted that MDD females would show increased activity in regions involved in emotional regulation, such as the DLPFC, during recall of positive AMs and increased activity in regions implicated in self-referential processing, such as the medial PFC and precuneus during recall of negative AMs. Furthermore, we hypothesized that MDD males would have increased activity in temporal regions involved in memory and spatial processing such as the hippocampus during specific AM recall regardless of valence.

We examined differences in hemodynamic activity during positive and negative specific AM recall separately. Although research in non-depressed subjects suggests overgeneral AM recall is best fit by a one-factor model to which the emotional valence of the memory does not contribute (Griffith *et al.* 2009; Heron *et al.* 2012), these studies included only healthy participants. It remains unclear, however, whether a one-factor model is also appropriate for clinical samples that manifest emotional processing biases, such as MDD. Previous studies in MDD participants have been inconsistent regarding the need to include emotional valence as a factor in studies of overgeneral

https://doi.org/10.1017/S003329171700112X Published online by Cambridge University Press

AM recall, with some finding support for a one factor model (e.g. Soderlund *et al.* 2014; Champagne *et al.* 2016) but others finding differential effects of valence (e.g. Brittlebank *et al.* 1993; Rawal & Rice, 2012). We previously found differential hemodynamic activity in MDD *v.* control participants during positive as well as negative AM recall (Young *et al.* 2013*b*, 2016), arguing against the application of one-factor models and leading us to instead form emotion-specific hypotheses. Such analyses also allow closer comparison with the results of our previous studies of AM recall (Young *et al.* 2013*c*). To provide information regarding the extent to which observed sex differences in MDD participants represented abnormal neural function or reflected sex differences more generally, we included data previously collected from HCs (Young *et al.* 2013*c*).

Methods and materials

Participants

Forty unmedicated right-handed adults (20 males) with MDD in a current depressive episode according to Diagnostic and Statistical Manual (DSM)-IV-TR criteria (APA, 2000) and 40 HCs ages 18–55 participated. Volunteers, recruited from the community via advertisements and evaluated at the Laureate Institute for Brain Research, underwent medical and psychiatric screening using the Structural Clinical Interview for *DSM-IV* disorders (SCID; First *et al.* 2002) and an unstructured interview with a psychiatrist. Exclusion criteria included current pregnancy, general MRI exclusions, serious suicidal ideation, psychosis, major medical or neurological disorders, exposure to any medication likely to influence CNS function within the 3 weeks prior to entry, and meeting *DSM-IV-TR* criteria for alcohol and/or substance abuse within the previous 1 year, or for a lifetime history of substance dependence (excepting nicotine). Additional exclusion criteria for HCs included a current or previous personal or family history of an Axis I psychiatric condition. After receiving a complete explanation of the study procedures, all participants provided written informed consent to participate, as obtained according to the Declaration of Helsinki and approved by the Western IRB. Subjects received financial compensation for their participation.

Intelligence testing was performed using the two-subtest version of the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999). Depressive symptoms were rated on the day of scanning the 21-item Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960).

fMRI AM task

Details of the fMRI data acquisition parameters and AM task have been reported previously (Young *et al.* 2013*a*, *c*). Briefly, blood-oxygen-level- dependent (BOLD) fMRI was performed on a 3 T GE MR750 scanner and an eight-channel head coil. Participants were presented with a cue word (20 positive, 20 negative, 20 neutral) and instructed to recall a past experience for 12s and rate their retrieved memory on specificity and valence. Participants were instructed on the definitions of the standard memory categorizations of specific, categorical, extended, and semantic (Williams & Broadbent, 1986; Williams *et al.* 2007), and were able to provide and classify examples of each type before scanning. AM recall was compared with a semantic example generation condition in which participants were presented with a cue word (10 positive, 10 negative, 10 neutral) and instructed to think of seven examples from the semantic category, then rate the ease with which they generated examples and report the number generated. Following each cue presentation and set of ratings, participants engaged in a riser detection task involving non- word letter strings as a control for visual input/attention.

Following the scan, the experimenter presented participants with all AM cue words again in the same order as during fMRI. Participants were asked to describe the memory to allow the experimenter to corroborate participants' ratings of specificity. In addition to the standard classifications, a 'can't recall' category was included if either the participant was unable to recall the memory retrieved during fMRI or the memory provided did not match the specificity rating given by the participant. An independent rater scored all responses to establish inter-rater reliability (agreement = 94%). Only memories for which both participant and experimenters' ratings of specificity agreed were included in the analysis. Participants also rated each memory on three additional properties: arousal (five point scale ranging from very low to very high), vividness (five point scale ranging from 'not at all vivid' to 'perfectly clear and vivid') and age when memory occurred (childhood, adolescence, adulthood after age 18 but >1 year prior to scan, adulthood between 6 months and 1 year prior to scan, adulthood within 6 months prior to scan).

Data analysis

fMRI analysis was performed using AFNI (http://afni.nimh. nih.gov/afni), and included slice timing correction, within-subject realignment, coregistration between anatomical and functional images, spatial normalization to a common stereotaxic array (Talairach & Tournoux, 1988), and spatial smoothing (Gaussian

https://doi.org/10.1017/S003329171700112X Published online by Cambridge University Press

kernel, 4 mm full width at half maximum). General linear model analysis was performed using 3dDeconvolve. For each participant, the response to each event type was modeled as a boxcar function convolved with a canonical hemodynamic response function. The main effects of interest were cue word presentations that prompted specific memory recall of positive memories, specific memory recall of negative memories, example generation in response to positive cues, and example generation in response to negative cues. In addition to regressors modeling main effects and subject motion, each design matrix included regressors modeling rating selection and cue presentations in which other types of memories were recalled (categorical, extended, semantic). The riser detection task was modeled as the baseline.

At the group level, a Sex × Diagnosis 3dANOVA was performed for Specific Positive Memories *v*. Positive Example Generation and Specific Negative Memories *v*. Negative Example Generation to examine emotion-dependent differences during specific AM recall. The significance criterion for detecting differences was set at $p_{corrected} < 0.05$ determined using 3dClustSim (cluster size > 30 voxels, thresholded at voxel $p < 0.001$). A supplementary analysis was conducted using AFNI's 3dLME to specifically examine the sex × diagnosis × valence interaction to further justify emotion-focused contrasts (online Supplement Table ST2 and Figure ST1). In the main manuscript only the emotion-focused contrasts are presented.

