

# Weak reward source memory in depression reflects blunted activation of VTA/SN and parahippocampus

Daniel G. Dillon,<sup>1</sup> Ian G. Dobbins,<sup>2</sup> and Diego A. Pizzagalli<sup>1</sup>

<sup>1</sup>Center for Depression, Anxiety and Stress Research & McLean Imaging Center, McLean Hospital, Harvard Medical School, Belmont, MA 02478, USA and <sup>2</sup>Department of Psychology, Washington University in Saint Louis, Saint Louis, MO 63130, USA

**Reward responses in the medial temporal lobes and dopaminergic midbrain boost episodic memory formation in healthy adults, and weak memory for emotionally positive material in depression suggests this mechanism may be dysfunctional in major depressive disorder (MDD). To test this hypothesis, we performed a study in which unmedicated adults with MDD and healthy controls encoded drawings paired with reward or zero tokens during functional magnetic resonance imaging. In a recognition test, participants judged whether drawings were previously associated with the reward token ('reward source') or the zero token ('zero source'). Unlike controls, depressed participants failed to show better memory for drawings from the reward source vs the zero source. Consistent with predictions, controls also showed a stronger encoding response to reward tokens vs zero tokens in the right parahippocampus and dopaminergic midbrain, whereas the MDD group showed the opposite pattern—stronger responses to zero vs reward tokens—in these regions. Differential activation of the dopaminergic midbrain by reward vs zero tokens was positively correlated with the reward source memory advantage in controls, but not depressed participants. These data suggest that weaker memory for positive material in depression reflects blunted encoding responses in the dopaminergic midbrain and medial temporal lobes.**

**Keywords:** depression; memory; reward; fMRI

## INTRODUCTION

Major depressive disorder (MDD) is characterized by anhedonia, the loss of interest in pleasurable activities (American Psychiatric Association, 2000). Translational research has advanced our understanding of anhedonia by linking it to research on brain reward systems in healthy humans and non-human animals. In particular, the identification of dopamine circuits that support reinforcement learning (Schultz, 1998) and motivated responding (Niv *et al.*, 2007) provides a foundation for the neuroscientific study of anhedonia (Dichter *et al.*, 2009; Pizzagalli *et al.*, 2009; Treadway and Zald, 2011). Here, we extend this approach to investigate the neurobiology of a reliable behavioral phenomenon: weak memory for positive material in depression.

Several studies document better memory for positive vs neutral or negative material in healthy adults, but not their depressed counterparts. For example, Dunbar and Lishman (1984) presented depressed inpatients and healthy controls with positive, negative and neutral words. Thirty minutes later, controls showed better memory for positive vs neutral and negative words, whereas the depressed group showed better memory for negative vs positive words. Furthermore, compared with controls, the depressed group showed worse accuracy for positive words but better accuracy for negative words. Thus, depression blunted memory for positive material while enhancing memory for negative material.

Similarly, Hamilton and Gotlib (2008) presented emotional pictures to controls and adults with MDD. One week later, controls showed better memory for positive vs negative pictures, whereas the MDD

group showed better memory for negative pictures than controls. Enhanced memory for negative material in depression was linked to stronger amygdala activation at encoding, but it was unclear why controls remembered positive pictures better than negative pictures, or why this positive memory effect was absent in MDD.

Several functional magnetic resonance imaging (fMRI) studies in healthy adults have suggested an answer to these questions by demonstrating that rewarding stimuli (e.g. monetary incentives) boost episodic memory formation via activation of the dopaminergic midbrain and medial temporal lobe (MTL) regions, such as the hippocampus and parahippocampus. For example, in a recent study, healthy volunteers studied objects preceded by high-value or low-value cues indicating how much money could be earned for remembering the objects in an upcoming test (Wolosin *et al.*, 2012). Memory accuracy was better for objects paired with high-value vs low-value cues, and *Memory* × *Reward* interactions were observed in the right parahippocampus and the ventral tegmental area/substantia nigra (VTA/SN), which comprises the dopaminergic midbrain. These interactions were driven by greater activation for subsequently remembered (vs forgotten) objects presented after the high-value cues, demonstrating that the monetary incentives boosted encoding activation (see also Adcock *et al.*, 2006; Wittmann *et al.*, 2008).

These results suggest that weak memory for positive material in MDD may reflect blunted encoding responses in the VTA/SN and MTL. To test this hypothesis, we performed an fMRI study in which depressed adults and healthy controls encoded drawings followed by reward or zero tokens. For brevity, we refer to drawings paired with reward and zero tokens as coming from the 'reward source' and 'zero source', respectively. A recognition test immediately followed. We predicted that controls—but not depressed participants—would show better memory for drawings from the reward source vs the zero source. Moreover, we expected this effect to be mirrored by differential encoding activation in the VTA/SN and MTL, with these regions showing stronger responses to reward (vs zero) tokens in controls, but not depressed participants. Similarly, we expected VTA/SN and MTL activation to correlate with memory accuracy in controls, but not depressed participants.

Received 7 May 2013; Revised 31 July 2013; Accepted 23 September 2013

Advance Access publication 26 September 2013

The authors gratefully acknowledge Nancy Hall Brooks, Andrew Cohen, Sunny Dutra and Laura Murray for their assistance with data collection. Portions of this work were presented as a poster at the 67th Annual Meeting of the Society of Biological Psychiatry (May 3–5, 2012, Philadelphia, PA, USA). This work was supported by the Brain & Behavior Research Foundation (NARSAD Young Investigator grant awarded to D.G.D.) and by the National Institute of Mental Health (grant numbers K99MH094438-01A1 to D.G.D. and R01MH68376 to D.A.P.). I.G.D. was supported by R01MH073982.

Correspondence should be addressed to Daniel G. Dillon, Center for Depression, Anxiety and Stress Research, McLean Hospital, 115 Mill Street, Belmont, MA 02478, USA. E-mail: ddillon@mclean.harvard.edu

## MATERIALS AND METHODS

### Participants

Results are from 21 unmedicated individuals diagnosed with MDD and 21 healthy controls (for details on diagnostic interviews used to establish eligibility, see the [Supplementary Material](#)). Additional data were excluded due to excessive head motion (>4 mm or degrees incremental: controls,  $n=4$ ; MDD,  $n=1$ ), failure to respond consistently during encoding (>14 encoding trials with no response: controls,  $n=1$ ; MDD,  $n=2$ ) and disclosure of antidepressant use (MDD,  $n=1$ ). Participants were compensated \$25/h for the fMRI session. All procedures were approved by the Harvard Committee on the Use of Human Subjects in Research and the McLean Hospital Institutional Review Board.

### Self-report measures

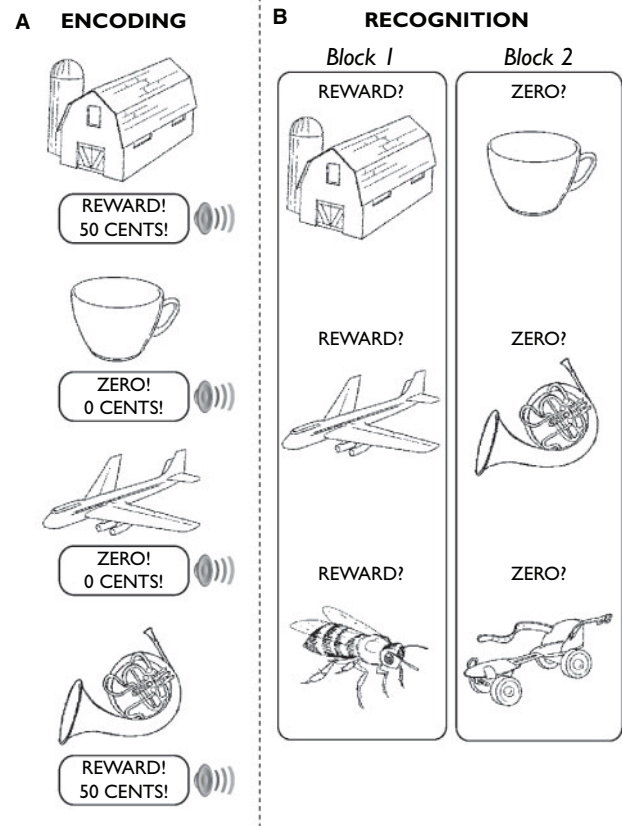
Participants completed the Beck Depression Inventory-II (BDI-II; [Beck et al., 1996](#)), Mood and Anxiety Symptoms Questionnaire (MASQ; [Watson et al., 1995](#)) and the Snaith-Hamilton Pleasure Scale (SHAPS; [Snaith et al., 1995](#)). The Wechsler Test of Adult Reading (WTAR; [The Psychological Corporation, 2001](#)) served as a brief IQ assessment. Total scores were computed for the BDI-II and SHAPS, and WTAR scores were standardized using age norms. Four scale scores were computed for the MASQ: General Distress-Depression (MASQ-GDD), Anhedonic Depression (MASQ-AD), General Distress-Anxiety (MASQ-GDA) and Anxious Arousal (MASQ-AA). Group differences on all measures were assessed by unpaired  $t$ -test.

