

Abnormal temporal difference reward-learning signals in major depression

P. Kumar,¹ G. Waiter,^{2,3} T. Ahearn,^{2,3} M. Milders,⁴ I. Reid¹ and J. D. Steele¹

¹Department of Mental Health, ²Department of Medical Physics, ³Department of Radiology and ⁴Department of Psychology, University of Aberdeen, Aberdeen, AB25 2ZH, UK

Correspondence to: Dr J. D. Steele, MD, PhD, M.R.C.Psych., Clinical Research Centre, Royal Cornhill Hospital, Aberdeen, AB25 2ZH, UK

E-mail: d.steele@abdn.ac.uk

Anhedonia is a core symptom of major depressive disorder (MDD), long thought to be associated with reduced dopaminergic function. However, most antidepressants do not act directly on the dopamine system and all antidepressants have a delayed full therapeutic effect. Recently, it has been proposed that antidepressants fail to alter dopamine function in antidepressant unresponsive MDD. There is compelling evidence that dopamine neurons code a specific phasic (short duration) reward-learning signal, described by temporal difference (TD) theory. There is no current evidence for other neurons coding a TD reward-learning signal, although such evidence may be found in time. The neuronal substrates of the TD signal were not explored in this study. Phasic signals are believed to have quite different properties to tonic (long duration) signals. No studies have investigated phasic reward-learning signals in MDD. Therefore, adults with MDD receiving long-term antidepressant medication, and comparison controls both unmedicated and acutely medicated with the antidepressant citalopram, were scanned using fMRI during a reward-learning task. Three hypotheses were tested: first, patients with MDD have blunted TD reward-learning signals; second, controls given an antidepressant acutely have blunted TD reward-learning signals; third, the extent of alteration in TD signals in major depression correlates with illness severity ratings. The results supported the hypotheses. Patients with MDD had significantly reduced reward-learning signals in many non-brainstem regions: ventral striatum (VS), rostral and dorsal anterior cingulate, retrosplenial cortex (RC), midbrain and hippocampus. However, the TD signal was increased in the brainstem of patients. As predicted, acute antidepressant administration to controls was associated with a blunted TD signal, and the brainstem TD signal was not increased by acute citalopram administration. In a number of regions, the magnitude of the abnormal signals in MDD correlated with illness severity ratings. The findings highlight the importance of phasic reward-learning signals, and are consistent with the hypothesis that antidepressants fail to normalize reward-learning function in antidepressant-unresponsive MDD. Whilst there is evidence that some antidepressants acutely suppress dopamine function, the long-term action of virtually all antidepressants is enhanced dopamine agonist responsiveness. This distinction might help to elucidate the delayed action of antidepressants. Finally, analogous to recent work in schizophrenia, the finding of abnormal phasic reward-learning signals in MDD implies that an integrated understanding of symptoms and treatment mechanisms is possible, spanning physiology, phenomenology and pharmacology.

Keywords: major depressive disorder; temporal difference signals; reward-learning; citalopram

Abbreviations: dAC = dorsal anterior cingulate; rAC = rostral anterior cingulate; RC = retrosplenial cortex; MDD = major depressive disorder; TD = temporal difference; BDI = Beck depression inventory; BOLD = blood oxygen level dependent; FDR = false discovery rate

Received March 25, 2008. Revised May 29, 2008. Accepted June 2, 2008. Advance Access publication June 25, 2008

Introduction

Central features of major depressive disorder (MDD) are anhedonia, disturbances of motivation, psychomotor speed and concentration (Ebmeier *et al.*, 2006). These functions

are regulated in part by the dopamine (DA) system. Convergent evidence from neuroimaging, post-mortem, behavioural and pharmacological studies, points to reduced DA function in MDD (Pizzagalli *et al.*, 2005; Ebmeier *et al.*, 2006;

Dunlop and Nemeroff, 2007; Gershon *et al.*, 2007; Steele *et al.*, 2007b). Despite this, most work to date has focused on the serotonergic (5-HT) and noradrenergic systems (Dunlop and Nemeroff, 2007). However, it is increasingly recognized that many patients with MDD do not achieve full remission (APA, 2000; Ebmeier *et al.*, 2006; Dunlop and Nemeroff, 2007). Antidepressants may fail to alter DA function in treatment-resistant illness (Dunlop and Nemeroff, 2007; Gershon *et al.*, 2007). Further studies on the contribution of anhedonia and hypothesized abnormalities of the DA system to the pathophysiology of MDD are required to improve outcomes for patients with treatment-unresponsive illness (Dunlop and Nemeroff, 2007; Gershon *et al.*, 2007).

Considerable recent work has demonstrated that DA neurons code a highly specific ‘phasic’ (brief duration) reward-learning signal, described by temporal difference (TD) theory (Montague *et al.*, 1996; Dayan and Abbott, 2001), which allows the formation of stimulus-outcome associations. There is compelling animal (Montague *et al.*, 1996; Schultz, 2002) and human neuroimaging (McClure *et al.*, 2003a; O’Doherty *et al.*, 2004; Tobler *et al.*, 2006; Seymour *et al.*, 2007) evidence linking the TD reward-learning signal with DA function. To date, no other neuronal type has been reported to exhibit the specific TD reward-learning signal (Schultz and Dickinson, 2000), although an aversive TD learning signal exists in humans, perhaps serotonin related (Seymour *et al.*, 2005, 2007). The TD mechanism is believed to contribute to the attribution of ‘incentive salience’, the process by which a stimulus grasps attention and motivates goal-directed behaviour by associations with reinforcing events (Berridge, 2007; McClure *et al.*, 2003b; Robbins and Everitt, 2007). Abnormally reduced TD reward-learning signals in MDD would imply reduced salience of, and attention to, rewarding events, such as occurs in anhedonia, and form a link between core phenomenology and physiology.