For each region demonstrating a significant difference in activity (for main effects and interactions), we extracted beta weights from those clusters and conducted independent sample *t* tests in SYSTAT to determine which groups were driving the effects. The criterion for significance was set at $p < 0.001$, corrected for 36 comparisons (Bonferroni). Effect sizes were calculated for main effects using Cohen's d and for interactions using partial eta squared.

Assessment of behavioral performance during fMRI

Behavioral data were analyzed using SYSTAT13 (Systat Software Inc., USA). Only MDD participants were included in these analyses, and follow up tests including HCs were planned only for contrasts for which significant sex differences were found in the MDD participants [A sex × group × valence × specificity analysis of variance (ANOVA) can be found in the supplement]. Potential sex differences in age, IQ, depression ratings, performance on the riser detection and example generation control tasks, and the percent of memories recalled at each specificity level were assessed using a one-way ANOVA. No sex differences were expected on these characteristics; these assessments explored whether a nonspecific difference may have confounded any difference identified in the AM retrieval condition; thus the associated *p* values were not corrected for multiple testing.

The *a priori* hypothesis testing focused on the properties of specific and categorical memories, as too few exemplars of the other types of AMs were retrieved to allow sufficient power to detect differences. To increase power, the following variables were collapsed: low and very low arousal formed a low arousal variable, high and very high arousal produced a high arousal variable, low and very low vividness produced a low vividness variable, high and very high vividness produced a high vividness variable, somewhat positive and positive produced a positive variable, and somewhat negative and negative produced a negative variable. Repeated measures ANOVAs were performed for repeated measure Type (Specific or Categorical) and either Valence (positive, negative, neutral), Arousal (low, medium, high), Vividness (low, medium, high), or Age (childhood, adolescence, after age 18 but longer than 1 year prior to scan, between 6 months and 1 year prior to scan and <6 months before scan) for the dependent variable Percent of Memories recalled. The threshold criterion for significance was set at $p < 0.002$, corrected for 28 multiple comparisons (Bonferroni).

Finally, to determine whether AM overgenerality was differentially associated with depressive symptoms in males and females as previously reported (Ros *et al.* 2014), Pearson correlations were performed between HDRS scores and the percent of specific, specific positive, and specific negative AMs recalled separately for MDD males and females.

Results

Behavioral results

Table 1 presents participant characteristics, general task performance, and associated statistics for MDD participants. MDD males and females did not differ in any variable measured, including mean age, IQ, HDRS score, accuracy on the riser detection task, number of examples generated for the distinctly valenced categories, the ease with which such examples were generated, or for the percent of memories recalled for each specificity type.

Properties of specific memory recall for the MDDs are presented in online Supplementary Table S1 and no significant interaction with Sex was observed ($Fs < 1.96$, $ps > 0.15$).

For MDD males and females combined, there was no significant correlation between HDRS and AM recall scores. In MDD males there was no correlation between HDRS scores and any AM variable

https://doi.org/10.1017/S003329171700112X Published online by Cambridge University Press

Table 1. *Participant characteristics by sex*

	Males	Females	F (1,38)	p value	Effect size (d)
Demographics					
Age	36.1 (10.2)	35.5 (8.13)	0.05	0.82	0.07
IQ [as determined by WASI]	110 (10.6)	108 (12.7)	0.47	0.50	0.17
HDRS	18.8 (5.16)	19.6 (6.63)	0.20	0.66	0.14
Control tasks					
Riser % correct	71.1 (14.6)	72.3 (14.9)	0.08	0.79	0.08
% Example generation rated easy					
Positive	85.4 (13.1)	87.7 (22.0)	0.57	0.62	0.13
Negative	85.3 (13.9)	83.1 (19.8)	0.54	0.62	0.13
Neutral	88.5 (12.7)	80.7 (15.4)	1.88	0.18	0.55
# of examples generated					
Positive	5.33 (0.99)	5.07 (1.21)	0.61	0.44	0.24
Negative	5.11 (1.24)	5.48 (1.14)	0.64	0.45	0.31
Neutral	5.76 (1.01)	5.51 (1.14)	0.62	0.44	0.23
Memory type (*% Recalled*)					
Specific	40.9 (16.6)	39.7 (13.2)	0.08	0.78	0.08
Categorical	32.8 (12.2)	35.5 (12.9)	0.53	0.47	0.22
Extended	3.81 (2.75)	2.53 (2.68)	2.49	0.12	0.54
Semantic	5.53 (4.77)	7.42 (5.85)	1.36	0.25	0.35
No memory	2.08 (3.86)	3.70 (5.83)	1.13	0.29	0.33
Can't recall post scan	13.6 (8.95)	11.1 (5.65)	2.55	0.12	0.33

Numbers in parentheses indicate one standard deviation of the mean.
HRSD, Hamilton Rating Scale for Depression; WASI, Wechsler Abbreviated Scale of Intelligence.

($rs < -0.27$, $ps > 0.26$ Fig. 1*a*, *c*). In MDD females, the correlation between specific AMs and HDRS scores approached significance ($r = -0.33$, $p = 0.08$; Fig. 1*b*), and the correlation between HDRS scores and the percent of specific AMs rated as positive in valence was significant ($r = -0.47$, $p = 0.02$; Fig. 1*d*). The correlation between HDRS scores and the percent of AMs rated as negative in valence, while in the opposite direction, was not significant in MDD females ($r = 0.25$, $p = 0.22$).

Imaging results

The linear mixed effects analysis (online Supplement Table ST2 and Fig. SF1) support the application of the two-factor model, as there was a significant group × sex × valence interaction in regional activity. We therefore ran analyses separately for positively and negatively valenced AMs. Table 2 shows the results of the separate whole brain sex × diagnosis ANOVAs for positive and negative memory recall. There were main effects of both sex and diagnosis (Table 2) and significant sex × diagnosis interactions (Table 2).

