OPEN ACCESS

University of Dundee

University of Dundee

Impaired reward-related learning signals in remitted unmedicated patients with recurrent depression

Geugies, Hanneke ; Mocking, Roel J. T. ; Figueroa, Caroline A. ; Groot, Paul F. C. ; Marsman, Jan-Bernard C. ; Servaas, Michelle N.

BRAIN
A JOURNAL OF NEUROLOGY

Impaired reward-related learning signals in remitted unmedicated patients with recurrent depression

Hanneke Geugies,[1,2] Roel J.T. Mocking,[3] Caroline A. Figueroa,[3,4] Paul F.C. Groot,[5] Jan-Bernard C. Marsman,[2] Michelle N. Servaas,[2] (iD) J. Douglas Steele,[6] Aart H. Schene[7,8] and Henricus G. Ruhé[1,3,4,8]

One of the core symptoms of MDD is anhedonia, an inability to experience pleasure. In patients with MDD, a dysfunctional reward-system may exist, with blunted temporal difference reward-related learning signals in the ventral striatum and increased temporal difference-related (dopaminergic) activation in the ventral tegmental area. Anhedonia often remains as residual symptom during remission; however, it remains largely unknown whether the abovementioned reward systems are still dysfunctional when patients are in remission. We used a Pavlovian classical conditioning functional MRI task to explore the relationship between anhedonia and the temporal difference-related response of the ventral tegmental area and ventral striatum in medication-free remitted recurrent depression patients ($n = 36$) versus healthy control subjects ($n = 27$). Computational modelling was used to obtain the expected temporal difference errors during this task. Patients, compared to healthy controls, showed significantly increased temporal difference reward learning activation in the ventral tegmental area ($P_{FWE,SVC} = 0.028$). No differences were observed between groups for ventral striatum activity. A group × anhedonia interaction [$t(57) = -2.29$, $P = 0.026$] indicated that in patients, higher anhedonia was associated with lower temporal difference activation in the ventral tegmental area, while in healthy controls higher anhedonia was associated with higher ventral tegmental area activation. These findings suggest impaired reward-related learning signals in the ventral tegmental area during remission in patients with depression. This merits further investigation to identify impaired reward-related learning as an endophenotype for recurrent depression. Moreover, the inverse association between reinforcement learning and anhedonia in patients implies an additional disturbing influence of anhedonia on reward-related learning or vice versa, suggesting that the level of anhedonia should be considered in behavioural treatments.

1 University Medical Center Groningen, University Center for Psychiatry, Mood and Anxiety Disorders, University of Groningen, The Netherlands
2 University Medical Center Groningen, Department of Neuroscience, Neuroimaging Center, University of Groningen, The Netherlands
3 Department of Psychiatry, Amsterdam University Medical Center, location AMC, University of Amsterdam, The Netherlands
4 Warneford Hospital, Department of Psychiatry, University of Oxford, UK
5 Department of Radiology and Nuclear Medicine, Amsterdam University Medical Center, location AMC, University of Amsterdam, The Netherlands
6 Medical School (Neuroscience), University of Dundee, Scotland, UK
7 Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, The Netherlands
8 Department of Psychiatry, Radboud University Medical Center, Nijmegen, The Netherlands

Correspondence to: H.G. Ruhé, MD, PhD
Radboudumc, Reinier Postlaan 4, Nijmegen, 6500 HB, The Netherlands
E-mail: h.g.ruhe@gmail.com

Correspondence may also be addressed to: H. Geugies

University Medical Center Groningen, University Center for Psychiatry, Mood and Anxiety Disorders, Hanzeplein 1, Groningen, 9700 RD, The Netherlands

E-mail: hannekegeugies@gmail.com

Introduction

Major depressive disorder (MDD) is a highly prevalent and disabling disease (Mathers and Loncar, 2006). Although treatment of a depressive episode can induce remission of symptoms, depressive episodes unfortunately tend to recur after a period of recovery (Frank *et al.*, 1991). The incidence of recurrences varies (depending on the population and setting) but may reach 80% within 5 years (Bockting *et al.*, 2009). Therefore, recurrence is a major contributor to the immense (in)direct annual costs of MDD (estimated > €113 billion in Europe) (Gustavsson *et al.*, 2011), which necessitates prevention of recurrence and knowledge of underlying aetiopathogenetic mechanisms.

An inability to experience pleasure/reward (anhedonia) is one of the core symptoms of depression (Ebmeier *et al.*, 2006) and often persists as a residual symptom after remission (Conradi *et al.*, 2011). The ability to experience reward appears important in providing resilience against recurrence. Positive emotional responses decrease stress-sensitivity (Wichers *et al.*, 2007), and predict recovery during antidepressant treatment (Wichers *et al.*, 2009). Furthermore, pleasure also has an important motivational function; it reinforces behaviour that leads to (potentially) pleasurable events (conditioning) (Pavlov, 1927). Patients with MDD often report either difficulties in experiencing normally positive events as pleasurable (i.e. consummatory anhedonia or 'liking') or deficits in motivation to pursue rewards (i.e. motivational anhedonia or 'wanting') (Treadway and Zald, 2011). Furthermore, patients with MDD have difficulties in learning new behaviours that might improve their mood or keep them well (Vrieze *et al.*, 2013).

Wanting, liking and learning have been identified as three important dissociable components of reward (Berridge *et al.*, 2009), where especially wanting and learning have been linked to dopaminergic neurotransmission in the reward-network consisting of the ventral striatum (Knutson *et al.*, 2001; Schott *et al.*, 2008) and ventral tegmental area (D'Ardenne *et al.*, 2008; Kumar *et al.*, 2008; Schott *et al.*, 2008). In the reward circuitry, the ventral tegmental area projects to the ventral striatum and receives projections from the habenula, which is involved in regulating the intensity of reward-seeking and distress-avoiding behaviour (Loonen and Ivanova, 2017).

Previous studies have shown that reward learning stimuli evoke short phasic firing patterns of dopaminergic neurons (Schultz, 1998; Tobler *et al.*, 2005), resembling temporal difference prediction errors (Schultz *et al.*, 1997; Kumar *et al.*, 2008). Temporal difference prediction errors are important for making a predictive association between stimuli and outcomes when stimuli are repeated and learned. Over time, dopaminergic neurons will predict a response as a result of previous associations between a stimulus and its rewarding value (classical conditioning/reinforcement learning). Briefly, before learning, delivery of an unexpected reward is followed by phasic dopamine activation. When the association between stimulus and reward has been consolidated, dopaminergic firing is activated at the presentation of the stimulus (cue), while firing to the reward itself is reduced when delivered as expected. However, when a learned cue is not followed by an expected reward, this results in a decrease in dopaminergic firing (below baseline), representing negative prediction errors.

Dysfunctions in anticipatory and consummatory reward processes in MDD have been investigated (Knutson *et al.*, 2008; Pizzagalli *et al.*, 2009; Smoski *et al.*, 2009), as well as temporal difference reward-related learning in depressed patients versus control subjects (Kumar *et al.*, 2008). Kumar and colleagues identified increased activation of dopaminergic neurons in the ventral tegmental area when thirsty patients with MDD were learning associations between a stimulus (picture) and a reward (water delivery) (Kumar *et al.*, 2008). Furthermore, the ventral striatum has been repeatedly reported to be hypoactive in MDD both in reinforcement-learning as in other reward processing paradigms (Kumar *et al.*, 2008; Pizzagalli *et al.*, 2009; Gradin *et al.*, 2011; Robinson *et al.*, 2012; Hall *et al.*, 2014).