A difficulty in interpreting a reduction in DA activity in MDD is that many antidepressants do not act directly on DA (Dunlop and Nemeroff, 2007). In contrast, almost all misused ‘recreational’ drugs do enhance DA release (Bardo, 1998; McBride *et al.*, 1999) and are misused because of their acute mood enhancing effects (Solomon, 1977). Although mood elevation occurs with short-term use, repeated longer-term use is often associated with anhedonia and anxiety (Solomon, 1977; McIntosh and Ritson, 2001). Stimulants are not particularly effective antidepressants (Satel and Nelson, 1989). In contrast, treatment of MDD frequently involves 5-HT increasing medications such as selective serotonergic reuptake inhibitors (SSRIs). This might appear surprising, as there is evidence for acute 5-HT increases inhibiting reward-related approach behaviours and being associated with anxiety (Graeff, 1993; Daw *et al.*, 2002). Clinically though, if an antidepressant is taken for several weeks, reduced anhedonia and anxiety with recovery from MDD often occurs (Taylor *et al.*, 2006). All antidepressants

are believed to work with a similar time course, and whilst there may be some symptomatic improvement by the end of the first week, recovery continues for at least 6 weeks (Taylor *et al.*, 2006). These observations suggest a common interaction between DA and 5-HT systems, with different effects of *acute* (DA system inhibition) versus *chronic* (DA system enhancement) antidepressant administration.

Here, a TD reward-learning approach was used, as considerable animal and human experimental evidence links the TD signal to DA neuronal activity, and not other neuronal types. Phasic TD reward-learning signals have not been previously investigated in MDD. Instead, ‘tonic’ (long timescale) measurements of DA in humans have been reported (Dunlop and Nemeroff, 2007), and in the few cases where phasic signals have been investigated (Elliott *et al.*, 1998; Steele *et al.*, 2007b), *learning* has not been investigated. Phasic reward-learning signals have very different properties compared with tonic DA signals (Daw *et al.*, 2002; Niv *et al.*, 2007). Importantly, long timescale DA measurement methods (e.g. receptor-binding studies, endocrine response trials) would not be expected to detect hypothesized phasic reward-learning signal abnormalities.

Therefore, using TD modelling, a Pavlovian task and event-related fMRI, we tested three main hypotheses: first, antidepressant-unresponsive MDD is associated with reduced TD reward-learning signals; second, healthy controls given an SSRI *acutely* have reduced TD reward signals; third, abnormal TD reward-learning signals in antidepressant-unresponsive MDD correlates with illness severity ratings, independent of medication status. SSRIs were given acutely to test the hypothesis of suppression of reward-learning signals (Daw *et al.*, 2002). In contrast, we hypothesized that a common effect of diverse antidepressants administered chronically is TD reward-learning signal enhancement, this enhancement being associated with clinical recovery and resolution of anhedonia. Consequently, patients who recovered the most with antidepressant administration were expected to have least blunting of TD reward-learning signals, reflecting partial recovery.

Materials and Methods

Subjects

Permission for the study was obtained from the local ethics committee and written informed consent obtained from all subjects. Data were obtained from patients with a DSM IV diagnosis of (unipolar) MDD without comorbidity, in the opinion of both the treating consultant and one of the authors (J.D.S.). Controls were matched on the basis of age, sex and National Adult Reading Test (Nelson and Wilson, 1991). Controls participated twice: once in an unmedicated state, and once after receiving the SSRI citalopram at a dose of 20 mg/day for 3 days. The order of participating in a medicated or unmedicated state was counter-balanced. As we aimed to investigate the effects of acute and not chronic antidepressant administration, 3 days and not several weeks of administration, was used. Three self-ratings were

completed by each subject: Beck depression inventory (BDI, Beck *et al.*, 1961), Spielberger state anxiety (Spielberger, 1983) and Snaith–Hamilton hedonia, a low score indicating anhedonia (Snaith *et al.*, 1995). A Hamilton-21 depression rating (Hamilton, 1960) was obtained as a measure of MDD illness severity. The same rater (J.D.S.) assessed all patients in the morning just before scanning.

Patients had a duration of symptomatic illness >3 months despite continuous antidepressant treatment and medications were stable for 1 month before scanning. ‘Unresponsive’ illness was defined as a Beck or Hamilton depression score greater than 21 on initial assessment, which was usually a few days before scanning. All recorded ratings were done in the morning, immediately before scanning. Patient medications as total dose per day were: escitalopram 15 mg, imipramine 200 mg, phenelzine 45–90 mg, trazodone 300 mg, mirtazapine 30–60 mg, venlafaxine 150–225 mg, amitriptyline 200 mg, lithium carbonate 600–800 mg (as antidepressant augmentation), citalopram 20 mg, fluoxetine 40 mg and sertraline 25–150 mg. Although such patients were deemed ‘unresponsive’ to the particular antidepressant they were receiving, they were not assumed to be ‘treatment-resistant’ as they had not received (and failed) extensive protocolized treatment trials (e.g. see Steele *et al.*, 2008). As it was hypothesized that treatment unresponsive patients would have reduced reward-learning signals independent of the particular antidepressant they were receiving, patients receiving different antidepressants were recruited.

Subject exclusion criteria were any other DSM IV Axis I or II diagnosis including personality disorder, a history of substance or alcohol misuse, structural brain abnormality, neurological disorder, use of non-antidepressant medication which might alter brain metabolism and ECT within the last few months. Subjects with claustrophobia were excluded as they were considered unlikely to tolerate scanning. Patients with other anxiety symptoms were not excluded and a dimensional measure of state anxiety was obtained as earlier.

Experimental task

A Pavlovian reward-learning paradigm was used, as the standard TD model describes DA function during such learning (Schultz, 2002). Subjects were asked to abstain from drinking fluids from the night before the scan to ensure they were thirsty. This is a routine requirement for many types of medical procedure and does not cause detectable biochemical alteration. Just before scanning subjects were told: ‘After either of the pictures drops of water may be delivered. You should try and learn which picture predicts the water. The picture which predicts the water may change’. The goal of the task was therefore made explicit.