Main effect of sex

For positive specific memories *v*. example generation in response to positive cues (Fig. 2*a*), regardless of diagnosis, males had increased activity relative to females in right parahippocampal gyrus, and females showed increased activity relative to males in left DLPFC. During recall of negative specific memories *v*. negative example generation in response to negative cues (Fig. 3*a*) females had increased left middle frontal gyrus (mPFC) and right precuneus activity relative to males.

Main effect of diagnosis

The main effect of diagnosis revealed that regardless of sex, participants with MDD had increased activity in left ventrolateral prefrontal cortex (VLPFC) during positive specific memory recall *v*. positive example generation (Fig. 2*b*), while HCs had increased activity in bilateral mPFC. During recall of specific negative memories *v*. negative example generation (Fig. 3*b*), participants with MDD had increased activity in the right dorsal anterior cingulate cortex (dACC) relative to HCs.

Sex × diagnosis interaction

Table 2 shows regions where a significant sex × diagnosis interaction occurred. During recall of specific positive AMs relative to positive Example Generation

https://doi.org/10.1017/S003329171700112X Published online by Cambridge University Press

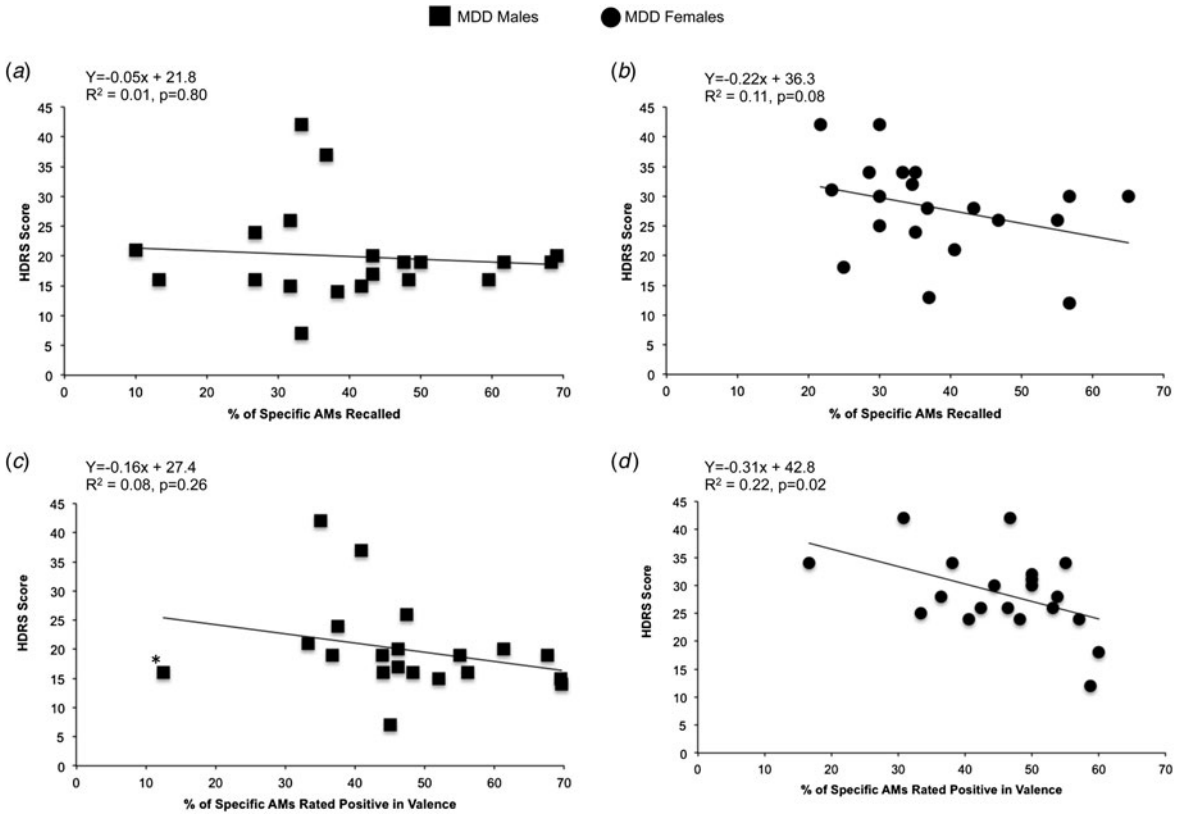

Fig. 1. Correlations between Memory Recall and Depression Scores by Sex. Correlation between HDRS scores and the percent of specific memories recalled in (*a*) females and (*b*) males. Correlation between HDRS scores and the percent of specific memories rated as positive in valence in (*c*) females and (*d*) males. *denotes an outlier. However, while removal of the outlier increases the R2 value to 0.12, the correlation still does not reach statistical significance ($p = 0.12$).

(Fig. 2*c*), in right dorsomedial prefrontal cortex (DMPFC), MDD males had increased activity relative to all other groups ($ps < 0.001$), while HC and MDD females did not differ from each other ($p = 0.36$). In right caudate, MDD females had decreased activity relative to all other groups ($ps < 0.001$), while the other groups did not differ significantly from each other ($p = 0.51$). For activity in bilateral precuneus, all groups differed significantly from each other ($ps < 0.001$).

During recall of specific negative AMs relative to negative Example Generation (Fig. 3*c*), groups differed in left PCC, right insula, bilateral precuneus, and left thalamus. In PCC and insula, MDD females had greater activity compared with all other groups ($ps < 0.001$). HC females also had increased activity relative to both male groups ($ps < 0.001$), who did not differ significantly from each other ($p = 0.70$). In precuneus, healthy males had decreased activity relative to all other groups ($ps < 0.001$), while in left precuneus, the other groups did not differ from each other ($ps > 0.06$), and in the right precuneus females had increased activity relative to males ($ps < 0.001$). Finally, thalamus activity was greatest in the female MDD group ($ps < 0.001$), but all groups differed significantly from each other ($ps < 0.001$).

Discussion

This study found sex differences in regional hemodynamic activity during AM recall in MDD participants in the absence of behavioral differences in the properties of the recalled AMs, consistent with previous findings in healthy males and females that the properties of recalled AMs do not differ (Rubin *et al.* 1999; Piefke *et al.* 2005; St Jacques *et al.* 2011; Young *et al.* 2013*c*). We extend this observation to currently depressed, unmedicated adults. Therefore, observed sex differences in regional hemodynamic activity do not appear attributable to differences in the phenomenological properties of recall measured in the current study.