Although evidence for a dysfunctional reward system in depressed patients is established (Martin-Soelch, 2009), there is still very little understanding whether these reward systems remain dysfunctional when patients are in remission. Previous studies conducted in subjects at risk for depression and with subthreshold depression have demonstrated that abnormalities in processing of wanting and liking aspects of reward may be a trait marker for MDD (McCabe *et al.*, 2009, 2012; Stringaris *et al.*, 2015; McCabe, 2016; Pan *et al.*, 2017). However, it remains largely unknown whether a dysfunction in processing of reward-related learning represents a trait rather than a

state-dependent abnormality, which may be of importance with regard to vulnerability for recurrence. Furthermore, little is known about the association between persistent anhedonia and deficits of reward processing in remitted patients (Dunlop and Nemeroff, 2007). We therefore quantified the response of the dopamine reward system (i.e. ventral striatum and ventral tegmental area) during a classical conditioning functional MRI task in medication-free patients with remitted recurrent depression (rrMDD), who were at high risk of recurrence (Mocking *et al.*, 2016). In addition, we hypothesized a link between abnormalities in the reward system and anhedonia levels. Based on earlier work in depressed patients during classical conditioning (Kumar *et al.*, 2008), we hypothesized decreased ventral striatum activation and increased ventral tegmental area activation in response to temporal difference reward-related learning in rrMDD versus controls, with positive associations of these abnormalities with anhedonia.

Materials and methods

Participants

As part of a larger neuroimaging study investigating vulnerability for recurrence in MDD (Mocking *et al.*, 2016), participants were recruited by advertisements and through previous clinical treatment and/or previous studies. In particular, patients aged 35–65 with a known recurrent depressive disorder, currently in stable remission without medication, were identified and approached for this study. Matched healthy control subjects were recruited via advertisements. We obtained permission from the local ethics committee and written informed consent from all participants (Mocking *et al.*, 2016). Dimensional assessment of illness severity was obtained by an observer-rated Hamilton Depression Rating Scale (HDRS17) (Hamilton, 1967), and a self-rated Snaith Hamilton Anhedonia and Pleasure Scale (SHAPS) (Snaith *et al.*, 1995). Sixty-two patients with MDD were scanned who satisfied the following criteria: (i) presence of a recurrent depression defined as $\geqslant 2$ depressive episodes according to the structured interview for DSM-IV (SCID); (ii) stable remission defined as a HDRS17 $\leqslant 7$ for at least eight subsequent weeks; and (iii) aged between 35–65. We scanned 41 healthy controls that were matched on the basis of age, sex and years of education. All participants were without any medications for >4 weeks. Exclusion criteria were: (i) a current diagnosis of alcohol or drug dependence; (ii) psychotic or bipolar disorder; (iii) primary anxiety disorder; (iv) MRI participation contraindications such as implanted metal; (v) electroconvulsive therapy within 2 months before scanning; and (vi) a history of head trauma or neurological disease. Healthy controls were excluded if they had personal (SCID) or first degree relatives with a psychiatric disorder.

Task

A Pavlovian classical conditioning task was used specifically to assess reward learning during passive observation (Kumar *et al.*, 2008) instead of an instrumental design that would have allowed to fit behavioural responses but potentially focusses on different aspects of learning. Participants were asked to refrain from liquids for $\geqslant 6\,h$ prior to scanning to ensure they were thirsty. The Pavlovian classical conditioning task consisted of four blocks of 30 trials of 8 s each. The task started with one block (30 trials) without juice delivery (the neutral condition), but with the to-be conditioned stimuli (but not yet conditioned). After the neutral block, three blocks followed that included juice delivery. One of two pictures was alternately shown on the screen [the conditioned stimulus (CS)] 2 s after the start of each trial. Two seconds thereafter, the conditioned stimulus was followed by the presence or absence of small amounts (0.2 ml) of rewarding juice [the unconditioned stimulus (US)] at different probabilities (80–20%) (Fig. 1). Every block, a change occurred (three times in total) in which the picture that was 'rewarding' (for 80% of the time) was switched with the non-rewarding picture. Before and after the task participants received 0.2 ml fluid after which they were asked how much money they were willing to pay to get more juice (wanting) and how much they enjoyed the taste of the juice (liking). A visual analogue scale ranging from −2 (receive money/unpleasant, respectively) to 2 (pay money/pleasant, respectively) was used to assess wanting and liking, with the centre of the scale being neutral. Juice delivery was via a polythene tube that was attached to a syringe-driver pump (B Braun-Infusomat P) positioned in the scanner control room, interfaced with the stimulus presentation computer. Stimuli were presented using E-prime 2 (Psychology Software Tools, Pittsburgh, PA). The participants were instructed to try to find out which picture predicted the juice delivery and notified that this association could change over time. With changing probabilities of juice delivery, temporal difference reward-learning signals were calculated (Kumar *et al.*, 2008). Other tasks within the same MRI session were carried out after the Pavlovian task to avoid possible confounding effects.

Data acquisition

Magnetic resonance images were acquired on a Phillips 3 T Achieva XT MRI scanner using a 32-channel SENSE head coil. T_2*-weighted gradient-echo-planar images were collected with the following parameters: repetition time 1500 ms, echo time 28 ms, 25 slices, 1125 volumes, field of view: 240×240 mm and matrix 80×80; voxel size: $3 \times 3 \times 3$ mm. Slices were oriented with 30° tilt from the AC-PC transverse plane and acquired in ascending order. High resolution T_1-weighted anatomical images were acquired with the following parameters: repetition time 8.3 ms, echo time 3.8 ms, 220 slices, field of view: 240×188 mm and matrix 240×240; voxel size: $1 \times 1 \times 1$ mm. Cardiac and respiratory signals were acquired concurrently during the scan and used to facilitate physiological noise correction in the analysis.

Data preprocessing

Images were preprocessed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm) implemented in MATLAB R2013a (The MathWorks Inc., Natick, MA). Structural and functional images were reoriented in anterior-posterior commissure alignment to facilitate co-registration. Functional images were re-aligned to the first functional image and were co-registered to the T_1-weighted image. Structural images were segmented into
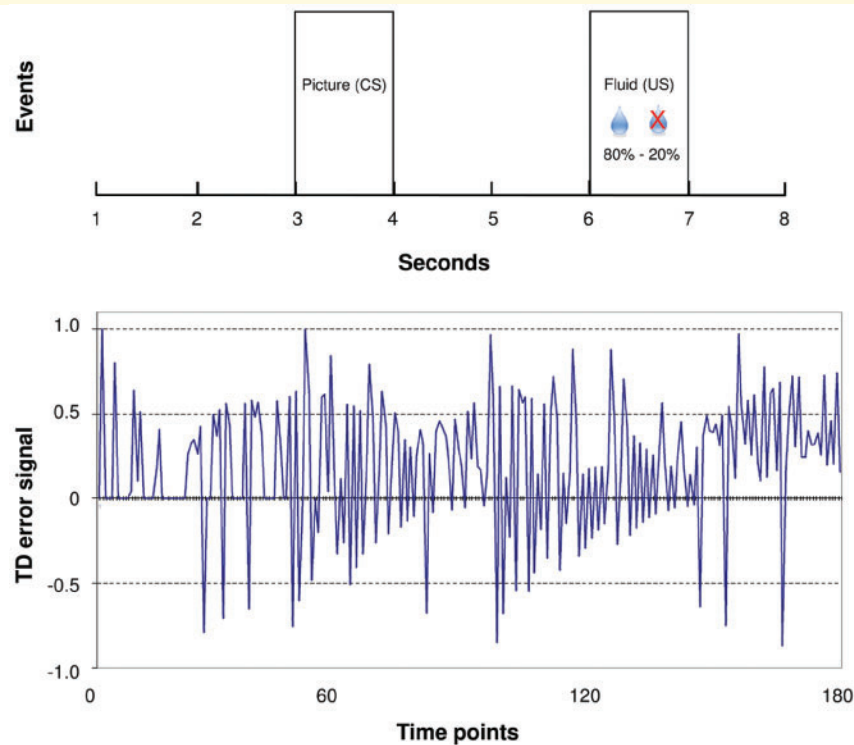
Figure 1 Pavlovian reinforcement task paradigm. (A) Timing of the conditioned (CS) and unconditioned stimulus (US) within one trial. **(B)** Example of a temporal difference (TD) error signal of one subject.