As shown in Fig. 1, there were 100 trials of 6 s duration each. Two seconds after the start of each trial, one of two fractal pictures (conditioned stimulus, CS) were presented (random order) for 1 s, then 4 s after the start of each trial, 0.1 ml of water (unconditioned stimulus, US) was delivered according to a probabilistic pattern. This volume was chosen empirically as subjects could clearly perceive the water, yet it minimized the need for swallowing and so the risk of inducing movement artefacts. Immediately after scanning, subjects completed linear analogue scales of perceived pleasantness of the delivered water and reportable association between pictures and water delivery, for the first and last blocks. The null hypothesis of no difference in water pleasantness between groups was tested using *t*-tests. The null

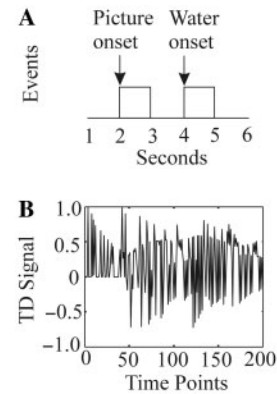


Fig. 1 (A) Pavlovian task timing diagram for presentation of picture stimuli and water delivery. (B) Typical predicted TD signal for a subject. Consecutive pairs of conditioned and unconditioned stimuli time points for the 100 trials are shown.

hypothesis of no difference in accuracy of association was tested using two Pearson’s 3×2 Chi-square tests, one test for each of the two blocks.

Water delivery was via a polythene tube attached to an electronic syringe pump (World Precision Instruments Ltd, Stevenage, UK) positioned in the scanner control room and interfaced to the image presentation and log file generating computer. There were five blocks of 20 trials each with the following probabilities of water delivery: picture 1 (80%) picture 2 (0%), picture 1 (50%) picture 2 (20%), picture 1 (0%) picture 2 (90%), picture 1 (20%) picture 2 (20%), picture 1 (80%) picture 2 (0%). Pre-study pilot testing had indicated that subjects could not identify where the boundaries between the blocks were, due to the probabilistic nature of the water delivery and the few numbers of reinforced trials in each block. As controls were scanned twice, to avoid the effects of previous learning, a parallel version of the task was used with different pictures and slightly different probabilities of water delivery: picture 1 (80%) picture 2 (0%), picture 1 (50%) picture 2 (50%), picture 1 (80%) picture 2 (20%), picture 1 (0%) picture 2 (50%), picture 1 (80%) picture 2 (0%). The order of controls being scanned in a medicated versus unmedicated state was counterbalanced. Null hypotheses of no difference in rating scores were tested using paired and unpaired *t*-tests as indicated.

Two other tasks were done by the subjects during the scanning session. Both were done after the Pavlovian task to avoid a possible effect on the Pavlovian task results. Analyses of these tasks will be reported separately.

Image acquisition and pre-processing

The scanning session lasted 1 h during which subjects participated in the Pavlovian learning study and a T1 image was obtained to exclude structural brain abnormality. For blood oxygen level dependent (BOLD) response imaging, T_2^* -weighted gradient echo planar images were obtained using a GE Medical Systems Signa 1.5 T MRI scanner. A total of 30 axially orientated 5 mm thick contiguous sequential slices were obtained for each volume, 246 volumes being obtained with a TR of 2.5 s, TE 30 ms, flip 90° , FOV 240 mm and matrix 64×64 . The first four volumes were discarded to allow for transient effects. Image acquisition was asynchronous with respect to stimulus and feedback presentation events.

For pre-processing, image data was converted to Analyze format and SPM2 (Friston, 2004) was used for analysis. Images were slice time corrected then realigned to the first image in each time series. The average realigned image was used to derive parameters for spatial normalization to the SPM2 template, then the parameters applied to each image in the time series. The resultant time-series realigned and spatially normalized images were smoothed with an 8 mm Gaussian kernel.

Temporal difference learning model

Each subject's log file was used to extract the sequence and timing of the CS and US and used to calculate a predicted TD signal profile. We used a standard temporal difference model (Dayan and Abbott, 2001) which assumes a discrete number of states representing the US and CS, and the time between their presentation. The presence or absence of a CS at time t was coded in binary form in the stimulus representation vector $x_i(t)$ (Dayan and Abbott, 2001) from the timing of events in each subject's log file. The estimation of the value (V) of each state was

$$\hat{V}(t) = \sum_i w_i x_i(t)$$

where w_i were weights, updated on each trial as below. The TD error signal $\delta(t)$ was defined as

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

where $r(t)$ was the delivered reinforcement (coded conventionally as '1' for water delivery, '0' for no-water delivery and all other time points) obtained from each subject's log file, and γ a discount factor which determined how less important later reinforcers were, compared with earlier reinforcers. Learning occurred by updating the weights on each trial as

$$\Delta w_i = \alpha \sum x_i(t) \delta(t)$$

where α was the learning rate. Each trial was assumed to consist of six time points. As associations were learned, the TD error signal moved 'backwards in time' from the US to the time of the CS, and when associations changed the error signal moved forwards again to the time of the US (Dayan and Abbott, 2001). The learning rate α and the discount factor γ have to be chosen. As in previous studies, $\gamma = 1.0$ and $\alpha = 0.1$ were used (O'Doherty *et al.*, 2006). The effects of other plausible choices were also investigated. The TD model used the same set of parameters across all subjects and all groups, since the image analysis tested the null hypothesis of no difference between groups (Pessiglione *et al.*, 2006).

Image analysis

For image analysis an event-related random-effects design was used. For first level analysis, for each subject, the covariate of interest was the event times multiplied by the predicted TD error signal, the result convolved with the SPM2 haemodynamic response function. The covariates of no interest were: the picture and water delivery event onsets convolved with the haemodynamic response function, six motion realignment terms to allow for any residual movement artefacts not removed by pre-processing realignment, and a constant term modelling the baseline of unchanged neural activity. For each subject, the first level covariate image of interest was the SPM2 'beta' image: an estimate of the 'strength' of the observed TD signal defined as the linear

regression coefficient (TD-LRC; positive or negative) between the predicted TD signal and observed BOLD signal, at each voxel.