During recall of positive specific memories, females compared with males had increased DLPFC activity, regardless of diagnosis. The DLPFC is involved in cognitive control, including modulation of emotional behavior, and higher DLPCF activity is associated with reductions in hippocampal/parahippocampal

https://doi.org/10.1017/S003329171700112X Published online by Cambridge University Press

Table 2. *Results of the sex × diagnosis × valence ANOVA*

				β Weight					
				MDD		HC			
(a) Area	*x*, *y*, *z*	cluster size	*t* Value	Males	Females	Males	Females	Effect size (*d*)	Direction of effect
Main effect of sex									
Positive specific memories *v.* positive example Generation									
L Superior Frontal/BA 8/DLPFC	−17, 19, 48	63	4.26	0.16 (0.15)	0.43 (0.39)	0.15 (0.17)	0.39 (0.20)	1.03	F > M
R Parahippocampal G	29, −27, −`20	37	4.07	0.32 (0.28)	0.04 (0.25)	0.28 (0.16)	0.03 (0.28)	1.06	M > F
Negative specific memories *v.* negative example Generation									
L middle frontal G	−31, 17, 54	56	3.78	0.17 (0.33)	0.51 (0.39)	0.19 (0.32)	0.48 (0.30)	0.94	F > M
R Precuneus	1, −72, 37	78	3.52	−0.12 (0.48)	0.32 (0.44)	−0.11 (0.64)	0.28 (0.31)	0.86	F > M
Main effect of diagnosis									
Positive specific memories *v.* positive example generation									
L inferior frontal G/BA 45/VLPFC	−59, 13, 22	98	4.49	0.47 (0.46)	0.62 (0.36)	0.08 (0.42)	0.10 (0.30)	1.15	MDD > HC
L Middle Frontal G	−25, −11, 44	37	3.67	0.01 (0.12)	−0.02 (0.08)	0.12 (0.10)	0.11 (0.10)	1.07	HC > MDD
R Middle Frontal G/BA 8	27, 9, 38	41	3.33	−0.03 (0.19)	−0.04 (0.15)	0.19 (0.06)	0.15 (0.04)	1.94	HC > MDD
Negative specific memories *v.* negative example generation									
R dACC	1, 13, 33	162	3.31	0.62 (0.56)	0.52 (0.39)	−0.20 (0.88)	−0.31 (0.87)	1.18	MDD > HC

				β Weight					
(b) Area	*x*, *y*, *z*	cluster size	*F* Value	Males	Females	Males	Females	Effect size (η_p^2)	Direction of effect
Sex × diagnosis interaction									
Positive specific memories *v.* positive example Generation									
R Middle Frontal G/BA 9/DMPFC	9, 49, 14	288	11.53	0.27 (0.24)	−0.16 (0.19)	0.08 (0.17)	−0.16 (0.20)	0.25	mMDD > mHC > fHC = fMDD
R Caudate	19, 19, 8	35	12.58	0.08 (0.04)	−0.06 (0.03)	0.06 (0.08)	0.05 (0.04)	0.08	mMDD = mHC = fHC > fMDD
L Precuneus	−14, −49, 51	94	8.22	0.15 (0.17)	−0.21 (0.17)	0.36 (0.27)	0.06 (0.13)	0.10	mHC > mMDD > fHC > fMDD
R Precuneus	9, −57, 56	37	15.01	0.11 (0.12)	−0.17 (0.07)	0.16 (0.10)	0.01 (0.09)	0.13	mHC > mMDD > fHC > fMDD
Negative specific memories *v.* negative example generation									
L PCC	−1, −35, 28	94	6.88	−0.03 (0.25)	0.29 (0.17)	0.01 (0.28)	0.15 (0.15)	0.07	fMDD > fHC > mMDD = mHC
R insula	35, −13, 14	93	13.32	−0.09 (0.23)	0.18 (0.07)	−0.15 (0.26)	0.08 (0.14)	0.06	fMDD > fHC > mMDD = mHC
L precuneus	−3, −49, 62	34	14.36	0.16 (0.34)	0.25 (0.27)	−0.34 (0.31)	0.22 (0.34)	0.14	fMDD = fHC = mMDD > mHC
R precuneus	11, −77, 46	72	14.52	0.03 (0.23)	0.34 (0.26)	−0.66 (0.50)	0.29 (0.22)	0.16	fMDD = fHC > mMDD > mHC
L thalamus	−1, −13, 14	45	14.37	−0.22 (0.60)	0.32 (0.36)	0.09 (0.48)	0.16 (0.39)	0.20	fMDD > fHC > mHC > mMDD

BA, Brodmann Area; C, cortex; dACC, dorsal anterior cingulate cortex; DLPFC, dorsolateral prefrontal cortex; DMPFC, dorsomedial prefrontal cortex; G, gyrus; L, left; PCC, posterior cingulate cortex; R, right; VLPFC, ventrolateral prefrontal cortex.

Coordinates correspond to the stereotaxic array of Talairach & Tournoux (1988). Cluster size refers to the number of contiguous voxels for which the voxel t value corresponds to $p_{corrected} < 0.05$. Numbers in parentheses indicate on standard deviation of the mean.