grey matter, white matter, and CSF. T_1-weighted images were used to create a study-specific group template using the DARTEL algorithm (Ashburner, 2007). Subsequently, functional images were normalized to Montreal Neurological Institute (MNI) space using this intermediate group template. Voxel sizes remained $3 \times 3 \times 3$ mm during DARTEL spatial normalization, and images were smoothed with a 4 mm Gaussian kernel. Physiological cardiac and respiratory noise signals were modelled and eliminated retrospectively by the DRIFTER algorithm (Sarkka *et al.*, 2012), a Bayesian method for physiological noise modelling and removal, allowing accurate dynamical tracking of the variations in the cardiac and respiratory frequencies. Frequency trajectories of the physiological signals were estimated by the interacting multiple models filter algorithm (reference signal 1 = respiratory signal: sampling interval = 500 Hz, array of possible frequencies = 10:70 bpm; reference signal 2 = cardiac signal: sampling interval = 500 Hz, array of possible frequencies = 40:140 bpm). The estimated frequency trajectories were then used in a state space model in combination with a Kalman filter and Rauch–Tung–Striebel smoother, which separated the signal into a cleaned activation-related signal, physiological noise, and white measurement noise components. Details regarding this algorithm are described in Sarkka *et al.* (2012).

Temporal difference learning model

From each participant, the E-prime log files were used to extract the timing of the unconditioned stimulus and the conditioned stimulus. All eight time points were modelled, with the

conditioned stimulus defined at time point 3 and the unconditioned stimulus at time point 6. The calculation of the temporal difference prediction errors was derived from Kumar *et al.* (2008), who used a standard temporal difference model derived from Dayan and Abbott (2001). As in previous studies, a same set of parameters was used for all subjects (Kumar *et al.*, 2008, 2018; Daw, 2011; Gradin *et al.*, 2011). The predicted value (V) at any time t was defined as:

$$\hat{V}(t) = \sum_i w_i x(t) \tag{1}$$

where $x_i(t)$ is coded with a 1 or a 0 (for all time points) for the presence or absence of a conditioned stimulus at time t. w_i corresponds to a weight that was updated on each trial in order to capture learning by:

$$\Delta w_i = \alpha \sum x_i(t) \delta(t) \tag{2}$$

where α is corresponding to a factor chosen in advance, which represents the learning rate. As recommended for model-based functional MRI analysis (Wilson and Niv, 2015), we selected multiple plausible learning rates from the literature (0.1 and 0.4 from Kumar *et al.*, 2008 and O'Doherty *et al.*, 2006; 0.2 from O'Doherty *et al.*, 2003, 2004; 0.45 from Gradin *et al.*, 2011; 0.5 from Lawson *et al.*, 2017) and explored which learning rate fitted our data best. We chose $\alpha = 0.45$ as the optimal learning rate based on optimal signal-to-noise ratio calculations and estimation of efficiency values of SPM designs (Liu *et al.*, 2017 and Supplementary material for details regarding the calculation of estimation efficiency). To ensure

our results were robust, we compared temporal difference (TD)-related activation in the CS × TD + US × TD contrast across the range of learning rates (Supplementary material).

The temporal difference error signal was defined as:

$$\delta(t) = r(t) + \gamma\hat{V}(t+1) - \hat{V}(t) \qquad (3)$$

where $r(t)$ is coded with a 1 or a 0 (for all time points) for delivery of juice or no-juice, respectively and γ corresponds to a factor chosen in advance, which determined the importance of later reinforcements compared with previous ones. Following previous studies, $\gamma = 1.0$ was used (Kumar et al., 2008; Gradin et al., 2011). This means that the model did not include discounting effects and assumed that such effects did not differ between groups, which is a common assumption in model-based functional MRI literature (O'Doherty et al., 2003, 2006; Kumar et al., 2008; Gradin et al., 2011).

Statistical analysis

Sample characteristics

Analyses were performed with SPSS v22.0 (SPSS Inc., USA). We used $P < 0.05$ as threshold for significance. Independent sample t-tests, χ^2-tests and non-parametric Mann-Whitney U-tests were used to compare demographics (age, sex, education, IQ) and clinical variables (HDRS, SHAPS, number of lifetime episodes, age of onset) between rrMDD and healthy control subjects.

Behavioural data

Group differences in wanting and liking ratings were analysed using repeated-measures analysis of variance with group (rrMDD, healthy controls) as the between-subjects factor and time (pre-task and post-task) as the within-subjects factor. Because groups differed slightly but significantly, we used HDRS scores as a covariate, to exclude effects driven by (small) HDRS differences.

Imaging data

In SPM12, an event-related random effects design was used for the analysis. For each participant, first-level haemodynamic responses for each stimulus (conditioned and unconditioned) were modelled using a canonical haemodynamic response function model. The temporal difference prediction errors were entered into the model as parametric modulators for the conditioned and unconditioned stimulus conditions. To look at main cue and delivery task effects separately, we modelled a conditioned stimulus > neutral and a unconditioned stimulus > neutral condition. We also modelled a pooled contrast (conditioned stimulus + unconditioned stimulus > neutral) to see if the task would elicit ventral striatum activity regardless if it was during cue (conditioned stimulus, CS) or delivery (unconditioned stimulus, US). Given our primary hypothesis about temporal difference (TD) related activation, we modelled the contrast CS × TD + US × TD. Separate contributions of the conditioned and unconditioned stimulus temporal difference errors were also modelled by a CS × TD and US × TD condition. A high-pass filter of 128 s was used to remove low frequency noise. Realignment parameters and their first derivatives were added to the model to address residual movement not corrected by realignment.

A priori regions of interest were the ventral tegmental area and ventral striatum. Region of interest selection was based on the definition used by D'Ardenne et al. (2008), who applied a comparable task and analysis, specifically tailored to image dopaminergic signals in the ventral tegmental area and ventral striatum. At second-level, we used a one-sample t-test to investigate main effects of cue/delivery (conditioned stimulus + unconditioned stimulus > neutral, conditioned stimulus > neutral and unconditioned stimulus > neutral contrasts), and main effect of prediction error (CS × TD + US × TD). We used independent two-sample t-tests to look at differences between patients and controls (CS × TD + US × TD, and CS × TD and US × TD separately). The main effect of cue/delivery images were thresholded at $P < 0.05$ uncorrected to display the extent of the signal (Kumar et al., 2008). As we had clear a priori regions of interest, a small volume correction (SVC), based on ventral tegmental area and ventral striatum coordinates from previous research (D'Ardenne et al., 2008), with a sphere of radius 5 mm, was applied with significance defined as $P < 0.05$ familywise error corrected. A second analysis was performed with HDRS scores as a covariate.

We then evaluated the association between the ventral tegmental area temporal difference signal and anhedonia (SHAPS) (Franken et al., 2007) with a multiple regression analysis. Here the ventral tegmental area temporal difference signal was the dependent variable, while SHAPS scores, group and the group × SHAPS interaction were examined with HDRS scores as a covariate.

Based on the suggestions of anonymous reviewers we performed additional sensitivity analyses. These are described in the Supplementary material.

Data availability

The data that support the findings of this study are available upon reasonable request.