Three second-level analyses were done. The first consisted of testing the null hypothesis of no TD signal. This was done by taking the covariate of interest ('beta') images for each subject, for each of the three data groups (unmedicated controls, medicated controls, patients), and entering them into three one-group t -tests. As is conventional, images were thresholded at $P < 0.001$ uncorrected to display the spatial extent of (de)activations and significant (de)activations defined as $P < 0.05$ 'whole brain' corrected using the false discovery rate (FDR) method implemented for SPM2. Based on the previous studies (O'Doherty *et al.*, 2004; Steele *et al.*, 2007a, b), *a priori* defined regions of interest were: rostral anterior cingulate (rAC), dorsal AC (dAC), ventral striatum (VS), amygdala, hippocampus, insula, ventral tegmental area (VTA) and retrosplenial cortex (RC). These regions have been repeatedly reported to have abnormal function in neuroimaging studies of major depression, be associated with emotion in healthy subjects and/or exhibit a TD signal in humans. As we had clear *a priori* expectation as to which regions might exhibit TD signal change of interest, an exploratory analysis was not done. Small volume corrections (SVCs) to significance levels for one group t -tests were also avoided, to set a more stringent level of significance. Where significant (de)activations were identified in *a priori* defined regions of interest, coordinates of maximal TD signal strength were identified. Brain 'activations' therefore consisted of regions where the observed BOLD had a significant positive TD-LRC across subjects, 'deactivations' consisted of regions with a significant negative TD-LRC across subjects.

The second (second level) analyses consisted of testing the null hypothesis of no difference between unmedicated control and patient groups, and unmedicated and medicated control groups, using (paired or unpaired as indicated) two group t -tests. Images were thresholded at $P < 0.05$ uncorrected to display the spatial extent of the signal and significance defined as $P < 0.05$ FDR corrected using 10 mm diameter SVCs centred at the locations identified from the one-group t -tests. The third (second level) analysis consisted of testing the null hypothesis, for only the patient group, of no correlation between the observed TD signal strength (TD-LRC) for each patient, and clinical ratings of anxiety, anhedonia and depression. Images were again thresholded at $P < 0.05$ uncorrected and significance defined as $P < 0.05$ FDR corrected using 10 mm diameter SVCs centred at the locations identified from the one-group t -tests.

As is usual in the imaging literature, SVCs were applied as independent corrections. Whilst this controls type 1 error less strongly than applying a SVC via a single mask consisting of all the regions of interest, it was justified on the basis that all the tests were planned *a priori*, and that this method helps to minimize the risk of type II error (which occurs with increased type I error control). Overall then, this method of 2nd level analyses focused only on a limited number of regions of particular interest where relatively strong one-group TD signals were observed, then tested for differences between groups and correlations with illness severity, for only these regions.

Results

Behavioural rating scales

Data were obtained from 15 medicated patients and 18 controls in an unmedicated state, but only 15 of the

same controls in a medicated state. One control declined medication and two discontinued medication due to possible transient side-effects: nausea, headache and nervousness. Table 1 shows the percentage of correct reports of picture–water association for the first and last blocks, anhedonia and anxiety rating values, and ratings of linear analogue perceived pleasantness of water. Mean ratings for

Table 1 Details of subjects

	P	U	M	Significance
Age (years)	45.3 ± 12.3	42.0 ± 12.8	41.7 ± 12	NS
Females/total	9/15	11/18	9/15	NS
NART	111.6 ± 8.4	113.8 ± 8.2	114.0 ± 8.4	NS
WVP	77.6 ± 24.7	74.8 ± 22.9	75.9 ± 22.5	NS
First, Last	86, 27	72, 55	100, 73	NS
BDI	22.9 ± 8.2	3.0 ± 2.8	3.0 ± 2.7	$P < 0.0001^*$
SP	54.6 ± 11.5	30.2 ± 10.3	30.8 ± 9.2	$P < 0.0001^*$
SH	35.0 ± 6.7	51.7 ± 4.3	51.5 ± 4.6	$P < 0.0001^*$
Hamilton	23.2 ± 5.3			

Values are mean ± SD. NS = difference not significant; U = control subjects in an unmedicated state; M = control subjects in a medicated state; P = patients; NART = National Adult Reading Test; WVP = water pleasantness rating as percentage; 'First, Last' = average percentage of correctly reported picture–water associations for first and last blocks; SH = Snaith–Hamilton hedonia score; SP = Spielberger anxiety scale; 'Hamilton' = Hamilton depression rating scale; * = significant difference between patients versus controls, but not a significant difference between controls when unmedicated versus medicated.

control subjects in a medicated versus unmedicated state did not differ significantly. As expected, patients rated themselves more anhedonic, anxious and depressed than controls. Regarding water pleasantness ratings, patients did not differ from unmedicated controls ($t = 0.27$, $P = 0.39$), nor controls in an unmedicated compared with a medicated state ($t = 0.07$, $P = 0.42$). For the three groups, there were no significant differences in accuracy of reporting picture–water associations for the first block ($\chi = 5.095$, $P = 0.305$). Although a smaller fraction of patients correctly reported the associations for the last block, the difference was not significant, ($\chi = 5.156$, $P = 0.076$); however a trend is suggested. Overall these results are consistent with a similar learning effort across groups.

Within-group TD signal analyses

Figure 1 shows a typical predicted TD reward-learning signal. The complex pattern is due to learning, and re-learning, the changing associations between the CS and US. The results of the three one-group TD signal analyses are shown in Fig. 2. For unmedicated controls, significant activations were found in the bilateral VS, amygdala, caudate, dAC and thalamus. Additionally, significant deactivations were found in the rAC, RC and hippocampus. For controls in a medicated state, a smaller number of significant activation regions were identified in the bilateral amygdala and anterior insula. No regions of significant deactivation