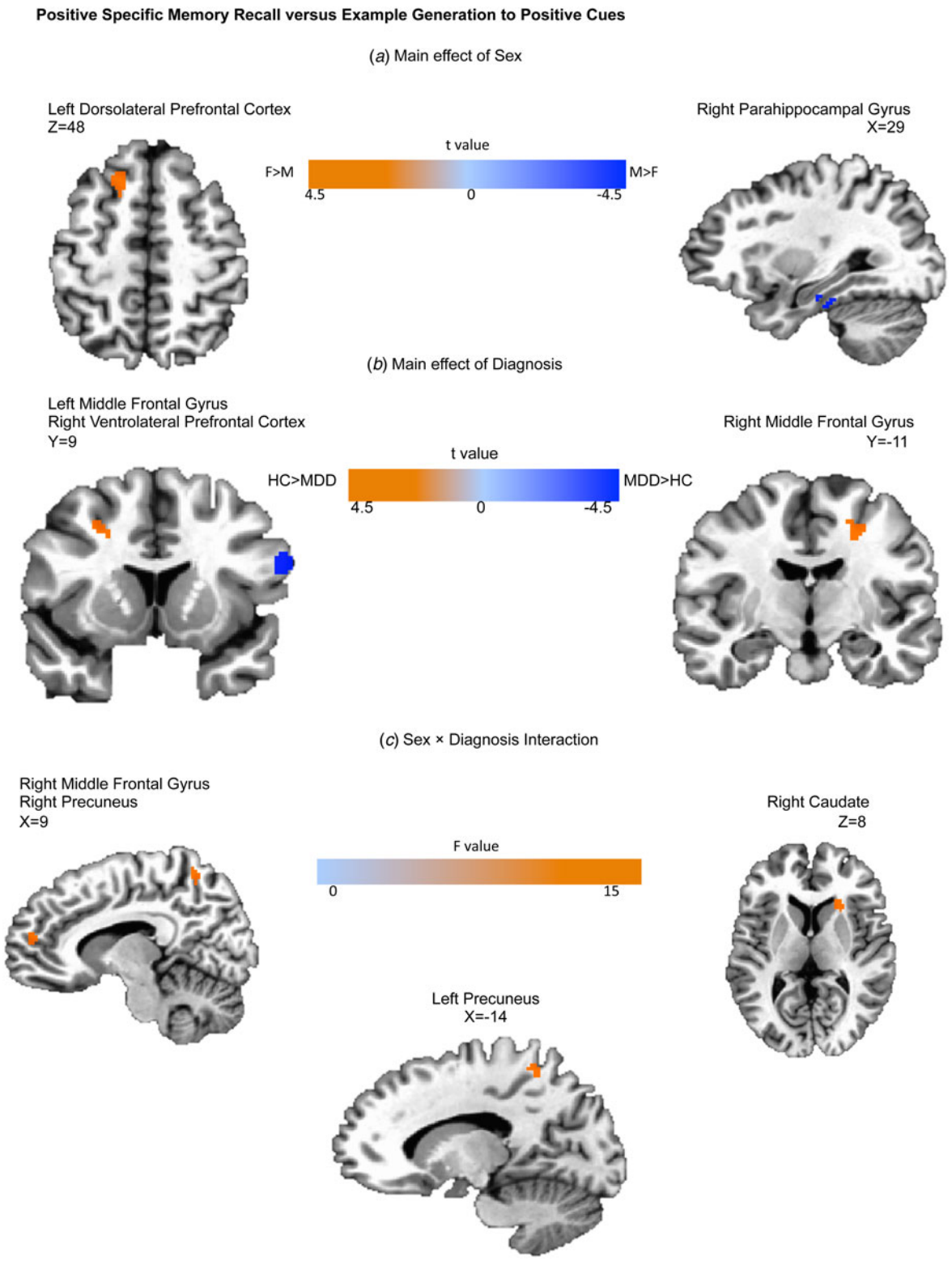

Fig. 2. Brain Regions showing a sex × diagnosis Interaction in Hemodynamic Activity during Positive Autobiographical Memory Recall. (*a*) Main effect of sex (*b*) main effect of diagnosis and (*c*) sex × diagnosis interaction for hemodynamic activity during positive specific memory recall *v*. example generation in response to positive cues ($p_{\text{corrected}} < 0.05$).

https://doi.org/10.1017/S003329171700112X Published online by Cambridge University Press