### Memory task

#### Encoding

During encoding, participants studied drawing-token associations ([Figure 1A](#)). There were 80 trials, divided over two runs. Each trial included a unique line drawing (3 s duration) ([Snodgrass and Vanderwart, 1980](#)), a jittered inter-stimulus interval (ISI: 2–5 s), a reward or zero token (1 s duration, 40 trials each) and a jittered inter-trial interval (ITI: 2–14 s). The reward token was a screen that read 'REWARD! 50 Cents!' accompanied by a chime sound. The zero token was a screen that read 'ZERO! 0 Cents!' accompanied by a tone. To ensure that participants remained attentive throughout encoding, they were asked to press a button when each drawing appeared. They were also instructed to try to remember whether each drawing was followed by the reward or zero token. The drawing-token pairings were arbitrary, and participants' button presses did not affect token delivery. Trial sequences and ITI lengths were determined with optseq ([Dale, 1999](#)). No more than three reward or zero tokens were presented sequentially, and a fixation cross was visible during the ISI and ITI. The task was programmed in E-Prime (Psychology Software Tools, Inc., Sharpsburg, PA, USA).

We use the term 'token' because participants did not receive additional money for the reward tokens, but were instructed to think of them as representing a portion of their fixed payment. Participants knew they would not receive additional money for the reward tokens, and that payment was not influenced by performance. We decoupled the tokens from payment because, while we wanted the reward tokens to elicit positive emotional responses, we also wanted all drawing-token associations to be equally important with respect to the memory test. By emphasizing that participants should try to remember all drawing-token associations—not just the drawings paired with reward tokens—and by making it clear that the reward tokens would not yield additional pay, we sought to minimize the possibility that group differences in reward source memory would emerge because of differences in motivation to earn extra money.



**Fig. 1** Experimental design. (A) During encoding, participants studied drawing-token associations. Half of the drawings were followed by a reward token ('reward source'), half were followed by a zero token ('zero source'). (B) At recognition, old and new drawings were presented in two blocks that featured either a 'Reward?' cue or a 'Zero?' cue. These cues prompted participants to make a source memory judgment for each drawing. A three-point confidence rating scale (low, medium and high) followed each drawing (not shown). Sixty drawings were tested under each cue (20 reward source, 20 zero source and 20 new).

Note that this approach differs from methods used in prior studies to show that the prospect of increased pay can improve memory (e.g. [Adcock et al., 2006](#)). In those studies, participants viewed cues indicating that remembering certain items would yield larger payments than remembering other items. In addition, most prior studies placed motivational cues *before* the items to be remembered, whereas we placed the reward and zero tokens *after* the drawings. This increases the likelihood that participants will attend carefully to each drawing, and minimizes the possibility that differences in source memory would be driven by differences in participants' attention to the drawings (because participants could not know to which source a drawing belonged until the reward or zero token was presented).

#### Recognition

The recognition test directly followed encoding ([Figure 1B](#)). There were 120 trials, which included 40 drawings from the reward source, 40 from the zero source and 40 new drawings. In one block, drawings (4 s duration) appeared under a 'REWARD?' cue and participants pressed a button for 'yes' (this is from the reward source) or 'no' (this is not from the reward source). In another block, drawings appeared under a 'ZERO?' cue and participants pressed a button for 'yes' (this is from the zero source) or 'no' (this is not from the zero source). Sixty drawings (20 reward source, 20 zero source and 20 new) were tested under each cue, and assignment of cues to the first and second blocks was counterbalanced across participants. A screen reading

'Confidence?' (2 s duration) followed each drawing, and participants pressed one of three buttons to indicate their level of confidence (low, medium and high) in each memory judgment. Jittered intervals (3–15 s) separated the trials.

The recognition cue manipulation investigated whether directing participants' attention toward or away from the reward source during retrieval would affect accuracy (March and Hicks, 1998). We predicted that controls—but not depressed participants—would show better reward source accuracy when responding to the 'REWARD?' cue vs the 'ZERO?' cue. This hypothesis was not supported. A preliminary  $Group \times Cue \times Source$  (reward, zero and new) ANOVA on accuracy data yielded a significant  $Group \times Source$  interaction but no significant interactions between *Cue* and either *Group* or *Source* (see [Supplementary Material](#)). Furthermore, there was no group difference in the percentage of correctly rejected new drawings (controls:  $86.18 \pm 15.36$ ; MDD:  $81.53 \pm 17.83$ ;  $t < 1$ ,  $P = 0.37$ ). Therefore, to focus on the  $Group \times Source$  interaction, we present accuracy and RT data collapsed across the recognition cues and omit responses to new drawings from further analysis. Importantly, we verified that group differences in source memory did not reflect response bias effects (see [Supplementary Material](#)). Because we hypothesized that group differences in source memory would reflect negative effects of MDD on encoding responses to reward tokens, we only present fMRI data from encoding.

### Post-task ratings

Following the memory task, the reward and zero tokens were presented once more with their accompanying sounds, and participants used five-point scales to rate each token for arousal (1 = low arousal, 5 = high arousal) and valence (1 = very unpleasant, 3 = neutral, 5 = very pleasant). Participants rated their current, 'in the moment' emotional response to each token, and did not retrospectively rate their responses during the encoding runs. No fMRI data were collected during these brief ratings, which were made in the scanner.

### fMRI acquisition

MRI data were collected on a 3 T magnet (Siemens, USA; 12-channel head coil). The MRI sessions included an auto-align localizer (van der Kouwe et al., 2005), a T1-weighted multi-echo magnetization prepared rapid acquisition gradient echo (MEMPRAGE; van der Kouwe et al., 2008) structural image [ $1.2 \text{ mm}^3$  voxels; 144 slices; repetition time (TR) = 2.2 s; echo time (TE) $1/2/3/4 = 1.54/3.36/5.18/7.01$  ms] and T2\*-weighted images sensitive to blood oxygen level-dependent contrast, acquired during the memory task ( $3.0 \text{ mm}^3$  voxels; 46 slices; TR = 3 s; TE = 30 ms; flip angle =  $85^\circ$ ; transverse acquisition).

### Behavioral data analysis

#### Post-task ratings

Arousal and valence ratings were entered in  $Group \times Token$  ANOVAs. *Gender* was included in all ANOVAs conducted on behavioral data, but it did not interact with any variable of interest and was removed to conserve degrees of freedom. Alpha was set to 0.05.

#### Encoding

Between-group *t*-tests compared the number and median RT of encoding responses. Median data were used in RT analyses to minimize the influence of extreme values.

#### Source accuracy

A  $Group \times Source$  (reward, zero) ANOVA was conducted on percent correct data. Reward minus zero source accuracy difference scores were

also computed for each participant and entered in a between-group *t*-test.

### Recognition RT

Median RTs were entered in a  $Group \times Source$  ANOVA.

### Confidence

Mean confidence ratings were entered in a  $Group \times Source$  ANOVA.

### Source accuracy regression

To determine if *Group* predicted reward minus zero source accuracy scores after accounting for variables unspecific to depression, a linear regression tested for effects of *Group* (entered second: control = 0, MDD = 1) after accounting for age, education, WTAR, MASQ-GDA and MASQ-AA scores (entered first).

### fMRI data analysis

#### Pre-processing

Five volumes from the start of each run were discarded for stabilization of longitudinal magnetization, and slice-time correction was accomplished with SPM2 (Wellcome Department of Cognitive Neurology, London, UK). Next, the data were corrected for head motion, normalized to MNI152 templates ( $2 \text{ mm}^3$  voxels) and spatially smoothed (6 mm full-width at half-maximum) using FSL software (Jenkinson et al., 2002; Smith, 2002).