Results

Patient disposition and sample characteristics

From the 62 rrMDD patients and 41 healthy control subjects that were scanned, we excluded three patients and two healthy controls because of abnormal brain anatomy and five patients and four healthy controls because of corrupted or missing task data. During the analysis phase, 18 patients and eight healthy controls were excluded because of missing or corrupted physiological data needed for filtering of cardiac and respiratory noise, leaving a sample of 36 patients and 27 healthy controls included in the final analyses. Excluded subjects did not significantly differ in sample characteristics from the included sample. No significant differences were observed between rrMDD patients and healthy controls (Table 1), except higher residual symptomatology (HDRS; $U = 224$, $P < 0.001$) and anhedonia (SHAPS; $U = 253$, $P = 0.002$) in rrMDD patients.

Table 1 Demographic and clinical characteristics

Characteristic		rrMDD (*n* = 36)	Healthy controls (*n* = 27)	Test-statistic (df)	*P*
Age, years	Mean (range)	47 (36–65)	41 (36–63)	$U = 806$	0.24
Sex	Male/female	10/26	8/19	$\chi^2(1) = 0.03$	0.87
Education levels[a]	n (1/2/3/4/5/6/7)	0/0/0/2/14/14/6	0/0/0/0/13/10/4	$\chi^2(3) = 1.86$	0.60
IQ	Mean (SD)	108 (8.9)	105 (9.9)	$t(56) = 1.12$	0.71
HDRS intake	Median (IQR)	3 (1–5)	0 (0–1)	$U = 181$	<0.001
HDRS MRI	Median (IQR)	3.5 (2–6)	1 (0–2)	$U = 224$	<0.001
SHAPS	Median (IQR)	24 (20–28)	17 (14–23)	$U = 253$	0.002
Lifetime episodes, *n*	Mean (SD)	9.2 (11.3)	-	-	-
Age of onse, years	Mean (SD)	25.7 (10.9)	-	-	-

IQR = interquartile range.
[a]Level of educational attainment (Verhage, 1964). Levels range from 1 to 7 (1 = primary school not finished, 7 = preuniversity/university degree).

Figure 2 Liking and wanting ratings. (A) Liking ratings: no significant main effect of group [$F(1,57) = 1.00$, $P = 0.322$], no significant main effect of time [$F(1,57) = 2.67$, $P = 0.108$] and no significant group × time interaction [$F(1,57) = 2.52$, $P = 0.118$]. Depicted are the estimated marginal means (means adjusted for any other variables in the model) with standard errors. **(B)** Wanting ratings: no significant main effect of group [$F(1,57) = 1.77$, $P = 0.188$], no significant main effect of time [$F(1,57) = 0.06$, $P = 0.803$] and no significant group × time interaction [$F(1,57) = 0.002$, $P = 0.961$]. Depicted are the estimated marginal means (means adjusted for any other variables in the model) with standard errors. HC = healthy controls.

Behavioural results

For the wanting and liking ratings (corrected for HDRS differences) no main effect of group or time was observed. No significant group × time interactions were identified (Fig. 2).

Functional MRI results

We observed main effect activation of the ventral striatum during delivery of cues and reward (conditioned stimulus + unconditioned stimulus > neutral, conditioned stimulus > neutral and unconditioned stimulus > neutral contrasts) (Table 2 and Supplementary Fig. 2). We also found a main effect of prediction error in the ventral tegmental area and the ventral striatum (CS × TD + US × TD contrast) (Table 2 and Supplementary Fig. 3). We found increased temporal difference-related activation (CS × TD + US × TD contrast) in the ventral tegmental

area in rrMDD patients compared to healthy controls ($P_{\text{FWE,SVC}} = 0.028$) (Table 3 and Fig. 3). The significance of this group difference was $P_{\text{FWE,SVC}} = 0.048$ after correction for HDRS scores between groups (Supplementary Fig. 4). Temporal difference signals in the ventral striatum did not differ significantly between groups. When comparing rrMDD versus healthy controls in the CS × TD and the US × TD contrast separately, differences in temporal difference-related ventral tegmental area activation were not significant (Table 3).

Association between ventral tegmental area temporal difference signal and anhedonia ratings

The regression model with SHAPS scores, group, group × SHAPS interaction and HDRS explained 21% of the variance [$F(4,57) = 3.78$, $P = 0.009$]. This

Table 2 Within-group activation

		Contrast	Location	MNI coordinates	z	Significance[a]
Main effect	Cue + reward delivery (CS + US > neutral)	rrMDD + healthy controls	VS	(−9, 12, −6)	2.62	0.004
	Cue delivery alone (CS > neutral)	rrMDD + healthy controls	VS	(−9, 12, −6)	3.36	0.000
				(6, 9, 0)	2.68	0.004
	Reward delivery alone (US > neutral)	rrMDD + healthy controls	VS	(−3, 6, −3)	1.83	0.034
				(9, 15, 0)	1.74	0.041
	Total TD signal (CS × TD + US × TD)	rrMDD + healthy controls	VTA	(0, −21, −3)	2.66	0.004
			VS	(−6, 3, −3)	2.05	0.020
				(6, 3, −3)	1.86	0.031

CS = conditioned stimuli; TD = temporal difference signal; US = unconditioned stimuli; VS = ventral striatum; VTA = ventral tegmental area.
[a]$P_{uncorrected}$ in order to display the extent of the signal.

Table 3 Between-group activation

		Contrast	Location	MNI coordinates	z	Significance[a]
Group differences	Total TD signal (CS × TD + US × TD)	rrMDD > healthy controls	VTA	(0, −21, −3)	2.79	0.028
			VS	(9, 0, -3)	2.91	0.154
				(−6, 3, -6)	2.64	0.361
		healthy controls > rrMDD	No clusters survived threshold			
	CS × TD	rrMDD > healthy controls	VTA	(0, −21, −3)	2.38	0.071
		healthy controls > rrMDD	No clusters survived threshold			
	US × TD	rrMDD > healthy controls	VTA	(0, −18, −15)	1.70	0.229
		healthy controls > rrMDD	No clusters survived threshold			

CS = conditioned stimuli; TD = temporal difference signal; US = unconditioned stimuli; VS = ventral striatum; VTA = ventral tegmental area.
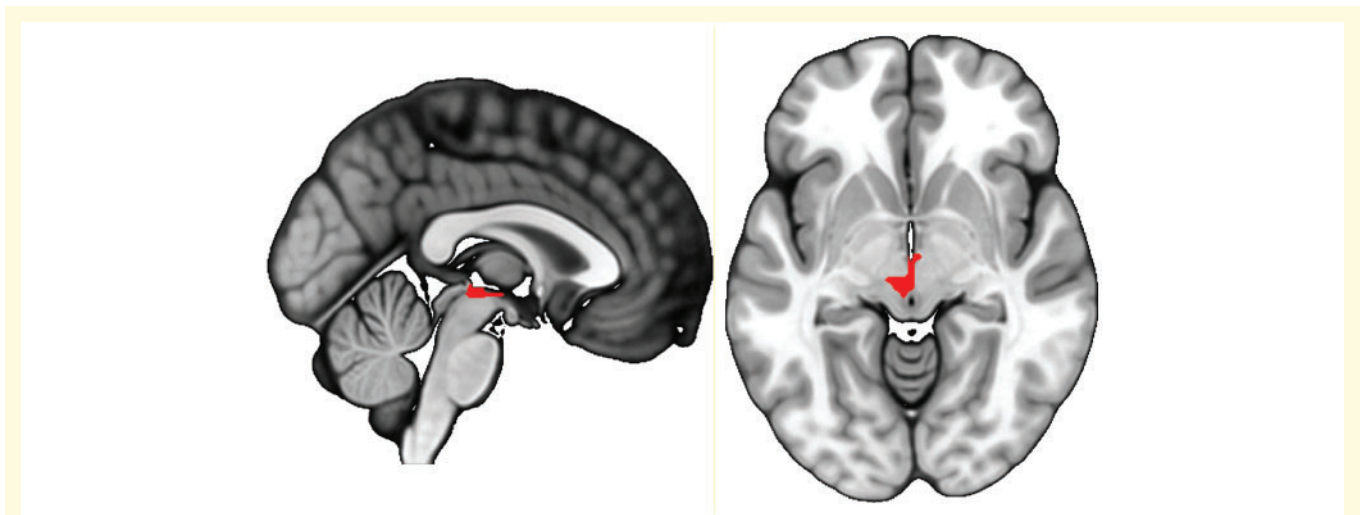[a]FWE peak level corrected + small volume corrected.