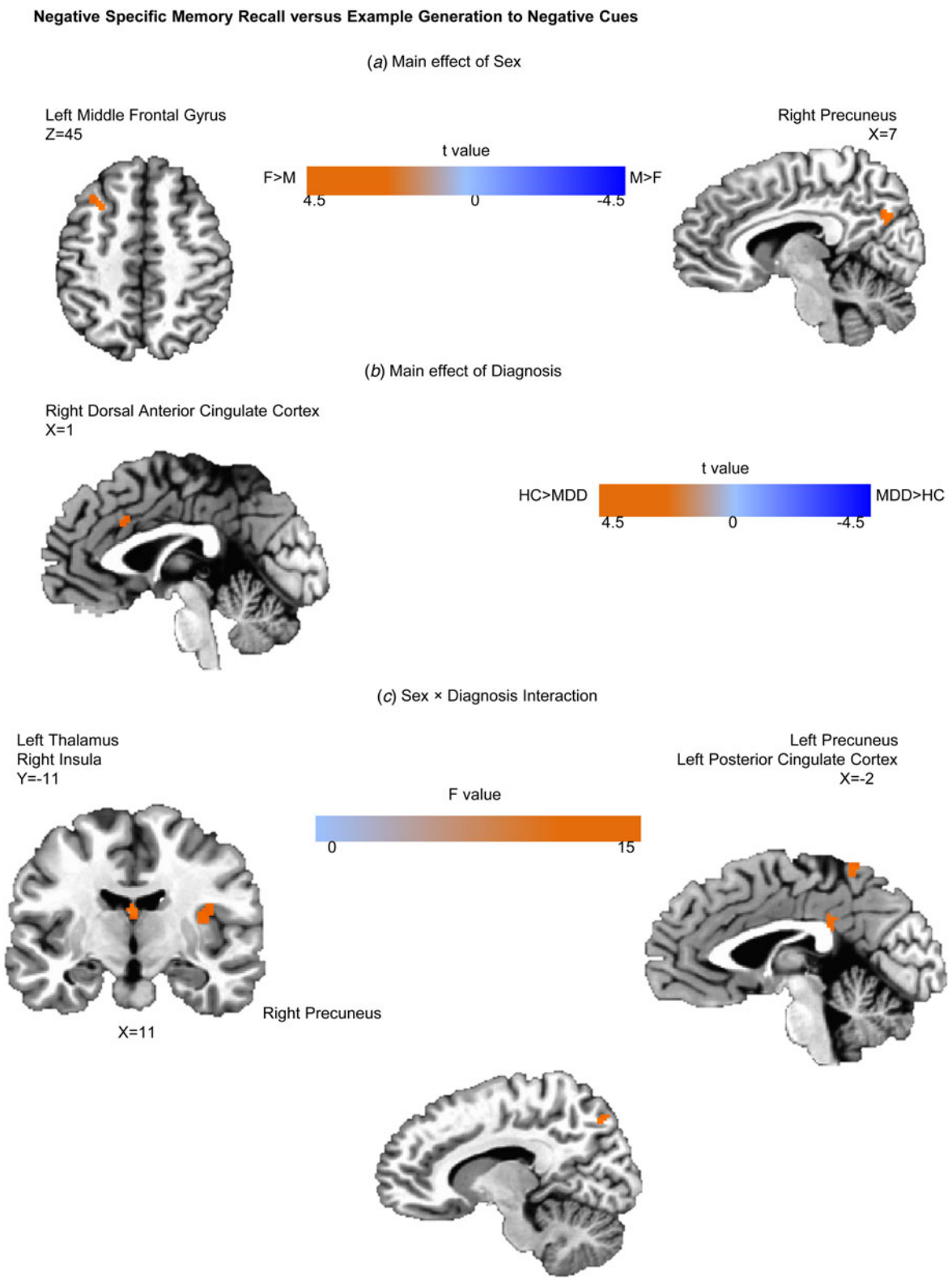

Fig. 3. Brain Regions showing a sex × diagnosis Interaction in Hemodynamic Activity During Negative Autobiographical Memory Recall. (*a*) Main effect of sex (*b*) main effect of diagnosis and (*c*) sex × diagnosis interaction for hemodynamic activity during negative specific memory recall *v*. example generation in response to negative cues ($p_{\text{corrected}} < 0.05$).

https://doi.org/10.1017/S003329171700112X Published online by Cambridge University Press

activity (Banich *et al.* 2009). In the current study, parahippocampal activity was higher in males than females, regardless of diagnosis, during positive AM recall. The parahippocampus is implicated in visuospatial processing and recalling contextual details (Aminoff *et al.* 2013). These results appear consistent with previous studies of sex differences in brain regions underlying AM recall that have reported increased hippocampal/parahippocampal activity in males relative to females (Piefke *et al.* 2005; St Jacques *et al.* 2011) and increased DLPFC activity in females relative to males (Piefke *et al.* 2005; Young *et al.* 2013*c*).

The mPFC, along with the precuneus, showed significant sex × diagnosis interactions during positive specific memory recall. While overall, males had more activity in these regions than females, depressed females exhibited a negative BOLD response lower than that of healthy females. Both regions have been implicated in self-referential processing (Northoff *et al.* 2006; Cavanna & Trimble, 2006) suggesting females find positive memories less self-relevant than males, with depressed females finding these memories even less self-relevant than healthy females.

Finally, the caudate was active during positive specific memory recall for all groups except MDD females, who showed a negative BOLD response. The caudate has been implicated in reward processing (Delgado *et al.* 2000), and found to be under responsive to receipt of reward in MDD (Pizzagalli *et al.* 2009). MDD males, however, were statistically indistinguishable from HCs in caudate activity.

The differential brain activity during positive specific memory recall suggests that males, regardless of diagnosis, recruit brain regions involved in the use of visuospatial strategies during positive AM recall to a greater extent than females. Furthermore, HCs recruit regions involved in self-referential processing to a greater extent than MDDs when recalling positive specific memories. However, depressed males are able to recruit these regions to a greater extent than depressed females, suggesting depressed males still process positive memories in a manner similar to their healthy counterparts.

When examining negative specific memory recall, the opposite pattern of hemodynamic activity emerged. Overall, females, regardless of diagnosis, and depressed patients, regardless of sex, had increased activity in regions implicated in self-referential (mPFC, precuneus; Cavanna & Trimble, 2006; Northoff *et al.* 2006) and salience processing (dACC; Seeley *et al.* 2007). However, activity was greatest in depressed females, with healthy females showing less activity than depressed females, but still greater activity than males. Our previous work in currently and formerly depressed participants using the AM task also found increased activity in regions implicated in self-referential processing (including the precuneus) relative to HCs during negative AM recall (Young *et al.* 2014, 2016). These results raise the possibility that previously reported differences between HCs and MDDs are largely driven by female participants, and support the inclusion of sex as an important variable to include in future studies using neuroimaging to differentiate depressed and healthy individuals.

The insula, a region involved in processing salient stimuli (Menon & Uddin, 2010), was also more active in MDD females than all other groups during negative specific memory recall. This finding suggests that for MDD females, specific negative memories are more salient, and therefore more likely to be attended to (Pauli & Roder, 2008) and mood relevant (Bower, 1981).

Finally, the thalamus was more active in MDD females than all other groups during recall of negative specific memories, with a negative BOLD response observed in MDD males. Increased thalamus engagement along with regions involved in self-referential processing has been suggested to underlie increased rumination and inability to suppress attention towards negative emotional states in MDD (Hamilton *et al.* 2015). As MDD females also engaged several regions involved in self-referential processing in addition to the thalamus during negative AM recall relative to all other groups, MDD females may engage in ruminative and self-referential processes to a greater extent than MDD males and HCs during negative AM recall, and this may be a factor contributing to the increased prevalence of MDD in females.

Differential activity during negative specific memory recall suggests females may find negative AMs more self-relevant and ruminated upon than in males. Particularly important for the present study, the PCC and mPFC have been linked to rumination and negative self-referential processing (Lemogne *et al.* 2011). That MDD females engage these regions to a greater extent than even HC females is consistent with the likelihood that rumination and negative self-focus is more prevalent in the depressed state. This hypothesis is speculative, however, as we did not collect data on ruminations or memory self-relevance.

The difference in how specific AMs are recalled may contribute to sex differences in MDD prevalence. While both males and females have the same cognitive deficit (overgeneral AM recall), only in females has this deficit been related to depressive symptoms (Ros *et al.* 2014), a finding confirmed in the current study. In many regions found to differentiate males and females with MDD, HC, and MDD males did not differ from each other, while MDD females differed from both HC males and females. These results suggest that MDD

https://doi.org/10.1017/S003329171700112X Published online by Cambridge University Press

males recall emotionally valenced memories in a similar manner as their healthy counterparts.