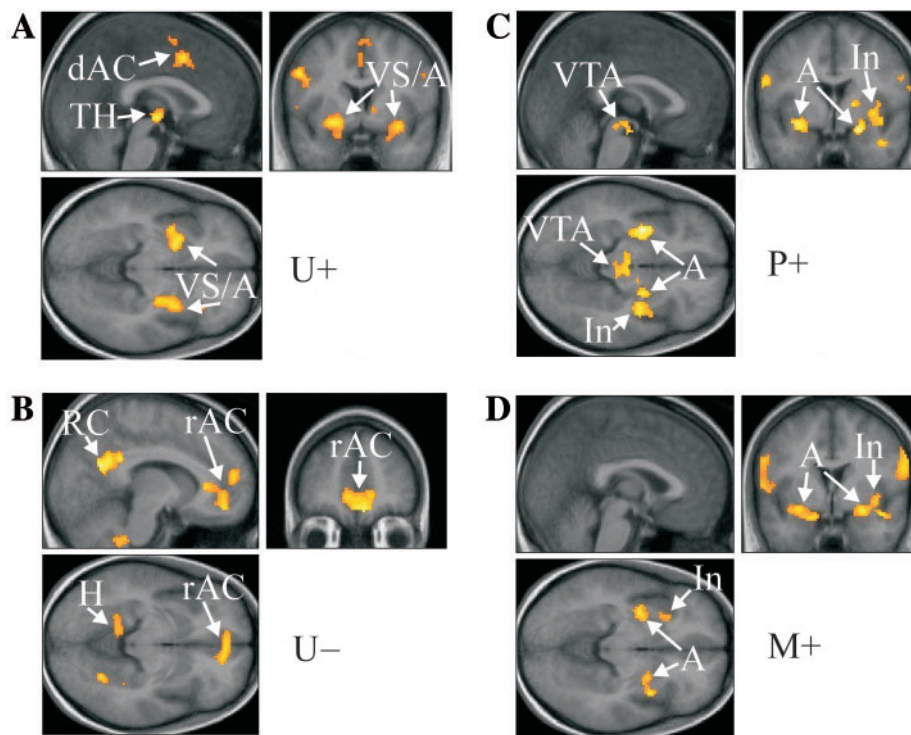


Fig. 2 (A) TD signal activation in unmedicated controls, U+. (B) TD deactivation in unmedicated controls, U-. (C) TD activation in patients, P+. (D) TD activation in medicated controls, M+. Images thresholded at $P < 0.001$ uncorrected, regions significant at $P < 0.05$ corrected. TH = thalamus; VS/A = ventral striatum/amygdala; H = hippocampus; In = insula; A = amygdala.

Table 2 Within group activation and deactivation

	Location	Coordinate	z	Significance*
U	Ventral striatum/pallidum	(-24,6,-10)	4.23	0.001
U	Ventral striatum/pallidum	(32,2,-12)	4.14	0.001
U	Amygdala	(-20,0,-20)	3.88	0.018
U	Amygdala	(26,-2,-14)	3.85	0.018
U	Caudate	(10,8,0)	4.20	0.001
U	Dorsal anterior cingulate	(-4,10,46)	4.62	0.009
U	Thalamus	(-2,-14,-6)	4.44	0.009
U	Rostral/subgenual AC	(2,54,6)	-4.41	0.015
U	Rostral/subgenual AC	(15,42,-3)	-4.47	0.015
U	Retrosplenial cortex	(-4,-60,26)	-4.83	0.012
U	Retrosplenial cortex	(9,-46,31)	-4.38	0.016
U	Hippocampus	(-17,-46,-10)	-3.44	0.032
M	Amygdala	(-25,-4,-15)	4.16	0.016
M	Amygdala	(26,0,-14)	3.99	0.016
M	Anterior insula	(-32,16,4)	4.47	0.016
M	Anterior insula	(36,20,2)	4.50	0.016
P	Midbrain/VTA	(0,-21,-10)	3.93	0.054
P	Amygdala	(-25,-2,-14)	4.72	0.054
P	Amygdala	(22,-2,-16)	4.68	0.054
P	Anterior insula	(42,4,-10)	3.76	0.054

P = patients; U = unmedicated controls; M = medicated controls; AC = anterior cingulate; -z-value indicates deactivation with predicted TD signal; * = FDR whole brain corrected.

were observed. For patients, a significant VTA TD signal was observed, with signal present also in the amygdala and anterior insula. No regions of significant deactivation were identified. Table 2 summarizes activation and deactivation details.

Differences in TD signal between MDD and control groups

Patients with antidepressant-unresponsive MDD, when compared with unmedicated controls, had reduced TD signals in the VS and dAC. The TD signal appeared increased in the VTA, rAC, RC and hippocampus. However, only the VTA signal was actually increased. The apparent increases in the rAC, RC and hippocampus were due to a lack of deactivation in patients: i.e. the TD signal was blunted in these regions in MDD. Comparing patients with controls in a medicated state, patients had an increased signal in the VTA and rAC. Again though, only the VTA signal was actually increased, and the apparent increase in the rAC was due to a lack of deactivation in patients. Figure 3A shows these regions and Fig. 4 shows the TD signal effect sizes with 90% confidence intervals for these regions. Table 3 lists details of these differences.

For controls in a medicated compared with unmedicated state, the TD signal appeared significantly increased in the