Figure 3 Temporal difference error-related activation comparing rrMDD and healthy controls. rrMDD patients show more activation related to temporal difference signals in the ventral tegmental area compared to healthy controls (Z = 2.79, P = 0.028 FWE corrected on peak-level, small volume corrected).

model showed a significant group × SHAPS interaction [$t(57) = −2.29$, $P = 0.026$] in addition to the main effect for group [$t(57) = 3.03$, $P = 0.004$] (Fig. 4). In rrMDD patients, higher anhedonia was associated with lower ventral tegmental area temporal difference activation. In healthy controls, higher anhedonia was associated with higher ventral tegmental area temporal difference activation.
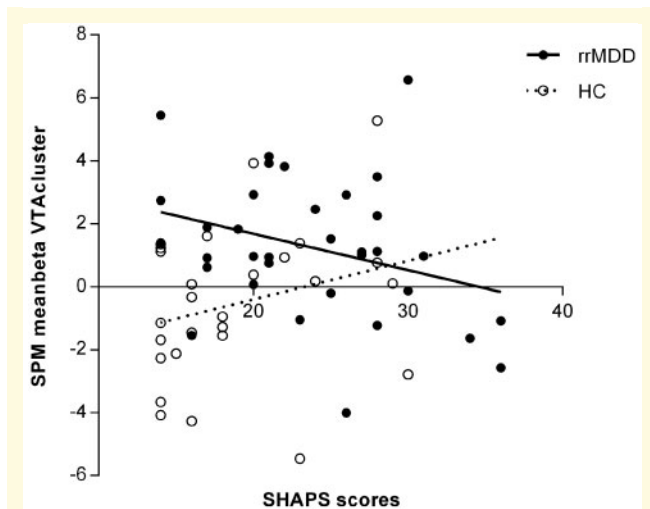
Figure 4 Association of ventral tegmental area activation and anhedonia (SHAPS). Significant group × SHAPS interaction [$t(57) = -2.29$, $P = 0.026$] and a main effect for group [$t(57) = 3.03$, $P = 0.004$]. HC = healthy controls.

Discussion

This study explored the response of the ventral tegmental area and ventral striatum during a classical conditioning functional MRI task in medication-free patients with rrMDD compared to healthy control subjects. We found significantly increased temporal difference reward learning activation in the ventral tegmental area in rrMDD patients compared to healthy controls. No differences between the groups were observed for ventral striatum activity. Moreover, we investigated the relationship with anhedonia and showed that in rrMDD patients, higher anhedonia was associated with lower ventral tegmental area temporal difference reward learning activation, while in healthy controls, higher anhedonia was associated with higher ventral tegmental area activation.

This study did not demonstrate the difference in basic wanting and liking processing, as described in depressed patients (Treadway and Zald, 2011). Furthermore, wanting and liking properties did not differ over time between both groups. This result is in agreement with McCabe *et al.* (2009), who also found no significant differences between recovered depression patients and healthy controls on ratings of wanting (pleasantness) and liking. This suggests that these differences are either not present, or are smaller in a remitted state. This notion is further corroborated by our functional MRI findings, where we found no group differences in basic processing of reward in the ventral striatum. Previous functional MRI studies in depressed patients found reduced ventral striatum activity (Pizzagalli *et al.*, 2009; Smoski *et al.*, 2009; Robinson *et al.*, 2012), although not consistently (Knutson *et al.*, 2008; Rothkirch *et al.*, 2017; Rutledge *et al.*, 2017). Inconsistencies might be attributable to differences in study designs and/or patient

characteristics. However, studies investigating reward processing in remitted depression patients, consistently never reported ventral striatum differences (Dichter *et al.*, 2012; Ubl *et al.*, 2015; Hammar *et al.*, 2016). We therefore propose that the reduction in reward sensitivity and ventral striatum activation during reward delivery in depressed patients is likely to recover after achieving remission and therefore could be considered a state effect. Another explanation for a difference between ventral tegmental area and ventral striatum temporal difference activation can be based on findings by Klein-Flügge *et al.* (2011), who demonstrated that classic temporal difference reward prediction error activity was specific to the ventral tegmental area, but not the ventral striatum, which suggests decoupling between ventral tegmental area dopaminergic neuron firing and ventral striatum dopamine release.

In contrast to the suggested recovery of basic wanting and liking processing in patients with remitted depression, our results show that the underlying learning signals to learn the associations between reward outcome and stimuli are impaired. Kumar *et al.* (2008) demonstrated increased ventral tegmental area temporal difference-related activations during reward-learning in patients while depressed, which correlated with illness severity. These findings were interpreted as reflecting a compensatory response to an impaired function of other non-brainstem regions, such as the ventral striatum, of the mesolimbic pathway. However, the current results demonstrate that also in remitted recurrent depression, increased ventral tegmental area activity during reward-learning persists, while the difference in temporal difference-related activation in the ventral striatum seems to be restored.

However, Kumar *et al.* (2008) investigated a sample of depressed patients who were non-responsive to long-term antidepressants, and healthy control subjects in unmedicated and (acutely) medicated state. Interestingly, the temporal difference signals in the ventral striatum of medicated healthy controls (compared to the unmedicated healthy controls) were reduced and did no longer differ significantly from patients with MDD. Animal studies report different effects of acute versus chronic administration of antidepressants (Sekine *et al.*, 2007) and in patients with MDD, acute administration of antidepressants reduced temporal difference error-related neural activity in the ventral striatum (McCabe *et al.*, 2010; Chase *et al.*, 2013; Herzallah *et al.*, 2013). Therefore, it could be hypothesized that reduced temporal difference signals in the ventral striatum in medicated, depressed patients might reflect medication effects instead of state effects. Indeed, a recent paper corroboratively reported no differences in prediction error-related activity in the ventral striatum in unmedicated depressed patients versus healthy control subjects (Rothkirch *et al.*, 2017). We are aware that there are relatively few studies on unmedicated samples, and that previous cohorts are often slightly less severe than medicated cohorts. Therefore, it is difficult to make claims about medication based on the present unmedicated cohort, and more direct

comparisons are needed. However, the described effects of medication could provide an additional explanation for our findings of comparable temporal difference-related activity in the ventral striatum.

Our finding of increased ventral tegmental area temporal difference signals in rrMDD patients versus healthy control subjects is in line with the report in unresponsive medicated patients with MDD (Kumar *et al.*, 2008) and suggests a trait-like abnormality, i.e. impaired reward-related learning is associated with MDD, and seems to be state-independent, which are both important criteria of the endophenotype concept (Gottesman and Gould, 2003), relevant for recurrent depression. Nevertheless, to the best of our knowledge, the heritability (another endophenotype characteristic) of impaired reward-related learning has yet to be demonstrated.

The phasic dopamine firing into temporal difference signals has been well described (Schultz *et al.*, 1997; Schultz, 1998; Tobler *et al.*, 2005), which makes it valid to interpret temporal difference signal impairments as a dysfunction of the dopaminergic system. The role of the (dysfunctional) dopamine system in the pathophysiology of MDD has been emphasized by Dunlop and Nemeroff (2007). They suggest the existence of subtypes of depression stemming from abnormal dopaminergic neurotransmission, and suggest further research regarding the involvement of dopamine circuit dysfunction in non-response to treatment, or treatment resistance. Given that 20% of recurrent depressive episodes become chronic despite treatment (Judd *et al.*, 1998), and with the present findings in mind, future studies focusing on reward-related learning impairments in treatment-resistant depression are warranted.