#### Source memory analysis

SPM8 (Wellcome Department of Cognitive Neurology) was used for general linear modeling. Delta functions placed at drawing and token onset times were convolved with SPM's canonical hemodynamic response function, and nuisance regressors accounted for run-to-run fluctuations in mean image intensity. The data were high-pass filtered (cut-off period: 128 s).

The fMRI analysis was used to identify brain regions supporting a group difference in reward vs zero source memory. For each participant, we computed contrasts comparing the response to each token (reward, zero) vs fixation. For the purposes of the group-level ANOVA, we also binned the drawings according to which token (reward or zero) followed them, and then contrasted the response to drawings vs fixation. Finally, we entered the contrast images in a whole-brain, random effects  $Group \times Stimulus$  (drawing, token)  $\times Token Type$  (reward, zero) ANOVA. We expected  $Group \times Stimulus \times Token Type$  interactions in the MTL and VTA/SN, driven by stronger responses to reward vs zero tokens in controls vs depressed participants.

To strengthen the link to memory, this analysis was restricted to trials that led to successful remembering. There was no group difference in the number of correctly remembered drawings from the reward source [controls:  $29.29 \pm 6.75$ ; depressed:  $25.57 \pm 7.33$ ;  $t(40) = 1.71$ ,  $P = 0.10$ ] or the zero source [controls:  $24.48 \pm 5.35$ ; depressed:  $26.81 \pm 5.46$ ;  $t(40) = -1.40$ ,  $P = 0.17$ ]. Unfortunately, low trial counts precluded analysis of memory failures (i.e. 17 participants had 10 or fewer incorrect reward source responses).

The voxelwise *P*-value was set to 0.005. Based on prior work (e.g. Davachi et al., 2003), we expected group differences in reward source memory to reflect variation in parahippocampal or hippocampal encoding activation, in addition to the VTA/SN. However, we had no hypotheses distinguishing between these regions or the two hemispheres. Therefore, when evaluating the  $Group \times Stimulus \times Token Type$  interaction in the MTL, we corrected for multiple comparisons over a bilateral hippocampal/parahippocampal mask from the Automated Anatomical Labeling atlas (Tzourio-Mazoyer et al.,



2002). In the VTA/SN, we corrected for multiple comparisons over an 8 mm radius sphere centered on the MNI coordinates [2, -14, -16], derived from a high-density fMRI study of the VTA/SN (Krebs *et al.*, 2011; see [Supplementary Material](#)). For all other contrasts and brain regions, we corrected for multiple comparisons over the whole brain. Only activations whose peak activation was significant at  $P < 0.05$  (corrected) are reported. MarsBaR (Brett *et al.*, 2002) was used to extract beta weights for additional analysis, and MRICron (Rorden and Brett, 2000) was used for display.

## RESULTS

### Self-report data

As expected, the MDD group reported more severe symptoms of depression, anhedonia and anxiety than controls (Table 1). Based on the mean BDI-II score ( $23.62 \pm 7.17$ ), the MDD group was moderately depressed. Additional clinical data are presented in the [Supplementary Material](#).

### Behavioral data

#### Post-task ratings

The reward token was rated more pleasant [reward:  $4.29 \pm 0.94$ ; zero:  $2.05 \pm 1.21$ ; *Token*,  $F(1, 40) = 95.42$ ,  $P < 0.001$ ] and arousing [reward:  $3.67 \pm 0.98$ ; zero:  $2.76 \pm 1.41$ ; *Token*,  $F(1, 40) = 10.15$ ,  $P = 0.003$ ] than the zero token. There were no effects involving *Group* ( $P_s > 0.18$ ).

#### Encoding

There were no group differences in the number of on-time responses [controls:  $78.71 \pm 2.12$ ; MDD:  $76.81 \pm 4.09$ ;  $t(40) = 1.89$ ,  $P = 0.07$ ] or the median RT [controls:  $969.97 \pm 267.09$  ms; MDD:  $904.45 \pm 185.55$  ms;  $t(40) = 1.37$ ,  $P = 0.18$ ].

### Source memory accuracy

As predicted, there was a *Group*  $\times$  *Source* interaction,  $F(1, 40) = 6.29$ ,  $P = 0.02$  (Figure 2A). Controls showed better accuracy for drawings from the reward source *vs* the zero source,  $t(20) = 3.23$ ,  $P = 0.004$ ,  $d = 0.71$ . In contrast, depressed participants showed numerically better accuracy for drawings from the zero source *vs* the reward source, but this was unreliable,  $t(20) < 1$ ,  $P = 0.50$ .

A between-group *t*-test on reward source accuracy was not significant. However, as shown in Figure 2B, the mean reward minus zero source accuracy difference score was more positive in controls *vs* the MDD group,  $t(40) = 2.51$ ,  $P = 0.02$ ,  $d = 0.78$ . Furthermore, 17 of 21 controls (binomial  $P = 0.007$ ) but only 9 of 21 depressed participants (binomial  $P = 0.66$ ) showed a positive difference score,  $\chi^2_1 = 6.46$ ,  $P = 0.01$  (Figure 2C). Reward minus zero source accuracy scores did

not differ between depressed participants with *vs* without comorbid anxiety and were not correlated with BDI-II, SHAPS, MASQ-GDD or MASQ-AD scores in the MDD group.

### Confidence rating

No significant effects emerged ( $P_s > 0.08$ ).

### Recognition RT

The *Group*  $\times$  *Source* interaction was marginal,  $F(1, 40) = 3.08$ ,  $P = 0.09$  (Figure 1D). In controls, RTs were faster for drawings from the reward source *vs* the zero source,  $t(20) = 2.17$ ,  $P = 0.04$ . In contrast, in depressed participants RTs were numerically faster for drawings from the zero source *vs* the reward source, but this was unreliable,  $t(20) < 1$ ,  $P = 0.60$ . Depressed participants responded faster than controls for drawings from the zero source,  $t(40) = 2.17$ ,  $P = 0.04$ .

### Source accuracy regression

*Group* emerged as the only significant predictor ( $\beta = -0.64$ ,  $P = 0.02$ ) of reward minus zero source accuracy after accounting for age, education and WTAR, MASQ-GDA and MASQ-AA scores ( $P_s > 0.13$ ). Furthermore, *Group* explained unique variance above and beyond all variables entered in the first step of the model [ $\Delta R^2 = 0.13$ ,  $\Delta F(1, 35) = 6.06$ ,  $P = 0.02$ ]. Thus, the group difference in reward *vs* zero source accuracy was not explained by variables unspecific to depression.

### fMRI

A complete list of activations is provided in Table 2. Based on our *a priori* hypothesis, we focus on the response to reward and zero tokens in brain regions that showed a *Group*  $\times$  *Stimulus*  $\times$  *Token Type* interaction. See the [Supplementary Material](#) for additional results, including VTA/SN and right parahippocampal responses to drawings.

### Group $\times$ Stimulus $\times$ Token Type interactions

Triple interactions emerged in the VTA/SN (Figure 3A) and right parahippocampus (Figure 3B). To follow-up these effects, beta weights indexing responses to reward and zero tokens were extracted from both clusters and submitted to *Group*  $\times$  *Token Type* follow-up ANOVAs. Significant interactions were confirmed in the VTA/SN,  $F(1, 40) = 12.32$ ,  $P = 0.001$  and right parahippocampus,  $F(1, 40) = 11.40$ ,  $P = 0.002$ . Similar analyses focused on the response to drawings in these two brain regions revealed no significant differences between the groups (see [Supplementary Material](#)).

In the VTA/SN (Figure 3A), the response to reward tokens was stronger in controls *vs* depressed participants,  $t(40) = 2.34$ ,  $P = 0.02$ ,  $d = 0.72$ . In contrast, the response to zero tokens was marginally stronger in depressed participants *vs* controls,  $t(40) = -1.97$ ,  $P = 0.06$ ,  $d = 0.61$ . To parallel the memory analysis, we computed a difference score for each participant by subtracting the zero token beta weight from the reward token beta weight (Figure 3A). A between-group *t*-test on these difference scores was significant,  $t(40) = 3.51$ ,  $P = 0.001$ ,  $d = 1.11$ , as the mean reward minus zero difference score was more positive in controls. Finally, paired *t*-tests indicated that controls showed a marginally stronger response to reward *vs* zero tokens,  $t(20) = 1.85$ ,  $P = 0.08$ ,  $d = 0.41$ , while the depressed group showed the opposite result—a stronger response to zero *vs* reward tokens,  $t(20) = -3.68$ ,  $P = 0.001$ ,  $d = 0.81$ .