Our data further suggest that depression may exaggerate sex differences in regional brain activity that occur normally during AM recall. Furthermore, it might be hypothesized that the sex differences during AM recall may contribute to pathophysiology of mood disorders as part of a 'two-hit' model whereby the differential brain activity evident in healthy females adds vulnerability under stress exposure, which in turn exaggerates these differences leading to the development of depressive symptoms in females. Longitudinal studies are needed to test this hypothesis of how brain differences in AM recall actually influence sex differences in risk or expression of MDD in females.

Several limitations of the study merit comment. Firstly, the inclusion of only unmedicated individuals limits the generalizability of our findings. However, the medication exclusion was deemed necessary to detect potential differences in emotion networks, as psychotropic medications influence limbic and prefrontal hemodynamic activity (Wang *et al.* 2008), and may mask important group differences, especially given that females are more likely to use and respond to antidepressant medications (Khan *et al.* 2005; Pratt *et al.* 2011). Furthermore, our HCs were free of any Axis I disorder, while our MDD sample could have co-morbid anxiety. Therefore, it is possible that results may be influenced by anxiety symptoms, and not simply depressed mood. Future studies that characterize the influence of anxiety symptoms are therefore warranted. Secondly, our sample size may have been underpowered to detect behavioral differences in the properties of the recalled memories. However, effect sizes of the behavioral differences were small in most cases, supporting the conclusion that depressed males and females do not differ in the measured phenomenological properties of AM recall. Thirdly, we only measured a small number of phenomenological properties and many of our conclusions are based on reverse inference where we make assumptions that differential neural activity reflects differences in constructs (such as rumination, effort, and self-focus) that were not directly measured. We consider these interpretations hypotheses that can be directly tested in future AM recall studies by incorporating momentary measures of such constructs as rumination, self *v.* other focus, and effort. Finally, due to the nature of AM recall, it is impossible to control the number of memories participants recalled in each mnemonic and valence category beyond the systematic use of cue words. Future studies could address this limitation by either increasing the number of cue words used or developing alternative methods for cueing AMs to elicit more balanced recall of specifically targeted AM types (i.e. positive specific), allowing for more memories to be generated in order to increase power to detect sex by diagnosis interactions.

In conclusion, sex must be carefully evaluated when conducting studies on MDD. Results suggest that many regions consistently found to differ between MDDs and HCs, including the mPFC and precuneus, may be driven by MDD females. As many studies have more depressed females than males, efforts should be made to have a sex-balanced MDD sample and evaluate sex effects within the sample. This research will provide enhanced understanding of the heterogeneity of MDD and lead to better targets for individualized or sex-based treatments for depression. Our results suggest that reducing rumination on negative AMs, while increasing self-referential processing and the reward value of positive AMs, may be an effective strategy at reducing depression in females.

Supplementary material

The supplementary material for this article can be found at https://doi.org/10.1017/S003329171700112X.

Acknowledgements

This research was supported by the Laureate Institute for Brain Research and The William K. Warren Foundation. The funders had no influence on the design or conduct of the study, collection, management, analyses, or interpretation of the data, or in the preparation, review or approval of the manuscript.

Declaration of Interest

WCD is currently an employee of Janssen Research & Development, LLC, of Johnson & Johnson, Inc., and also holds equity in Johnson & Johnson. The other authors have no financial conflicts of interest or disclosures to declare.

Ethical Standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

American Psychological Association (APA) (2000). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*. American Psychiatric Association: Washington, DC.

Aminoff EM, Kveraga K, Bar M (2013). The role of the parahippocampal cortex in cognition. *Trends in Cognitive Sciences* **17**, 379–390.

Banich MT, Mackiewicz KL, Depue BE, Whitmer AJ, Miller GA, Heller W (2009). Cognitive control mechanisms, emotion and memory: a neural perspective with implications for psychopathology. *Neuroscience Biobehavioral Reviews* **33**, 613–630.

Bower GH (1981). Mood and memory. *American Psychologist* **36**, 129–148.

Briceno EM, Rapport LJ, Kassel MT, Bieliauskas LA, Zubieta JK, Weisenbach SL, Langenecker SA (2015). Age and gender modulate the neural circuitry supporting facial emotion processing in adults with major depressive disorder. *American Journal of Geriatric Psychiatry* **23**, 304–313.

Brittlebank AD, Scott J, Williams JM, Ferrier IN (1993). Autobiographical memory in depression: state or trait marker? *British Journal of Psychiatry* **162**, 118–121.

Buckner RL, Andrews-Hanna JR, Schacter DL (2008). The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences* **1124**, 1–38.

Cavanna AE, Trimble MR (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* **129**, 564–583.

Champagne K, Burkhouse KL, Woody ML, Feurer C, Sosoo E, Gibb BE (2016). Brief report: overgeneral autobiographical memory in adolescent major depressive disorder. *Journal of Adolescence* **52**, 72–75.

Davis M, Whalen PJ (2001). The amygdala: vigilance and emotion. *Molecular Psychiatry* **6**, 13–34.

Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology* **84**, 3072–3077.

First MB, Spitzer RL, Gibbon M, Williams JBW (2002). *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition* (*SCID-I/P*). New York State Psychiatric Institute, Biometrics Research: New York, NY.

Griffith JW, Sumner JA, Debeer E, Raes F, Hermans D, Mineka S, Zinbarg RE, Craske MG (2009). An item response theory/confirmatory factor analysis of the Autobiographical Memory Test. *Memory* **17**, 609–623.

Hamilton JP, Farmer M, Fogelman P, Gotlib IH (2015). Depressive rumination, the default-mode network, and the dark matter of clinical neuroscience. *Biological Psychiatry* **78**, 224–230.

Hamilton M (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry* **23**, 56–62.

Heron J, Crane C, Gunnell D, Lewis G, Evans J, Williams JM (2012). 40 000 memories in young teenagers: psychometric properties of the Autobiographical Memory Test in a UK cohort study. *Memory* **20**, 300–320.

Kessler RC (2003). Epidemiology of women and depression. *Journal of Affective Disorders* **74**, 5–13.