The significant group × anhedonia interaction indicated that rrMDD patients with higher levels of anhedonia have reduced ventral tegmental area temporal difference signals. Reduced ventral tegmental area activity was also reported by Dillon *et al.* (2014a), who investigated reward memory in unmedicated adults with MDD. Furthermore, the group × anhedonia interaction indicated that healthy controls with higher levels of anhedonia have increased ventral tegmental area temporal difference signals. Interestingly, a study in healthy participants reported that higher levels of anhedonia were not associated with the ventral tegmental area, but instead associated with reduced activity in other key areas of the reward circuitry linked to the ventral tegmental area (basal forebrain, ventral striatum). Therefore, the observed increased ventral tegmental area activity in healthy controls might be compensatory to overcome a diminished reward sensitivity in more anhedonic healthy controls (Keller *et al.*, 2013).

In contrast, the opposite relation between anhedonia and ventral tegmental area temporal difference activation in MDD, even in the remitted state, could be interpreted in accordance with Eldar and Niv (2015), who suggested that reward prediction errors are strongly related to mood. If remitted depressed individuals are recovering from depression, it may be that they experience larger positive prediction errors as they find rewarding events more rewarding than they are used to. Hence a larger reward prediction error might be observed. This would explain why remitted depression patients with greater residual anhedonia have smaller prediction error responses.

Another explanation can be based on Liu *et al.* (2017), who found that in depressed, unmedicated MDD, especially in response to expected punishment, higher levels of anhedonia were associated with attenuated habenula activation. The habenula is not only important in punishment processes (i.e. expectation of aversive stimuli), but also plays a central role in reward processing (i.e. absence of rewards) (Lawson *et al.*, 2014), specifically via projections to the ventral tegmental area. Studies investigating habenula function in humans and animal models of MDD showed that the habenula is hyperactive in MDD (Shumake and Gonzalez-Lima, 2013; Dillon *et al.*, 2014b; Lecca *et al.*, 2014; Benarroch, 2015; Zhao *et al.*, 2015; Liu *et al.*, 2017). As the habenula is known to inhibit ventral tegmental area dopaminergic firing (Matsumoto and Hikosaka, 2007), and the absence of a reward is in particular a strong activator of the habenula (Proulx *et al.*, 2014), this could explain the negative correlation between anhedonia and ventral tegmental area temporal difference signals in rrMDD patients. More anhedonic rrMDD patients, experiencing less/absence of rewards, might have further increased habenula hyperactivity, resulting in increased (habenula-driven) inhibition of dopaminergic firing in the ventral tegmental area. By a stronger decrease in reward expectancy this could even strengthen anhedonia and associated depressive behaviour in a vicious cycle. Via this mechanism, anhedonia might have a modifying effect on the effectiveness of behavioural treatments, commonly used to alleviate MDD, which, however, remains to be established (Treadway and Zald, 2011). Notably, in rats, a decrease of habenula firing has been associated with reduction of depressive-like behaviour (Li *et al.*, 2011), and deep brain stimulation in the habenula resulted in remission of symptoms in a patient with treatment-resistant depression (Sartorius *et al.*, 2010). Unfortunately, due to low power, our present study design was not suitable to specifically explore negative temporal difference errors coding for the absence of a reward. Therefore, the role of the habenula in the association between anhedonia and temporal difference signals remains speculative, requiring verification in future studies.

Regardless whether a functional impairment of the ventral tegmental area or the habenula underlies the association with anhedonia, it would be interesting to investigate whether the observed impairments in reinforcement learning are associated with recurrence. A link between recurrence and impaired reinforcement learning would suggest that—in line with previous research—the focus of therapy should not only lie on diminishing negative affect but also enhancing positive affect by training patients to focus attention on positive reinforcers (Wichers *et al.*, 2010, 2012; Servaas *et al.*, 2017).

Focusing on positive experiences might train the ability to make associations between behaviour and pleasurable outcomes and might reinforce repetition of reward-provoking behaviour (operant conditioned learning). Training the ability for (rr)MDD patients to learn about rewarding feedback in daily life and remediate impaired reinforcement learning should be investigated in future studies, while considering anhedonia as a moderator.

Strengths and limitations

This is the first study exploring reinforcement learning during remission in a relatively large group of unmedicated patients with MDD. Nevertheless, potential limitations are present. First, as in the original task (Kumar *et al.*, 2008), the experimental task lacked an active response to the appearance of the pictures on the screen. This excludes the possibility of any behavioural confound in the Pavlovian learning. Although this passive conditioning task was specifically used to assess particular aspects of learning, participants might have lost their engagement or attention to the task and we were not able to assess individualized learning rates. In new experiments, an active response (e.g. button press) will be embedded in the task, which will facilitate the possibility to fit the model to the data and select parameters that show the best overall fit to the signals. Furthermore, future analyses could benefit from novel methods that extract parameters by fitting computational models to neural data alone or to a combination of behavioural and neural data at the same time (Purcell *et al.*, 2010; Turner *et al.*, 2013; Frank *et al.*, 2015; Turner *et al.*, 2016; van Ravenzwaaij *et al.*, 2017). Second, the direct measurement of dopamine signalling with functional MRI is impossible. Nevertheless, strong evidence supports that blood oxygen level-dependent signals in reward-related brain areas reflect dopaminergic release (Pessiglione *et al.*, 2006; Knutson and Gibbs, 2007). Third, by modelling the temporal difference error signal and comparing patients and controls, we reject the null hypothesis of no differences between groups. These differences between groups could be due to either actual difference in dopaminergic learning signals between groups, or differences between groups (and individuals in the groups) in learning learning-rate and/or discount factor, which are used to model the temporal difference errors. However, previous research found no differences in model parameters between patients with MDD and healthy controls (Gradin *et al.*, 2011). Moreover, using a single set of model parameters across all participants and groups showed more robust results in multi-subject functional MRI studies (Daw, 2011). Therefore, we interpret our findings as representing differences in dopaminergic temporal difference signals between groups. A fourth limitation is that the *a priori* choices that were made for our analysis (e.g. learning rate selection, choice of smoothing kernel) are one out of many approaches that can be considered. We chose to explore plausible learning rates from literature instead of exploring

an entire range of learning rates between 0 and 1. This method was chosen because the primary aim was to investigate the difference between patients and controls and not to methodologically explore how to model learning rates. Furthermore, it has been suggested in literature that even gross deviations in the learning rate lead to only minimal changes in the neural results and that precise model fitting is not always necessary for model-based functional MRI (Wilson and Niv, 2015). When exploring our neural results in the range we described, we indeed found comparable results when using different learning rates. A fifth limitation is that a currently depressed group or scanning of the subjects when depressed was not incorporated in the present analysis. This hampers the ability to draw inferences about persistence. However, in its present form, the study can be very helpful for the identification of factors that remain impaired during remission in depressive patients with a history of recurrence. Lastly, no individual levels of thirst were obtained at the start of the experiment. Nevertheless, participants confirmed that they refrained from liquids for ⩾6 h prior to scanning, which made it fair to assume sufficient levels of thirstiness.

Conclusion

In summary, we demonstrated impaired reward-related learning in unmedicated patients with a recurrent MDD during remission, which may be an (endo)phenotype linked to depression vulnerability. Our findings add to evidence for state-independent, impaired temporal difference learning signals in the ventral tegmental area, which requires further investigation as an endophenotype for (recurrent) MDD. Furthermore, the association between impaired reinforcement learning and anhedonia in rrMDD patients strengthens the need to focus on this residual symptom and investigate remediation of hedonic capacity and processing of reward-related learning in rrMDD.

Acknowledgements

We thank three anonymous reviewers for their thoughtful comments which clarified our methods and suggested us to perform additional sensitivity analyses.