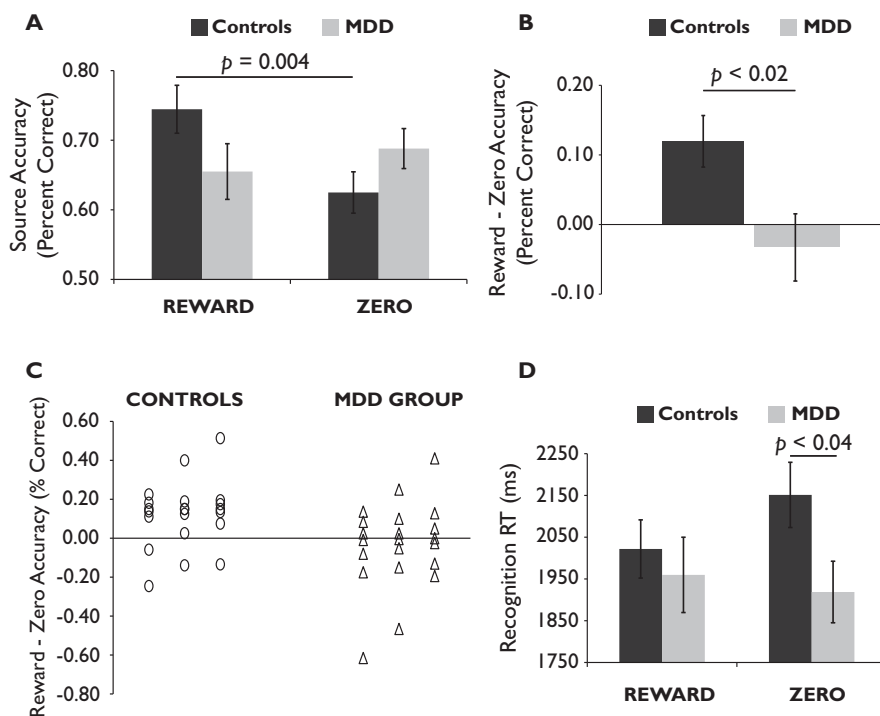
In the right parahippocampus, a slightly different pattern was observed. As shown in Figure 3B, there was no group difference in response to reward tokens,  $t(40) < 1$ ,  $P = 0.60$ ,  $d = 0.16$ . However, the depressed group responded more strongly to zero tokens than controls,

**Table 1** Demographics and mean (s.d.) self-report data

Variable	Controls	Depressed	<i>P</i>
Gender	9 f, 12 m	11 f, 10 m	0.54
Age	36.62 (13.32)	34.33 (12.16)	0.57
Education (years)	16.67 (2.06)	16.19 (2.34)	0.49
BDI-II	0.81 (1.40)	23.62 (7.17)	<0.001
MASQ-GDD	13.95 (2.84)	38.05 (9.98)	<0.001
MASQ-AD	43.71 (9.87)	83.00 (9.09)	<0.001
MASQ-GDA	12.71 (2.08)	24.19 (5.50)	<0.001
MASQ-AA	18.00 (1.61)	23.71 (6.86)	0.001
SHAPS	0.24 (0.54)	5.10 (3.75)	<0.001
WTAR <sup>a</sup>	116.68 (7.31)	113.11 (11.37)	0.26

f, female; m, male.

<sup>a</sup>WTAR data from non-native English speakers were not analyzed (controls,  $n = 2$ ; MDD,  $n = 2$ ).



**Fig. 2** (A) In contrast to the controls, the MDD group failed to show better accuracy for drawings from the reward source vs the zero source. (B) The mean reward minus zero source accuracy difference score was more positive in controls vs the MDD group. (C) More controls (17/21) than depressed participants (9/21) showed a positive reward minus zero source accuracy difference score. (D) In controls only, correct recognition decisions were rendered faster for drawings from the reward source vs the zero source. Compared with controls, the MDD group made faster correct recognition decisions for drawings from the zero source. Bars indicate standard error.

**Table 2** Activations in the Group  $\times$  Stimulus  $\times$  Token Type fMRI ANOVA

Region	x	y	z	Voxels	Z	FWE-corrected P-value
Group $\times$ Stimulus $\times$ Token Type						
Right parahippocampal gyrus <sup>a</sup>	26	-26	-24	32	3.95	0.035
VTA/SN <sup>b</sup>	-2	-16	-16	26	4.18	0.002
Group $\times$ Stimulus						
No significant activations						
Group $\times$ Token Type						
No significant activations						
Stimulus $\times$ Token Type						
No significant activations						
Main effect of stimulus						
Right Heschl's gyrus	54	-16	4	14730	Infinite	<0.001
Left superior temporal gyrus	-54	-24	2	24633	Infinite	<0.001
Left intracalcarine cortex	-10	-88	2	14632	Infinite	<0.001
Right occipital pole	32	-92	14	2864	Infinite	<0.001
Left precentral gyrus	-38	-22	68	2445	7.01	<0.001
Left central opercular cortex	-40	-2	12	166	6.60	<0.001
Left occipital pole	-30	-90	14	1418	6.10	<0.001
Main effect of token type						
Right Heschl's gyrus	52	-18	4	348	5.63	0.001
Left Heschl's gyrus/planum temporal	-40	-28	6	415	5.44	0.003
Right postcentral gyrus/white matter	30	-32	40	257	4.87	0.045
Main effect of group						
No significant activations						

Activations localized with the Harvard-Oxford cortical/subcortical structural atlases. Voxelwise  $P$ -value set at 0.005. Only clusters with a FWE-corrected  $P$ -value <0.05 are reported (peak-level). Multiple comparison correction over: <sup>a</sup>structurally defined bilateral hippocampal/parahippocampal mask; <sup>b</sup>8 mm radius sphere centered at [2, -14, -16], based on Krebs et al. (2011).

$t(40) = -3.81$ ,  $P < 0.001$ ,  $d = 1.18$ . As in the VTA/SN, reward minus zero difference scores were more positive in controls vs depressed participants,  $t(40) = 3.38$ ,  $P = 0.002$ ,  $d = 1.04$ . Finally, paired  $t$ -tests revealed that controls showed a stronger response to reward vs zero

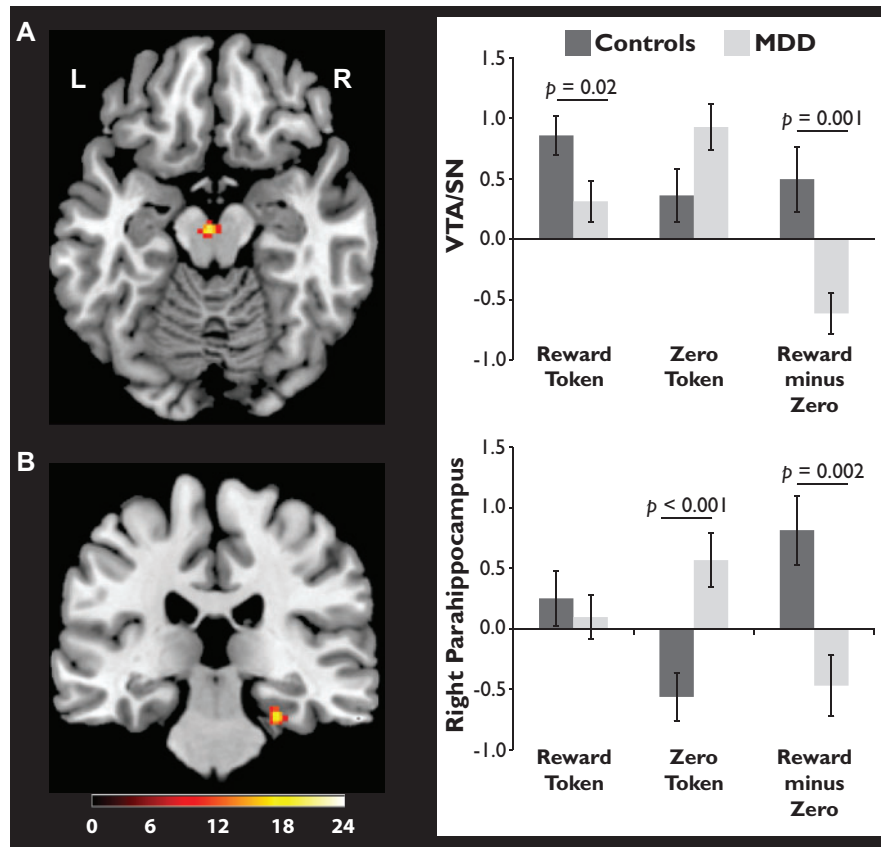
tokens,  $t(20) = 2.87$ ,  $P = 0.01$ ,  $d = 0.63$ , while the depressed group showed a marginally stronger response to zero vs reward tokens,  $t(20) = 1.86$ ,  $P = 0.08$ ,  $d = 0.41$ .