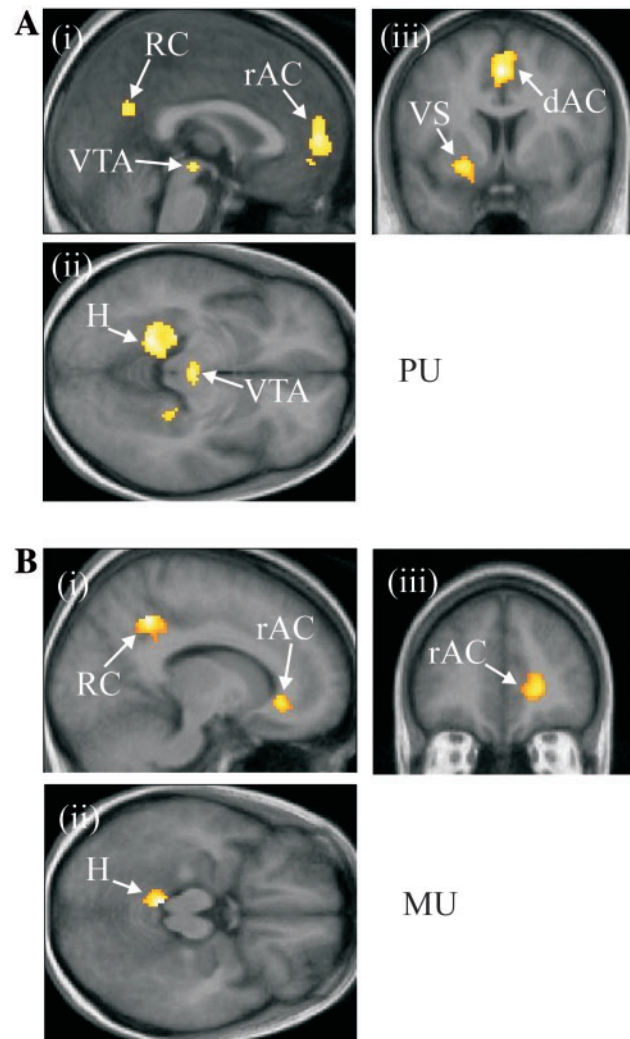


Fig. 3 (A) Difference in TD signal strength in patients compared with unmedicated controls, PU: blunted deactivation (i); blunted deactivation (ii); blunted activation (iii). (B) Difference in TD signal strength in medicated controls compared with unmedicated controls, MU: blunted deactivation (i), blunted deactivation (ii), blunted deactivation (iii). Regions significant at $P < 0.05$ corrected. H = hippocampus.

rAC, RC and hippocampus (Fig. 3B). However, as shown in Fig. 4, this was due to a lack of deactivation in the medicated state. Therefore, as hypothesized, the effect of acute medication administration was also to blunt the TD signal in these regions. Table 3 lists details of the significant between-group differences.

Correlations between TD signal and MDD severity ratings

Significant correlations between clinical ratings of MDD severity and the observed strength of the TD signal (TD-LRC) are summarized in Table 4 and illustrated in Fig. 5. Interpretation of the correlations depends on whether a region was an activation or deactivation.

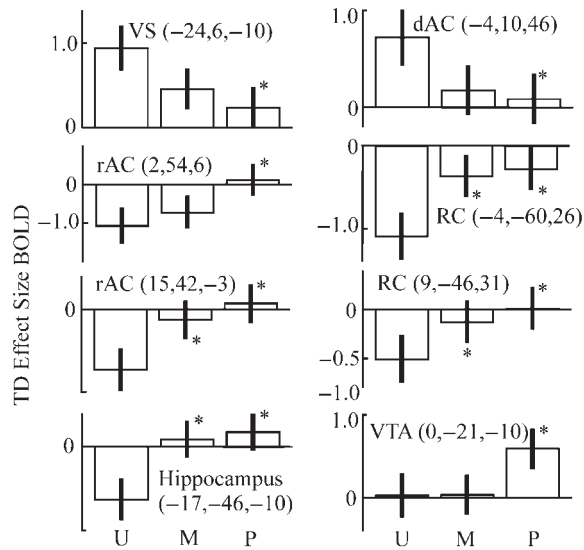


Fig. 4 Observed TD signal effect sizes with 90% confidence intervals for patients (P), unmedicated (U) and medicated (M) controls. *significant difference compared with unmedicated controls.

Table 3 Between group comparisons

	Location	Coordinate	z	Significance*
PU	Ventral striatum	(-24,6,-10)	-2.51	0.046
PU	Dorsal anterior cingulate	(-4,10,46)	-3.06	0.013
PU	Rostral/subgenual AC	(2,54,6)	3.40	0.004
PU	Retrosplenial cortex	(-4,-60,26)	3.05	0.011
PU	Midbrain	(0,-21,-10)	3.09	0.014
PU	Hippocampus	(-17,-46,-10)	3.47	0.002
PM	Midbrain	(0,-21,-10)	2.95	0.026
PM	Rostral/subgenual AC	(2,54,6)	2.89	0.032
MU	Rostral anterior cingulate	(15,42,-3)	3.49	0.009
MU	Retrosplenial cortex	(9,-46,31)	2.61	0.055
MU	Hippocampus	(-17,-46,-10)	2.58	0.050

PU = patients compared with unmedicated controls; PM = patients compared with medicated controls; MU = controls in a medicated state compared with unmedicated state; -z-value indicates a relative deactivation for the contrast of interest; * = FDR small volume corrected.

A significant VTA activation was observed in the MDD group and patients had a significantly stronger TD signal (larger positive TD-LRC) than unmedicated or medicated controls. Consistent with this, more severe MDD, defined by Hamilton, BDI and Spielberger ratings, had the strongest VTA TD signals (Fig. 5). A significant hippocampal deactivation was present in unmedicated controls, the magnitude of which was significantly less in patients. Consequently, the apparently increased hippocampal activity in MDD was due to a blunted deactivation (Fig. 4). Consistent with this, a weaker TD signal (larger positive

Table 4 TD signal correlations with MDD severity ratings

Location	Rating scale	Coordinate	z	Significance*
VTA	Hamilton	(0,-21,-10)	3.21	0.017
VTA	Spielberger	(0,-21,-10)	3.03	0.019
VTA	BDI	(0,-21,-10)	3.44	0.004
rAC	Spielberger	(0,50,4)	2.31	0.043
Hippocampus	Hamilton	(-17,-46,-10)	2.83	0.032
Amygdala	Snaith-Hamilton	(-25,-2,-14)	-2.50	0.047
Amygdala	Snaith-Hamilton	(22,-2,-10)	-2.80	0.040

-z-value indicates a negative correlation between the observed BOLD TD signal strength and rating of MDD illness severity; * = FDR small volume corrected.

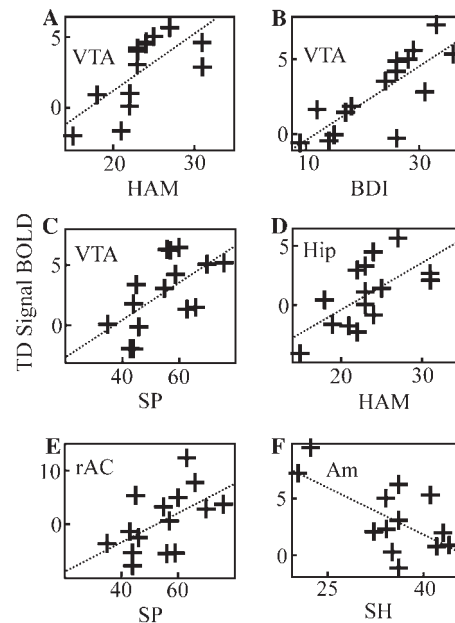


Fig. 5 Correlations between TD signal strength and major depression severity ratings. Regions significant at $P < 0.05$ corrected. Best fit linear regression lines also shown. HAM = Hamilton scale; Hip = Hippocampus; SP = Spielberger anxiety scale; Am = Amygdala; SH = Snaith-Hamilton anhedonia scale.