Funding

preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

Ashburner J. A fast diffeomorphic image registration algorithm. Neuroimage 2007; 38: 95–113.

Benarroch EE. Habenula: recently recognized functions and potential clinical relevance. Neurology 2015; 85: 992–1000.

Berridge KC, Robinson TE, Aldridge JW. Dissecting components of reward: 'liking', 'wanting', and learning. Curr Opin Pharmacol 2009; 9: 65–73.

Bockting CL, Spinhoven P, Wouters LF, Koeter MW, Schene AH; DELTA Study Group. Long-term effects of preventive cognitive therapy in recurrent depression: a 5.5-year follow-up study. J Clin Psychiatry 2009; 70: 1621–8.

Chase HW, Nusslock R, Almeida JR, Forbes EE, LaBarbara EJ, Phillips ML. Dissociable patterns of abnormal frontal cortical activation during anticipation of an uncertain reward or loss in bipolar versus major depression. Bipolar Disord 2013; 15: 839–54.

Conradi HJ, Ormel J, de Jonge P. Presence of individual (residual) symptoms during depressive episodes and periods of remission: a 3-year prospective study. Psychol Med 2011; 41: 1165–74.

D'Ardenne K, McClure SM, Nystrom LE, Cohen JD. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. Science 2008; 319: 1264–7.

Daw ND. Trial-by-trial data analysis using computational models. In: Delgado MR, Phelps EA, Robbins TW, editors. Decision making, affect, and learning. attention and performance XXIII. New York: Oxford University Press Inc.; 2011.

Dayan P, Abbott F. Theoretical neuroscience: computational and mathematical modeling of neural systems. Cambridge, MA: The MIT Press; 2001.

Dichter GS, Kozink RV, McClernon FJ, Smoski MJ. Remitted major depression is characterized by reward network hyperactivation during reward anticipation and hypoactivation during reward outcomes. J Affect Disord 2012; 136: 1126–34.

Dillon DG, Dobbins IG, Pizzagalli DA. Weak reward source memory in depression reflects blunted activation of VTA/SN and parahippocampus. Soc Cogn Affect Neurosci 2014a; 9: 1576–83.

Dillon DG, Rosso IM, Pechtel P, Killgore WD, Rauch SL, Pizzagalli DA. Peril and pleasure: an rdoc-inspired examination of threat responses and reward processing in anxiety and depression. Depress Anxiety 2014b; 31: 233–49.

Dunlop BW, Nemeroff CB. The role of dopamine in the pathophysiology of depression. Arch Gen Psychiatry 2007; 64: 327–37.

Ebmeier KP, Donaghey C, Steele JD. Recent developments and current controversies in depression. Lancet 2006; 367: 153–67.

Eldar E, Niv Y. Interaction between emotional state and learning underlies mood instability. Nat Commun 2015; 6: 6149.

Frank E, Prien RF, Jarrett RB, Keller MB, Kupfer DJ, Lavori PW, et al. Conceptualization and rationale for consensus definitions of terms in MDD: remission, recovery, relapse, and recurrence. Arch Gen Psychiatry 1991; 48: 851–5.

Franken IH, Rassin E, Muris P. The assessment of anhedonia in clinical and non-clinical populations: further validation of the Snaith-Hamilton Pleasure Scale (SHAPS). J Affect Disord 2007; 99: 83–9.

Frank MJ, Gagne C, Nyhus E, Masters S, Wiecki TV, Cavanagh JF, et al. fMRI and EEG predictors of dynamic decision parameters during human reinforcement learning. J Neurosci 2015; 35: 485–94.

Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry 2003; 160: 636–45.

Gradin VB, Kumar P, Waiter G, Ahearn T, Stickle C, Milders M, et al. Expected value and prediction error abnormalities in depression and schizophrenia. Brain 2011; 134: 1751–64.

Gustavsson A, Svensson M, Jacobi F, Allgulander C, Alonso J, Beghi E, et al. Cost of disorders of the brain in Europe 2010. Eur Neuropsychopharmacol 2011; 21: 718–79.

Hall GB, Milne AM, Macqueen GM. An fMRI study of reward circuitry in patients with minimal or extensive history of major depression. Eur Arch Psychiatry Clin Neurosci 2014; 264: 187–98.

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

Hammar A, Neto E, Clemo L, Hjetland GJ, Hugdahl K, Elliott R. Striatal hypoactivation and cognitive slowing in patients with partially remitted and remitted major depression. PsyCh J 2016; 5: 191–205.

Herzallah MM, Moustafa AA, Natsheh JY, Abdellatif SM, Taha MB, Tayem YI, et al. Learning from negative feedback in patients with MDD is attenuated by SSRI antidepressants. Front Integr Neurosci 2013; 7: 67.

Judd LL, Akiskal HS, Maser JD, Zeller PJ, Endicott J, Coryell W, et al. A prospective 12-year study of subsyndromal and syndromal depressive symptoms in unipolar MDDs. Arch Gen Psychiatry 1998; 55: 694–700.

Keller J, Young CB, Kelley E, Prater K, Levitin DJ, Menon V. Trait anhedonia is associated with reduced reactivity and connectivity of mesolimbic and paralimbic reward pathways. J Psychiatr Res 2013; 47: 1319–28.

Klein-Flügge MC, Hunt LT, Bach DR, Dolan RJ, Behrens TE. Dissociable reward and timing signals in human midbrain and ventral striatum. Neuron 2011; 72: 654–64.

Knutson B, Gibbs SE. Linking nucleus accumbens dopamine and blood oxygenation. Psychopharmacology 2007; 191: 813–22.

Knutson B, Adams CM, Fong GW, Hommer D. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. J Neurosci 2001; 21: RC159.

Knutson B, Bhanji JP, Cooney RE, Atlas LY, Gotlib IH. Neural responses to monetary incentives in major depression. Biol Psychiatry 2008; 63: 686–92.

Kumar P, Waiter G, Ahearn T, Milders M, Reid I, Steele JD. Abnormal temporal difference reward-learning signals in major depression. Brain 2008; 131: 2084–93.

Kumar P, Goer F, Murray L, Dillon DG, Beltzer ML, Cohen AL, et al. Impaired reward prediction error encoding and striatal-midbrain connectivity in depression. Neuropsychopharmacology 2018; 43: 1581–8.

Lawson RP, Nord CL, Seymour B, Thomas DL, Dayan P, Pilling S, et al. Disrupted habenula function in major depression. Mol Psychiatry 2017; 22: 202–8.

Lawson RP, Seymour B, Loh E, Lutti A, Dolan RJ, Dayan P, et al. The habenula encodes negative motivational value associated with primary punishment in humans. Proc Natl Acad Sci U S A 2014; 111: 11858–63.

Lecca S, Meye FJ, Mameli M. The lateral habenula in addiction and depression: an anatomical, synaptic and behavioral overview. Eur J Neurosci 2014; 39: 1170–8.

Li B, Piriz J, Mirrione M, Chung C, Proulx CD, Schulz D, et al. Synaptic potentiation onto habenula neurons in the learned helplessness model of depression. Nature 2011; 470: 535–9.

Liu WH, Valton V, Wang LZ, Zhu YH, Roiser JP. Association between habenula dysfunction and motivational symptoms in unmedicated MDD. Soc Cogn Affect Neurosci 2017; 12: 1520–33.

Loonen AJM, Ivanova SA. Circuits regulating pleasure and happiness: evolution and role in mental disorders. Acta Neuropsychiatr 2017; 11: 1–14.

Martin-Soelch C. Is depression associated with dysfunction of the central reward system? Biochem Soc Trans 2009; 37: 313–7.

Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 2006; 3: e442.

Matsumoto M, Hikosaka O. Lateral habenula as a source of negative reward signals in dopamine neurons. Nature 2007; 447: 1111–5.