#### Relationship between VTA/SN and right parahippocampal activation

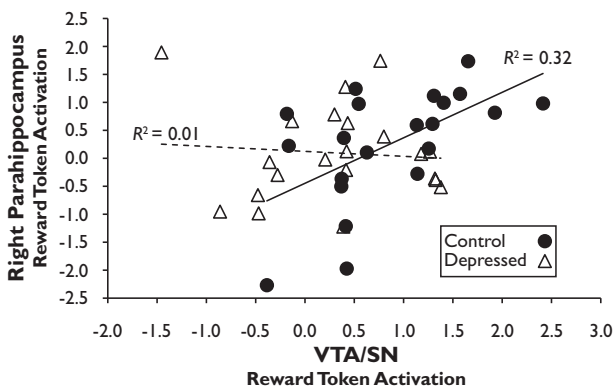
Prior work reported positive correlations between VTA/SN and MTL encoding responses in healthy volunteers (Shohamy and Wagner, 2008). To determine if this effect was present in our sample, we computed Pearson correlations between beta weights extracted from the VTA/SN and right parahippocampus. As shown in Figure 4, responses to reward tokens were positively correlated in controls,  $r = 0.57$ ,  $P < 0.01$ , but no relationship was seen in depressed participants,  $r = -0.08$ ,  $P = 0.73$ . Fisher's test for independent correlations confirmed a stronger relationship in controls vs depressed participants,  $Z = 2.18$ ,  $P = 0.03$ . Similar analyses of responses to zero tokens revealed no reliable correlations (controls:  $r = 0.29$ ,  $P = 0.20$ ; depressed:  $r = 0.24$ ,  $P = 0.30$ ).

#### Relationship between memory and activation of the VTA/SN and right parahippocampus

Finally, we correlated reward minus zero source accuracy with reward minus zero token difference scores in the VTA/SN (Figure 5A) and right parahippocampus (Figure 5B). One control was excluded from this analysis; see Supplementary Material for details. In controls, a significant correlation was seen in the VTA/SN ( $r = 0.69$ ,  $P = 0.001$ ) with a trend in the right parahippocampus ( $r = 0.42$ ,  $P = 0.07$ ). In depressed participants, no relationship was seen in either region (VTA/SN;  $r = 0.07$ ,  $P = 0.76$ ; right parahippocampus:  $r = -0.16$ ,  $P = 0.48$ ). The correlation with VTA/SN activation was stronger in controls vs depressed adults,  $Z = 2.30$ ,  $P = 0.02$ , with a trend in the right parahippocampus,  $Z = 1.84$ ,  $P = 0.07$ .



**Fig. 3** Group differences in the response to reward and zero tokens were observed in the (A) VTA/SN and (B) right parahippocampus. fMRI data are thresholded at  $P < 0.005$  (uncorrected); the scale bar corresponds to  $F$  values for the  $Group \times Stimulus \times Token Type$  interaction. Bar graphs (with standard error bars) show beta weights for reward and zero tokens vs fixation, plus the difference between them.



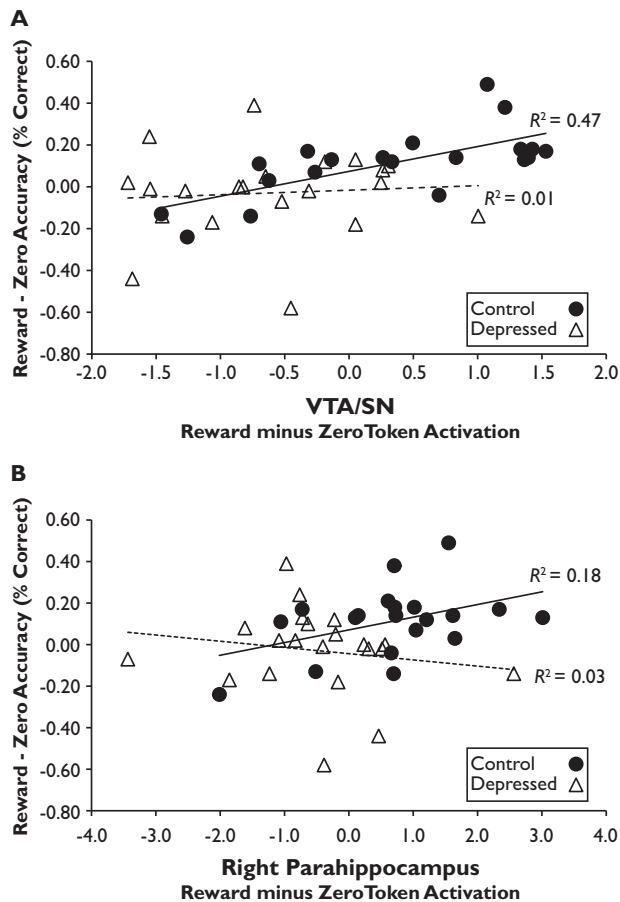
**Fig. 4** Relationship between responses to reward tokens in the VTA/SN and right parahippocampus. The correlation was significant in controls, but not depressed participants, and the strength of the correlation was significantly different between groups.

**DISCUSSION**

This study provides novel insight into weak memory for positive material in depression. In contrast to controls, the MDD group failed to show better memory for the reward source vs the zero source, yielding a group difference in reward minus zero source accuracy. This was not confounded by group differences in attention at encoding, as depressed participants responded as frequently and rapidly as controls. Furthermore, group differences in source memory were not driven by differential responses to the recognition cues, as the results held when responses to each recognition cue were considered separately,

and when  $d'$  was used instead of percent correct as the accuracy measure (see [Supplementary Material](#)). Instead, the group differences in source memory appear to reflect divergent encoding responses in the VTA/SN and right parahippocampus. In these regions, controls showed stronger responses to reward vs zero tokens, whereas depressed participants showed the opposite pattern: stronger responses to zero vs reward tokens. The consistency of this result across both regions is striking, although it should be noted that several within-group tests were marginally significant. Furthermore, the VTA/SN reward response was stronger in controls, and VTA/SN responses to the reward token (vs the zero token) were correlated with reward source (minus zero source) accuracy in controls, but not depressed participants. Collectively, these findings indicate that weak memory for positive material in depression reflects blunted encoding responses in the VTA/SN and MTL.

Stronger neural responses to zero tokens were seen in the MDD group. Valence ratings indicated that the zero tokens elicited unpleasant emotions, thus these results converge with reports of hypersensitivity to negative outcomes in depression ([Beats et al., 1996](#); [Holmes and Pizzagalli, 2008](#); [Beevers et al., 2012](#)), and echo prior findings of enhanced neural responses to negative material during encoding in MDD ([Hamilton and Gotlib, 2008](#)). Behaviorally, the MDD group showed a numerical advantage for zero vs reward source memory, but this was unreliable. This may reflect the fact that the zero token did not signal punishment; stronger findings would likely emerge if more intrinsically negative material had been used. Regardless, while this work highlights mechanisms supporting blunted memory for positive material in depression, the results also confirm increased sensitivity to non-rewarding outcomes in MDD.



**Fig. 5** Relationship between reward minus zero source accuracy and reward minus zero token difference scores in the (A) VTA/SN and (B) right parahippocampus. In controls, the correlation was significant in the VTA/SN and marginal in the right parahippocampus. In depressed participants, the correlation was not significant in either region. The strength of the correlations differed significantly between groups in the VTA/SN, with a marginal difference in the right parahippocampus.