TD-LRC) was associated with more severe MDD, as defined by Hamilton rating (Fig. 5). A significant rAC deactivation was present in unmedicated controls, the magnitude of which was significantly less in MDD. Again consistent with this, more severe MDD defined by Spielberger rating was associated with a weaker TD signal (larger positive TD-LRC). More severe MDD defined by Snaith-Hamilton anhedonia score was associated with significantly stronger amygdala TD signals (larger positive TD-LRC). No significant correlations were found for control data.

Stability of TD modelling and choice of learning parameters

Comparing the TD estimates of brain activity for learning rate 0.1 versus 0.4, and a discount factor of 1.0 versus 0.4,

no significant differences were found. This indicates that the results were not due to an idiosyncratic choice of learning parameters. Similar stability has been reported previously (O'Doherty *et al.*, 2003). Investigation of the effects of different learning parameters in relation to the observed group differences will be described in a future report.

Discussion

Our first hypothesis was that patients with antidepressant-unresponsive MDD had reduced reward-learning signals. If the observed TD reward-learning signals were a direct consequence of DA neuronal firing, then the finding of blunted TD reward-learning signals in the VS, rAC, dAC, RC and hippocampus supports this hypothesis, but not the VTA, as that signal was enhanced. A challenge to the 'direct' interpretation is the observation of both activations and deactivations described by the predicted TD signal. It has been suggested that phasic DA neuronal firing leads to DA release, which facilitates some form of longer duration post-synaptic activity, such as post-synaptic potentiation and inhibition (Menon *et al.*, 2007). Such longer duration post-synaptic DA-mediated responses could be the basis of the BOLD signal that correlates with the predicted TD signal (Menon *et al.*, 2007). This could parsimoniously account for both the activations and deactivations described by the predicted TD signal: the same widespread DA axonal projections, originating from the same VTA cell bodies, mediate all the observed signals. The increased VTA signal might have been a compensatory response to blunted reward-learning signals in non-brainstem regions, or as discussed later, an effect of medication. In a preliminary analysis, we have recently replicated the findings of increased brainstem TD signals and blunted non-brainstem TD signals in MDD using a different paradigm which will be reported separately.

Two neuroimaging studies of MDD have reported reduced VS activation and behavioural blunting associated with feedback of performance in a pseudo-instrumental gambling task (Steele *et al.*, 2007b) and various cognitive tasks (Elliott *et al.*, 1998). These findings were interpreted as reflecting a blunted response to rewarding feedback information implying reduced DA activity (Elliott *et al.*, 1997; Steele *et al.*, 2007b). In one study, behavioural blunting correlated with Snaith–Hamilton anhedonia (Steele *et al.*, 2007b). Here, we also report apparently increased activity in the rAC and hippocampus of patients with MDD, due to blunted phasic deactivation in patients but not controls. This is also consistent with previous work (Steele *et al.*, 2004). Whilst consistent, previous studies did not investigate the effects of reward-learning, using an established computational model of dopamine function.

Our second hypothesis was that acute administration of an SSRI to controls reduces TD reward-learning signals. In acutely medicated controls, the TD signal was found to be significantly blunted in the rAC, RC and hippocampus. Figure 4 suggests non-significant consistent trends for

other regions. As expected, the pattern of TD signal blunting with *acute* SSRI administration closely resembled the pattern for unresponsive MDD, although not as marked, in that medicated controls typically had a TD signal effect size intermediate between the unmedicated state and MDD. In patients, the VTA, hippocampus and rAC correlations with MDD severity ratings indicates that the TD signal differences between MDD and unmedicated controls should be reduced or absent in antidepressant *responsive* MDD. It is important to note that these correlations were present despite patients receiving diverse medications, suggesting observed differences between patients and unmedicated controls were unrelated to medication. Therefore, consistent with our second hypothesis, acute administration of an SSRI to controls reduced TD reward-learning signals, and consistent with our third hypothesis, regions of TD signal abnormality in antidepressant-unresponsive MDD correlated with ratings of illness severity.

Reduced TD signals suggest that either the neural TD signals were actually reduced, or that the chosen theoretical model did not match the observed BOLD as well as for regions where a higher signal was found. Having found abnormal TD signals it is important to try to account for them. It is possible that an alteration in the theoretical TD model for patients (e.g. altered learning rate) might match the data better such that the observed differences in TD signal strength between groups disappears. We are currently investigating this. However, this hypothesis cannot account for simultaneously decreased and increased TD signals, if it is assumed that the same model parameters apply throughout the brain.

A reduction in DA function with acute citalopram administration to controls is consistent with a number of animal studies, and an opponency (mutual inhibition) between 5-HT and DA has long been suggested (Di Mascio *et al.*, 1998; Daw *et al.*, 2002). Considering citalopram in particular, *acute* administration to animals had no effect on the number of spontaneously active VTA DA neurons in one study (Sekine *et al.*, 2007) and produced a reduction in another (Prisco and Esposito, 1995). In contrast, *chronic* (21 days) of citalopram administration produced an increase in the number of spontaneously active VTA DA neurons (Sekine *et al.*, 2007). It is possible that the effect of a delayed increase in VTA activity would be an increase in non-brainstem TD signals (e.g. ventral striatum), for patients who respond to long-term administration of an antidepressant. It is important to note however that an increase in the number of spontaneously active VTA DA neurons is not necessarily the same as an increase in the strength of the specific TD signal. Furthermore, the correlations between VTA TD signals and illness severity ratings suggest an illness rather than medication effect. Further work is required to clarify the mechanism associated with VTA TD signal increase.