Khan A, Brodhead AE, Schwartz KA, Kolts RL, Brown WA (2005). Sex differences in antidepressant response in recent antidepressant clinical trials. *Journal of Clinical Psychopharmacology* **25**, 318–324.

Lemogne C, Gorwood P, Bergouignan L, Pelissolo A, Lehericy S, Fossati P (2011). Negative affectivity, self-referential processing and the cortical midline structures. *Social Cognitive and Affective Neuroscience* **6**, 426–433.

Levesque J, Eugene F, Joanette Y, Paquette V, Mensour B, Beaudoin G, Leroux JM, Bourgouin P, Beauregard M (2003). Neural circuitry underlying voluntary suppression of sadness. *Biological Psychiatry* **53**, 502–510.

Menon V, Uddin LQ (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function* **214**, 655–667.

Nejad AB, Fossati P, Lemogne C (2013). Self-referential processing, rumination, and cortical midline structures in major depression. *Frontiers in Human Neuroscience* **7**, 666.

Nolen-Hoeksema S (1987). Sex differences in unipolar depression: evidence and theory. *Psychological Bulletin* **101**, 259–282.

Northoff G, Heinzel A, de Greck M, Bermpohl F, Dobrowolny H, Panksepp J (2006). Self-referential processing in our brain--a meta-analysis of imaging studies on the self. *Neuroimage* **31**, 440–457.

Pauli WM, Roder B (2008). Emotional salience changes the focus of spatial attention. *Brain Research* **1214**, 94–104.

Piccinelli M, Wilkinson G (2000). Gender differences in depression. Critical review. *British Journal of Psychiatry* **177**, 486–492.

Piefke M, Weiss P, Markowitsch H, Fink G (2005). Gender differences in the functional neuroanatomy of emotional episodic autobiographical memory. *Human Brain Mapping* **24**, 313–324.

Pizzagalli DA, Holmes AJ, Dillon DG, Goetz EL, Birk JL, Bogdan R, Dougherty DD, Iosifescu DV, Rauch SL, Fava M (2009). Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder. *American Journal of Psychiatry* **166**, 702–710.

Pratt LA, Brody DJ, Gu Q (2011). Antidepressant use in persons aged 12 and over: United States, 2005–2008. NCHS data brief, no 76. National Center for Health Statistics: Hyattsville, MD.

Rawal A, Rice F (2012). Examining overgeneral autobiographical memory as a risk factor for adolescent depression. *Journal of the American Academy of Child and Adolescent Psychiatry* **51**, 518–527.

Ros L, Ricarte JJ, Serrano JP, Nieto M, Aguilar MJ, Latorre JM (2014). Overgeneral autobiographical memories: gender differences in depression. *Applied Cognitive Psychology* **28**, 472–480.

Ross M, Holmberg D (1992). Are wives' memories for events in relationships more vivid than their husbands' memories? *Journal of Social and Personal Relationships* **9**, 585–604.

Rubin DC, Schulkind MD, Rahhal TA (1999). A study of gender differences in autobiographical memory: broken down by age and sex. *Journal of Adult Development* **6**, 61–71.

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience* **27**, 2349–2356.

Seidlitz L, Diener E (1998). Sex differences in the recall of affective experiences. *Journal of Personality and Social Psychology* **74**, 262–271.

Soderlund H, Moscovitch M, Kumar N, Daskalakis ZJ, Flint A, Herrmann N, Levine B (2014). Autobiographical episodic memory in major depressive disorder. *Journal of Abnormal Psychology* **123**, 51–60.

St Jacques PL, Conway MA, Cabeza R (2011). Gender differences in autobiographical memory for everyday events: retrieval elicited by SenseCam images versus verbal cues. *Memory* **19**, 723–732.

Stokes DJ, Dritschel BH, Bekerian DA (2004). The effect of burn injury on adolescents autobiographical memory. *Behaviour Research and Therapy* **42**, 1357–1365.

Talairach J, Tournoux P (1988). *Co-planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System - An Approach to Cerebral Imaging*. Thieme Medical Publishers: New York.

Wang L, LaBar K, Smoski M, Rosenthal M, Dolcos F, Lynch T, Krishnan R, McCarthy G (2008). Prefrontal mechanisms for executive control over emotional distraction are altered in major depression. *Psychiatry Research* **163**, 143–155.

Watkins E, Teasdale JD (2001). Rumination and overgeneral memory in depression: effects of self-focus and analytic thinking. *Journal of Abnormal Psychology* **110**, 353–357.

Wechsler D (1999). *Wechsler Abbreviated Scale of Intelligence* (*WASI*). Harcourt Assessment: San Antonio, TX.

Williams J, Broadbent K (1986). Autobiographical memory in suicide attempters. *Journal of Abnormal Psychology* **95**, 144–149.

Williams JM, Barnhofer T, Crane C, Herman D, Raes F, Watkins E, Dalgleish T (2007). Autobiographical memory specificity and emotional disorder. *Psychological Bulletin* **133**, 122–148.

Young KD, Bellgowan P, Bodurka J, Drevets WC (2013*a*). Behavioral and neurophysiological correlates of autobiographical memory deficits in patients with depression and individuals at high risk for depression. *Journal of the American Medical Association: Psychiatry* **70**, 698–708.

Young KD, Bellgowan P, Bodurka J, Drevets WC (2013*b*). Functional neuroimaging of sex differences in autobiographical memory recall. *Human Brain Mappping* **34**, 3320–3332.

Young KD, Bellgowan P, Bodurka J, Drevets WC (2014). Neurophysiological correlates of autobiographical memory deficits in currently and formerly depressed subjects. *Psychological Medicine* **44**, 2951–2963.

Young KD, Bellgowan P, Bodurka J, Drevets WC (2015). Functional neuroimaging correlates of autobiographical memory deficits in subjects at risk for depression. *Brain Sciences* **5**, 144–164.

Young KD, Siegle GJ, Bodurka J, Drevets WC (2016). Amygdala activity during autobiographical memory recall in depressed and vulnerable individuals; Association with symptom severity and autobiographical overgenerality. *American Journal of Psychiatry* **173**, 78–89.