#### What explains the group difference in VTA/SN and MTL responses?

The mechanism underlying group differences in VTA/SN and right parahippocampal responses is unclear, but differential activation of dopamine circuitry is a plausible candidate. Dopamine circuits linking the VTA/SN and MTL are believed to gate the entry of information into long-term memory, with dopamine bursts signaling which information is important enough to retain (Lisman and Grace, 2005). Along these lines, dopamine antagonists injected into the bilateral hippocampus of rodents prior to encoding render newly formed long-term memories unstable and susceptible to decay (Bethus *et al.*, 2010). Several lines of evidence suggest that depression may involve dopamine dysfunction (Treadway and Zald, 2011). Thus, relatively poor reward source memory in MDD may reflect weaker responses to reward tokens in VTA/SN dopamine cells that serve to stabilize memory representations in the MTL.

We speculate that the strong response to zero tokens in depressed participants may reflect activation of GABAergic cells in the VTA. In rodents, optogenetic methods have been used to identify separate populations of VTA GABAergic and dopaminergic neurons (Cohen *et al.*, 2012). While the dopaminergic neurons respond strongly to rewards and reward-predicting cues, the GABAergic neurons respond to aversive stimuli and inhibit dopamine firing. Thus, it is possible that the zero tokens may elicit stronger activation of these VTA cells in depressed *vs* healthy adults, which could translate into a stronger

response to zero tokens in downstream regions such as the parahippocampus.

Of course, these explanations are speculative because fMRI data cannot speak directly to dopaminergic or GABAergic function. Nevertheless, they are sensible given the VTA data from rodents (Cohen *et al.*, 2012), the framework linking dopamine-driven VTA/MTL interactions to episodic memory formation (Lisman and Grace, 2005), prior data from healthy volunteers in similar paradigms (Adcock *et al.*, 2006; Wittmann *et al.*, 2008; Wolosin *et al.*, 2012), and evidence of disrupted dopamine function in depression (Treadway and Zald, 2011). Therefore, it would be valuable to test these hypotheses using other methods capable of directly probing dopamine function in humans.

#### Clinical/translational implications

The results presented here are novel and require replication. Nonetheless, it is worth considering some translational implications. First, while a substantial literature already documents the fact that depressed individuals recall negative events with ease, these data suggest a complementary, more insidious effect of depressive illness: it makes encoding positive events more difficult. Thus, a depressed individual not only suffers from an abundance of negative memories but may also struggle to form positive memories, even when the raw material is available. Second, by relating blunted reward responses in the VTA/SN to poor source memory, the findings bridge two symptoms of depression that are typically considered separately—such as anhedonia and memory impairments. Over the last decade, researchers have made significant progress linking anhedonia to dysfunction in putative dopamine circuits, emphasizing the impact of this dysfunction on mood and motivation (Treadway and Zald, 2011). The data presented here suggest that the negative effects of anhedonia may extend beyond mood and motivation, and into memory. Thus, although many of the neurobiological details remain to be elucidated, one clinical point is clear: anhedonia merits its status as a cardinal symptom of depression, and should be a primary target of clinical interventions.

#### Limitations

Several limitations should be acknowledged. First, by decoupling the reward tokens from payment and performance, we may have inadvertently limited their impact on brain activation and behavior. Future studies using rewards that confer additional payment are warranted. Second, while the post-task arousal and valence ratings were helpful for verifying that the tokens elicited the intended emotional responses, more sensitive measures would be preferable. In particular, simultaneous psychophysiological measurements (e.g. heart rate) might reveal negative effects of depression on emotional responses to these stimuli. Third, while our *a priori* hypotheses concerning source memory and activation of the VTA/SN and MTL were supported, our prediction that the ‘REWARD?’ recognition cue would selectively enhance reward *vs* source memory in controls—but not depressed participants—was not confirmed (see Supplementary Material). Source memory retrieval is an effortful process that recruits the dorsolateral prefrontal cortex (DLPFC; Dobbins *et al.*, 2002), and depression has been consistently linked to DLPFC hypoactivation (Koenigs and Grafman, 2009) and difficulty with effortful tasks (Roy-Byrne *et al.*, 1986). Therefore, although our recognition cueing manipulation did not affect memory accuracy as expected, the potential for impaired source memory retrieval in depression—possibly mediated by DLPFC hypofunction—is worthy of further inquiry.

Finally, the connection between memory and brain activation in this study must be considered preliminary, because there were too few



incorrect responses to support a 'subsequent memory' analysis contrasting neural regions that support successful vs unsuccessful encoding. With this limitation in mind, we restricted our analysis to encoding trials that led to successful memory. This approach seemed preferable to analyzing the response to all reward and zero tokens, regardless of memory status, as this could reveal brain regions whose activation was entirely unrelated to encoding success. Furthermore, we found a reliable correlation between VTA/SN activation and memory accuracy in controls that was not present in the depressed group, with a similar trend in the right parahippocampus, which provides evidence linking brain activation to memory. Nonetheless, future studies should use a subsequent memory design (with sufficient 'forget' trials) to permit firmer claims about relationships between brain activation and memory in depression.

## CONCLUSION

This study relates blunted reward source memory in depression to aberrant activation of the VTA/SN and right parahippocampus during encoding. The results highlight neural mechanisms implicated in weak memory for positive material in depression, as well as interwoven cognitive and affective dysfunction in MDD.

## SUPPLEMENTARY DATA

Supplementary data are available at SCAN online.

## Conflict of Interest

Dr Pizzagalli has received consulting fees from ANT North America Inc. (Advanced Neuro Technology), AstraZeneca, Ono Pharma USA, Servier and Shire for projects unrelated to the present research. Dr Dillon and Dr Dobbins report no conflicts of interest.

## REFERENCES

- Adcock, R.A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., Gabrieli, J.D.E. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron*, 50, 507–17.
- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* 4th edn. Washington, DC: American Psychiatric Association.
- Beats, B.C., Sahakian, B.J., Levy, R. (1996). Cognitive performance in tests sensitivity to frontal lobe dysfunction in the elderly depressed. *Psychological Medicine*, 26, 591–603.
- Beck, A.T., Steer, R.A., Brown, G.K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: The Psychological Corporation.
- Beevers, C.G., Worthy, D.A., Gorlick, M.A., Nix, B., Chotibut, T., Maddox, T.W. (2012). Influence of depression symptoms on history independent reward and punishment processing. *Psychiatry Research*, 207, 53–60.
- Bethus, I., Tse, D., Morris, R.G.M. (2010). Dopamine and memory: modulation of the persistence of memory for novel hippocampal NMDA receptor-dependent paired associates. *Journal of Neuroscience*, 30, 1610–8.
- Brett, M., Anton, J.L., Valabregue, R., Poline, J.B. (2002). *Region of Interest Analysis Using an SPM Toolbox*. Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2–6, Sendai, Japan.
- Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature*, 482, 85–8.
- Dale, A.M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping*, 8, 109–14.
- Davachi, L., Mitchell, J.P., Wagner, A.D. (2003). Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 2157–62.
- Dichter, G.S., Felder, J.N., Petty, C., Bizzell, J., Ernst, M., Smoski, M.J. (2009). The effects of psychotherapy on neural responses to rewards in major depression. *Biological Psychiatry*, 66, 886–97.
- Dobbins, I.G., Foley, H., Schacter, D.L., Wagner, A.D. (2002). Executive control during episodic retrieval: multiple prefrontal processes subserve source memory. *Neuron*, 35, 989–96.
- Dunbar, G.C., Lishman, W.A. (1984). Depression, recognition-memory and hedonic tone a signal detection analysis. *British Journal of Psychiatry*, 144, 376–82.
- Hamilton, J.P., Gotlib, I.H. (2008). Neural substrates of increased memory sensitivity for negative stimuli in major depression. *Biological Psychiatry*, 63, 1155–62.
- Holmes, A.J., Pizzagalli, D.A. (2008). Spatiotemporal dynamics of error processing dysfunction in major depressive disorder. *Archives of General Psychiatry*, 65, 179–88.
- Jenkinson, M., Bannister, P., Brady, M., Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17, 825–41.
- Koenigs, M., Grafman, J. (2009). The functional neuroanatomy of depression: distinct roles for ventromedial and dorsolateral prefrontal cortex. *Behavioral Brain Research*, 201, 239–43.
- Krebs, R.M., Heipertz, D., Schuetze, H., Düzel, E. (2011). Novelty increases the mesolimbic functional connectivity of the substantia nigra/ventral tegmental area (SN/VTA) during reward anticipation: evidence from high-resolution fMRI. *Neuroimage*, 58, 647–55.
- Lisman, J.E., Grace, A.A. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron*, 46, 703–13.
- March, R.L., Hicks, J.L. (1998). Test formats change source-monitoring decision processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1137–51.
- Niv, Y., Daw, N.D., Joel, D., Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)*, 191, 507–20.
- Pizzagalli, D.A., Holmes, A.J., Dillon, D.G., et al. (2009). Reduced caudate and nucleus accumbens response to rewards in unmedicated subjects with Major Depressive Disorder. *American Journal of Psychiatry*, 166, 702–10.
- Rorden, C., Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioral Neurology*, 12, 191–200.
- Roy-Byrne, P.P., Weingartner, H., Bierer, L.M., Thompson, K., Post, R.M. (1986). Effortful and automatic cognitive processes in depression. *Archives of General Psychiatry*, 43, 265–7.
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, 80, 1–27.
- Shohamy, D., Wagner, A.D. (2008). Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron*, 60, 378–89.
- Smith, S.M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17, 143–55.
- Snaith, R.P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., Trigwell, P. (1995). A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. *British Journal of Psychiatry*, 167, 99–103.
- Snodgrass, J.G., Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215.
- The Psychological Corporation. (2001). *Wechsler Test of Adult Reading Manual*. San Antonio, TX: The Psychological Corporation.
- Treadway, M.T., Zald, D.H. (2011). Reconsidering anhedonia in depression: lessons from translational neuroscience. *Neuroscience and Biobehavioral Reviews*, 35, 537–55.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., et al. (2002). Automated Anatomical Labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15, 273–89.
- van der Kouwe, A.J.W., Benner, T., Fischl, B., et al. (2005). On-line automatic slice positioning for brain MR imaging. *Neuroimage*, 27, 222–30.
- van der Kouwe, A.J., Benner, T., Salat, D.H., Fischl, B. (2008). Brain morphometry with multiecho MPRAGE. *Neuroimage*, 40, 559–69.
- Watson, D., Clark, L.A., Weber, K., Assenheimer, J.S., Strauss, M.E., McCormick, R.A. (1995). Testing a tripartite model: II. Exploring the symptom structure of anxiety and depression in student, adult, and patient samples. *Journal of Abnormal Psychology*, 104, 15–25.
- Wittmann, B.C., Schiltz, K., Boehler, C.N., Düzel, E. (2008). Mesolimbic interaction of emotional valence and reward improves memory formation. *Neuropsychologia*, 46, 1000–8.
- Wolosin, S.M., Zeithamova, D., Preston, A.R. (2012). Reward modulation of hippocampal subfield activation during successful associative encoding and retrieval. *Journal of Cognitive Neuroscience*, 24, 1532–47.