McCabe C. Neural signals of 'intensity' but not 'wanting' or 'liking' of rewards may be trait markers for depression. J Psychopharmacol 2016; 30: 1020–7.

McCabe C, Cowen PJ, Harmer CJ. Neural representation of reward in recovered depressed patients. Psychopharmacology 2009; 205: 667–77.

McCabe C, Mishor Z, Cowen PJ, Harmer CJ. Diminished neural processing of aversive and rewarding stimuli during selective serotonin reuptake inhibitor treatment. Biol Psychiatry 2010; 67: 439–45.

McCabe C, Woffindale C, Harmer CJ, Cowen PJ. Neural processing of reward and punishment in young people at increased familial risk of depression. Biol Psychiatry 2012; 72: 588–94.

Mocking RJ, Figueroa CA, Rive MM, Geugies H, Servaas MN, Assies J, et al. Vulnerability for new episodes in recurrent MDD: protocol for the longitudinal DELTA-neuroimaging cohort study. BMJ Open 2016; 6: e009510.

O'Doherty JP, Buchanan TW, Seymour B, Dolan RJ. Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. Neuron 2006; 49: 157–66.

O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ. Temporal difference models and reward-related learning in the human brain. Neuron 2003; 38: 329–37.

O'Doherty JP, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 2004; 304: 452–4.

Pan PM, Sato JR, Salum GA, Rohde LA, Gadelha A, Zugman A, et al. Ventral striatum functional connectivity as a predictor of adolescent depressive disorder in a longitudinal community-based sample. Am J Psychiatry 2017; 174: 1112–9.

Pavlov IP. Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. Oxford, London: Oxford University Press; 1927.

Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 2006; 442: 1042–5.

Pizzagalli DA, Holmes AJ, Dillon DG, Goetz EL, Birk JL, Bogdan R, et al. Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with MDD. Am J Psychiatry 2009; 166: 702–10.

Proulx CD, Hikosaka O, Malinow R. Reward processing by the lateral habenula in normal and depressive behaviors. Nat Neurosci 2014; 17: 1146–52.

Purcell BA, Heitz RP, Cohen JY, Schall JD, Logan GD, Palmeri TJ. Neurally constrained modeling of perceptual decision making. Psychol Rev 2010; 117: 1113–43.

Robinson OJ, Cools R, Carlisi CO, Sahakian BJ, Drevets WC. Ventral striatum response during reward and punishment reversal learning in unmedicated MDD. Am J Psychiatry 2012; 169: 152–9.

Rothkirch M, Tonn J, Kohler S, Sterzer P. Neural mechanisms of reinforcement learning in unmedicated patients with MDD. Brain 2017; 140: 1147–57.

Rutledge RB, Moutoussis M, Smittenaar P, Zeidman P, Taylor T, Hrynkiewicz L, et al. Association of neural and emotional impacts of reward prediction errors with major depression. JAMA Psychiatry 2017; 74: 790–7.

Sarkka S, Solin A, Nummenmaa A, Vehtari A, Auranen T, Vanni S, et al. Dynamic retrospective filtering of physiological noise in BOLD fMRI: DRIFTER. Neuroimage 2012; 60: 1517–27.

Sartorius A, Kiening KL, Kirsch P, von Gall CC, Haberkorn U, Unterberg AW, et al. Remission of major depression under deep brain stimulation of the lateral habenula in a therapy-refractory patient. Biol Psychiatry 2010; 67: e9–11.

Schott BH, Minuzzi L, Krebs RM, Elmenhorst D, Lang M, Winz OH, et al. Mesolimbic functional magnetic resonance imaging activations during reward anticipation correlate with reward-related ventral striatal dopamine release. J Neurosci 2008; 28: 14311–9.

Schultz W. Predictive reward signal of dopamine neurons. J Neurophysiol 1998; 80: 1–27.

Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science 1997; 275: 1593–9.

Sekine Y, Suzuki K, Ramachandran PV, Blackburn TP, Ashby CR Jr. Acute and repeated administration of fluoxetine, citalopram, and paroxetine significantly alters the activity of midbrain dopamine neurons in rats: an in vivo electrophysiological study. Synapse 2007; 61: 72–7.

Servaas MN, Riese H, Renken RJ, Wichers M, Bastiaansen JA, Figueroa CA, et al. Associations between daily affective instability and connectomics in functional subnetworks in remitted patients with recurrent MDD. Neuropsychopharmacology 2017; 42: 2583–92.

Shumake J, Gonzalez-Lima F. Functional opposition between habenula metabolism and the brain reward system. Front Hum Neurosci 2013; 7: 662.

Smoski MJ, Felder J, Bizzell J, Green SR, Ernst M, Lynch TR, et al. fMRI of alterations in reward selection, anticipation, and feedback in MDD. J Affect Disord 2009; 118: 69–78.

Snaith RP, Hamilton M, Morley S, Humayan A, Hargreaves D, Trigwell P. A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. Br J Psychiatry 1995; 167: 99–103.

Stringaris A, Vidal-Ribas Belil P, Artiges E, Lemaitre H, Gollier-Briant F, Wolke S, et al. The brain's response to reward anticipation and depression in adolescence: dimensionality, specificity, and longitudinal predictions in a community-based sample. Am J Psychiatry 2015; 172: 1215–23.

Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. Science 2005; 307: 1642–5.

Treadway MT, Zald DH. Reconsidering anhedonia in depression: lessons from translational neuroscience. Neurosci Biobehav Rev 2011; 35: 537–55.

Turner BM, Forstmann BU, Wagenmakers EJ, Brown SD, Sederberg PB, Steyvers M. A Bayesian framework for simultaneously modeling neural and behavioral data. Neuroimage 2013; 72: 193–206.

Turner BM, Rodriguez CA, Norcia TM, McClure SM, Steyvers M. Why more is better: simultaneous modeling of EEG, fMRI, and behavioral data. Neuroimage 2016; 128: 96–115.

Ubl B, Kuehner C, Kirsch P, Ruttorf M, Flor H, Diener C. Neural reward processing in individuals remitted from major depression. Psychol Med 2015; 45: 3549–58.

van Ravenzwaaij D, Provost A, Brown SD. A confirmatory approach for integrating neural and behavioral data into a single model. J Math Psychol 2017; 76: 131–41.

Vrieze E, Pizzagalli DA, Demyttenaere K, Hompes T, Sienaert P, de Boer P, et al. Reduced reward learning predicts outcome in MDD. Biol Psychiatry 2013; 73: 639–45.

Wichers M, Barge-Schaapveld DQ, Nicolson NA, Peeters F, de Vries M, Mengelers R, et al. Reduced stress-sensitivity or increased reward experience: the psychological mechanism of response to antidepressant medication. Neuropsychopharmacology 2009; 34: 923–31.

Wichers M, Lothmann C, Simons CJ, Nicolson NA, Peeters F. The dynamic interplay between negative and positive emotions in daily life predicts response to treatment in depression: a momentary assessment study. Br J Clin Psychol 2012; 51: 206–22.

Wichers M, Myin-Germeys I, Jacobs N, Peeters F, Kenis G, Derom C, et al. Evidence that moment-to-moment variation in positive emotions buffer genetic risk for depression: a momentary assessment twin study. Acta Psychiatr Scand 2007; 115: 451–7.

Wichers M, Peeters F, Geschwind N, Jacobs N, Simons CJ, Derom C, et al. Unveiling patterns of affective responses in daily life may improve outcome prediction in depression: a momentary assessment study. J Affect Disord 2010; 124: 191–5.

Wilson RC, Niv Y. Is model fitting necessary for model-based fMRI? PLoS Comput Biol 2015; 11: e1004237.

Zhao H, Zhang BL, Yang SJ, Rusak B. The role of lateral habenula-dorsal raphe nucleus circuits in higher brain functions and psychiatric illness. Behav Brain Res 2015; 277: 89–98.