## Supplementary Material

*Weak Reward Source Memory in Depression Reflects Blunted Activation of VTA/SN and*

*Parahippocampus*

Daniel G. Dillon, Ian G. Dobbins, and Diego A. Pizzagalli

### Methods

#### **Diagnostic interview**

The Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Patient Edition (First et al., 2002) was used to identify eligible depressed and healthy participants. All depressed participants were unmedicated with no history of psychosis. Axis I comorbidity was confined to anxiety disorders, with the exception of one participant who met criteria for binge disorder. Past psychotropic medication was allowed (no use in the past 2 weeks for benzodiazepines, 6 weeks for SSRIs, 6 months for dopaminergic drugs). Controls reported no current or past Axis I diagnosis. Participants were 18 - 64 years old and right-handed. None presented with neurological conditions or significant medical history, or met criteria for lifetime substance dependence or substance abuse in the last year. Directly following the interview, participants completed 20 practice trials of the memory task in preparation for the MRI session. MRI sessions took place about 10 days after the interview (mean $\pm$ SD: 10.46 $\pm$ 0.95). Participants were compensated for the diagnostic interview (\$10/hour).

#### **VTA/SN small volume correction**

In a high-resolution fMRI study, Krebs et al. (2011) identified reward responses in the caudal VTA, with the peak activation at the following MNI coordinates: [3, -15, -16]. Using SPM8 software, we were unable to center a spherical ROI on odd number coordinates. Thus, we

evaluated the *Group x Stimulus x Token Type* interaction in the VTA/SN with an 8 mm sphere centered at [2, -14, -16], 1 mm medial and superior to the peak reported in Krebs et al. (2011).

## Results

### Clinical Data

The median length of the current major depressive episode was seven months. Four participants reported prior depressive episodes too numerous or indistinct to count; the remaining depressed participants reported a median of two prior depressive episodes. Ten depressed participants had one comorbid anxiety disorder (panic disorder,  $n = 1$ ; social phobia,  $n = 9$ ), while another had three (social phobia, specific phobia, panic disorder) plus binge eating disorder. The remaining ten depressed participants had no comorbid Axis-I diagnoses.

### Behavioral Data

**Preliminary accuracy analysis: *Group x Cue x Source* ANOVA.** In the main text, we focus on group differences in source memory accuracy collapsed across the two recognition cues. This approach was taken following a preliminary *Group x Cue x Source* (reward, zero, new) ANOVA that yielded a *Group x Source* interaction,  $F(2, 80) = 3.15, p = 0.048$ , along with main effects of *Cue*,  $F(1, 40) = 6.00, p = 0.02$ , and *Source*,  $F(2, 80) = 18.86, p < 0.001$ . As described in the main text, the *Group x Source* interaction was driven by better accuracy for drawings from the reward source vs. the zero source in controls, but not the MDD group. Furthermore, both groups showed better accuracy (i.e., correct rejections) for new drawings vs. old drawings from either source, accounting for the main effect of *Source*.

The *Cue* effect reflected better overall accuracy under the “Reward?” cue ( $74.58\% \pm 0.11$ ) vs. the “Zero?” cue ( $69.52\% \pm 0.13$ ). Critically, this effect did not differ across sources, *Cue x Source*,  $F < 1$ , or groups, *Cue x Group*,  $F(1, 40) = 2.27, p = 0.14$ . Furthermore, the triple

interaction was not significant, *Group x Cue x Source*,  $F < 1$ . Thus, we found no support for our prediction that the “Reward?” cue would selectively improve reward source memory in the controls. Because *Cue* did not interact with *Group* or *Source*, we collapsed across the recognition cues to avoid unnecessary complexity and focus on group differences in source memory.

**Does the *Group x Source* interaction reflect group differences in cue-driven response bias?** As noted above, our preliminary accuracy analysis did not reveal *Group x Cue* or *Group x Cue x Source* interactions. This indicates that there was no tendency for either group to respond “yes” to either cue more often than the other group. For example, there was no evidence that controls were more prone to reply “yes” to the ‘Reward?’ cue than the depressed participants. Nonetheless, we conducted two additional analyses to more firmly establish that the *Group x Source* interaction did not reflect cue-driven group differences in response bias.

First, we ran separate *Group x Source* ANOVAs on percent correct data rendered under each cue. As shown in Figure S1, a consistent pattern was observed. When considering only responses to the “Reward?” cue (Figure S1A), there was a significant *Group x Source* interaction,  $F(1, 40) = 4.27$ ,  $p = 0.045$ . In controls, a paired *t*-test confirmed better accuracy for the reward source vs. the zero source,  $t(20) = 2.29$ ,  $p = 0.03$ . By comparison, the same test in the MDD group revealed no difference between sources,  $t(20) < 1$ . When considering only responses rendered under the “Zero?” cue (Figure S1B), the interaction was not significant,  $F(1, 40) = 2.58$ ,  $p = 0.12$ . However, the controls again showed better accuracy for the reward source vs. the zero source,  $t(20) = 2.50$ ,  $p = 0.02$ . In the depressed group, there was again no difference in accuracy across the sources,  $t(20) < 1$ . By revealing the same pattern of results under both recognition cues, this analysis confirms that the *Group x Source* interaction was not driven by the cueing manipulation and did not reflect group differences in cue-driven response bias.

This conclusion was strengthened by our second analysis, which used  $d'$  instead of percent correct as the accuracy measure. For responses to the “Reward?” cue, we computed  $d'$  scores by comparing reward source hit rates and zero source correct rejection rates vs. false alarm rates to new drawings. For responses to the “Zero?” cue, we computed  $d'$  scores by comparing reward source correct rejection rates and zero source hit rates vs. false alarm rates to new drawings. We then entered these data into two *Group x Source* ANOVAs, considering responses to the “Reward?” cue and “Zero?” cue separately.

Results were entirely consistent with the analysis of percent correct data. For responses to the “Reward?” cue, a *Group x Source* interaction emerged,  $F(1, 40) = 5.25, p = 0.03$ . In controls,  $d'$  was significantly higher for the reward source ( $2.04 \pm 1.05$ ) vs. the zero source ( $1.63 \pm 0.90$ ),  $t(20) = 2.66, p = 0.02$ . By contrast, there was no difference in  $d'$  for the reward source ( $1.87 \pm 0.87$ ) vs. the zero source ( $1.94 \pm 1.08$ ) in the depressed group,  $t < 1$ . As with the percent correct data, when responses to the “Zero?” cue were considered alone, the *Group x Source* interaction was not significant,  $F(1, 40) = 2.48, p = 0.12$ . However, the controls again showed better accuracy for the reward source ( $1.98 \pm 1.10$ ) vs. the zero source ( $1.62 \pm 1.14$ ),  $t(20) = 2.56, p = 0.02$ . By comparison, there was no difference in reward ( $1.15 \pm 1.71$ ) vs. zero ( $1.20 \pm 0.98$ ) source accuracy in the depressed group,  $t(20) < 1$ . Thus, the  $d'$  and percent correct analyses yielded identical results. Because  $d'$  is sensitive to response bias, this consistency confirms that response bias did not drive the *Group x Source* interaction.

## **fMRI Data**

**Responses to drawings in the VTA/SN and right parahippocampus.** To complement the analysis of responses to tokens in the VTA/SN and right parahippocampus, we conducted *Group x Token Type* (reward, zero) ANOVAs on the neural response to drawings in these



regions. Because tokens were presented after the drawings disappeared from the screen, we did not expect *Group x Token Type* interactions when considering the neural response to drawings.

The raw data from these analyses are presented in Figure S2. Our expectation was met in the VTA/SN (Figure S2A), as neither the *Group x Token Type* interaction nor any other effect was significant (all  $ps > 0.10$ ). The lack of effects indicates that the group difference in VTA/SN responses to reward tokens (Figure 3, main text) was not confounded by a group difference in response to the drawings. Unexpectedly, the *Group x Token Type* interaction was significant in the right parahippocampus,  $F(1, 40) = 5.35, p = 0.03$  (Figure S2B). However, follow-up *t*-tests revealed no significant differences, either between groups or between conditions within groups. Specifically, the response to drawings that were followed by reward tokens was marginally stronger in depressed vs. control participants,  $t(40) = 1.73, p = 0.09$ , with no group difference in response to drawings that were followed by zero tokens,  $t(40) < 1$ . Paired *t*-tests comparing the responses to drawings followed by reward vs. zero tokens yielded no difference in controls ( $p = 0.19$ ) and a marginal result in the MDD group,  $t(20) = 1.89, p = 0.07$ .

In summary, these analyses did not reveal reliable group differences in the neural response to drawings. No effects approached significance in the VTA/SN, while in the right parahippocampus, there was a marginally stronger response to drawings followed by reward tokens in the depressed group. As there was no group difference in the right parahippocampal response to reward tokens (Figure 3, main text), this trend-level result does not confound the main findings. Overall, these data confirm that group differences in VTA/SN and right parahippocampal responses were temporally localized to the reward and zero tokens, as opposed to the drawings that preceded them.

**Regions showing a main effect of *Stimulus*.** Table 2 lists sectors of auditory, visual, and motor cortex that showed a main effect of *Stimulus*, indicating that they responded differently to drawings vs. tokens. To unpack these effects, beta weights were extracted from each region (averaged over all significant voxels) and submitted to follow-up ANOVAs. Significant findings are briefly presented in narrative fashion as they are not of central interest.

Activations in right Heschl's gyrus, left superior temporal gyrus, and left intracalcarine cortex reflected a stronger response to tokens vs. drawings. It is unclear why the left intracalcarine cortex showed this response profile, but Heschl's gyrus and the superior temporal gyrus correspond to bilateral auditory cortex, which would be expected to respond strongly to the sounds that accompanied each token. Stronger responses to drawings vs. tokens were observed in bilateral occipital pole, left precentral gyrus, and left central opercular cortex. Activation of the occipital pole likely reflects the fact that each drawing was unique and more visually complex than the reward and zero tokens, which were repeatedly presented; these factors could drive a stronger response in visual cortex. Left precentral gyrus corresponds to motor cortex, thus its activation is probably due to the right-handed button press made in response to each drawing. The operculum is part of somatosensory cortex (Eickhoff et al., 2006), thus its strong response to drawings may also have been driven by the button presses executed in response to each one.

**Regions showing a main effect of *Trial Type*.** Bilateral auditory cortex (Heschl's gyrus) and a cluster spanning the right postcentral gyrus and parietal white matter were characterized by main effects of *Trial Type*. Follow-up analysis of responses in auditory cortex revealed substantially stronger responses to zero tokens vs. reward tokens, with much smaller differences in response to the drawings. The differential response to the tokens was unsurprising, as the tone (delivered as part of the zero token) was subjectively louder than the chime (delivered as part of

the reward token). The right postcentral gyrus responded most strongly to reward tokens and drawings that preceded them. It is unclear why this response profile was seen.

**Exclusion of one control from correlational analysis.** One control with extreme reward minus zero token difference scores in the VTA/SN and right parahippocampus was excluded from the correlational analysis linking reward minus zero token difference scores in these regions to source memory. This participant's reward minus zero token difference score values exceeded the mean plus 3 SDs in the VTA/SN, and exceeded the mean plus 2 SDs in the right parahippocampus, which inflated the strength of the correlations in controls. We verified that excluding this participant did not affect the group difference in source memory accuracy or any of the follow-up analyses in the VTA/SN or parahippocampus.

## References

- Eickhoff, S. B., Amunts, K., Mohlberg, H., & Zilles, K. (2009) The human parietal operculum. II. Stereotaxic maps and correlation with functional imaging results. *Cerebral Cortex*, 16, 268-279.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002) Structured clinical interview for DSM-IV-TR Axis I disorders, research version, patient edition (SCID-I/P). New York, NY: Biometrics Research, New York State Psychiatric Institute.



### Supplementary Figure Captions

*Figure S1.* Memory accuracy (percent correct) for drawings from the reward and zero sources under the (A) “Reward?” and (B) “Zero?” cues. Regardless of which recognition cue was presented, controls showed better accuracy for the reward source vs. the zero source, but there was no accuracy difference between the two sources in depressed group. Analysis of  $d'$  data yielded identical results.

*Figure S2.* Neural responses to drawings in the (A) VTA/SN and (B) right parahippocampus. In the VTA/SN, no significant effects emerged. In the right parahippocampus, an unexpected *Group x Token Type* interaction was observed. However, follow-up tests revealed no significant differences between groups or between conditions within each group. y-axis indicates magnitude of the beta weight; bars indicate standard error.

Figure S1

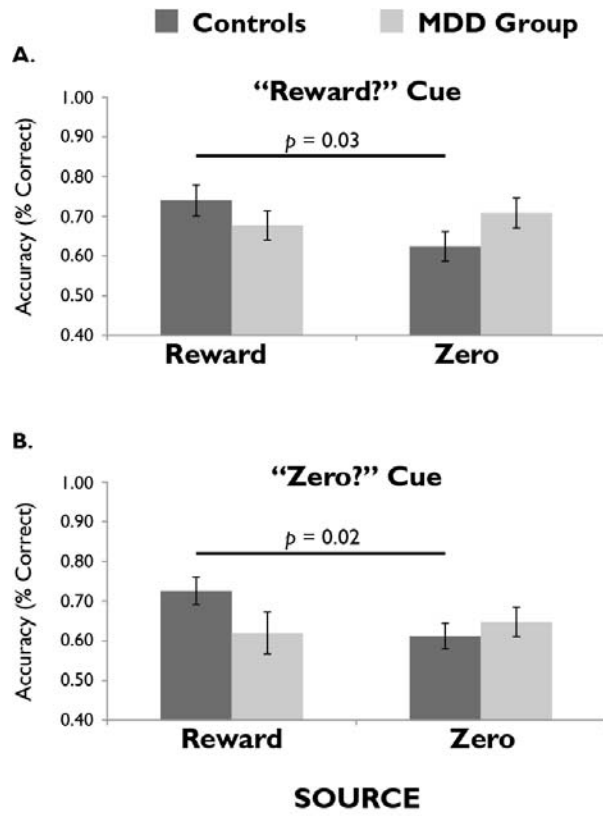


Figure S2