There is accumulating evidence that a common *long-term* effect of virtually all antidepressants is enhanced

motor-stimulant responses to DA agonists mediated via D_{2/3} VS receptors (Ebert *et al.*, 1996; Bonhomme and Esposito, 1998; Esposito, 2006; Dunlop and Nemeroff, 2007; Gershon *et al.*, 2007). In contrast, the *acute* effect of some of the same antidepressants may be DA suppression (Prisco and Esposito, 1995; Di Mascio *et al.*, 1998; Daw *et al.*, 2002; Esposito, 2006). A delayed enhancement of post-synaptic mesolimbic DA function may help explain the psychological consequences of MDD responsive to antidepressants. As a consequence of re-learning associations over time, in the context of increasing phasic DA function, reduced expectation of aversive events and increased expectation of rewarding events may occur.

As discussed earlier, TD theory is linked to the concept of incentive salience, which is the process by which a stimulus grasps attention and motivates goal-directed behaviour by associations with reinforcing events (McClure *et al.*, 2003b; Berridge, 2007; Robbins and Everitt, 2007). Reduced reward-learning signals imply reduced salience and attention to reward-learning stimuli, consistent with anhedonia. Patients with MDD who have been ill for some time are likely to incorporate their experience of abnormal salience into their larger cognitive schemas, as characterized by the cognitive theory of depression (Beck, 1979). Antidepressant administration could therefore remove the driving force for the illness, although full recovery would require prolonged psychological re-learning. It is otherwise difficult to explain the delayed resolution of cognitive distortions in MDD patients who respond to antidepressants but do not receive effective psychotherapy.

Although reduced TD signals in patients could imply reduced attention to reward-learning stimuli, the accuracy of verbal report of correct picture–water associations did not differ significantly between groups, although a trend was suggested for the last block. It is possible to incorporate a behavioural task into a Pavlovian task and changes in reaction time can provide more detailed information on learning. This was not done as we wished to avoid a possible behavioural confound with the Pavlovian task. Nevertheless, more detailed measures of behavioural learning might demonstrate a difference between groups. To address this, we have also obtained behavioural and imaging data from a probabilistic instrumental learning task. The analysis of that data will be presented separately.

Possible abnormalities of reward prediction error signals have been investigated for ketamine-induced psychotic symptoms in normal subjects (Corlett *et al.*, 2006) and schizophrenia (Murray *et al.*, 2007). Murray and colleagues study is of particular note as they also focused on the brainstem VTA predictive error signal in patients. They report an attenuated VTA response to reward prediction error, and an augmented response to neutral prediction error, in schizophrenia (Murray *et al.*, 2007). No brain regions of patients with schizophrenia were observed to have greater overall activation than controls (Murray *et al.*, 2007). In contrast, we report an overall increased VTA activation in MDD with correlations between abnormal

signal magnitude and indices of illness severity. Whilst it is encouraging that different abnormalities of VTA activity have been found in different psychiatric disorders, it is unclear if these results can be replicated and whether differences in task design influence the results. We are therefore obtaining additional data from patients with schizophrenia, and investigating the effects of different tasks.

As earlier, whilst there is extensive evidence supporting the hypothesis that phasic activity of DA neurons conforms to a TD model during reward-learning, it should not be assumed that only DA neurons exhibit this pattern of activity. Activations conforming to a TD model have also been reported for regions without strong DA innervation. Whilst it is encouraging that two recent human imaging studies have employed DA blocking and enhancing challenges and reported an alteration in predictive error signals in accordance with expectation (Pessiglione *et al.*, 2006; Menon *et al.*, 2007), our study did not explore this link. Consequently, the neuronal substrate of the reduced TD reward-learning signals in MDD, and particularly the increased VTA signal, remains unclear.

In summary, antidepressant-unresponsive MDD was associated with reduced phasic TD reward-learning signals in non-brainstem regions, consistent with a hypothesized failure of a delayed increase in DA system responsiveness (Dunlop and Nemeroff, 2007), acute administration of citalopram to controls reduced TD reward-learning signals, and blunting of TD signals in unresponsive MDD correlated with illness severity ratings. In contrast, the VTA TD signal was increased in patients, and the extent of this increase correlated with measures of illness severity. This is the first study to investigate hypothesized abnormal phasic reward-learning signals in MDD. The results suggest it may be possible to understand the therapeutic mechanism of action of antidepressants, and their failure in some patients, in a manner which links the biology, phenomenology and pharmacology of MDD. A similar approach to linking understanding of these three fields has been described for schizophrenia (Kapur, 2003; Smith *et al.*, 2007).

Acknowledgements

The work was supported by grants from the Miller McKenzie Trust and Scottish Office. We thank Peter Dayan for discussions on TD theory.

References

- APA. Practice guideline for the treatment of patients with major depression. Washington, DC: American Psychiatric Association; 2000.
- Bardo MT. Neuropharmacological mechanisms of drug reward: beyond dopamine in the nucleus accumbens. *Clin Rev Neurobiol* 1998; 12: 37–67.
- Beck AT. Cognitive theory of depression. New York: J. Wiley and Sons; 1979.
- Beck AT, Ward HC, Mendelson M. An inventory for measuring depression. *Arch Gen Psychiatry* 1961; 4: 561–71.
- Berridge KC. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl)* 2007; 191: 391–431.

- Bonhomme N, Esposito E. Involvement of serotonin and dopamine in the mechanism of action of novel antidepressant drugs: a review. *J Clin Psychopharmacol* 1998; 18: 447–54.
- Corlett PR, Honey GD, Aitken MR, Dickinson A, Shanks DR, Absalom AR, et al. Frontal responses during learning predict vulnerability to the psychotogenic effects of ketamine: linking cognition, brain activity, and psychosis. *Arch Gen Psychiatry* 2006; 63: 611–21.
- Daw ND, Kakade S, Dayan P. Opponent interactions between serotonin and dopamine. *Neural Netw* 2002; 15: 603–16.
- Dayan P, Abbott F. *Theoretical neuroscience: computational and mathematical modelling of neural systems*. Cambridge, MA: The MIT Press; 2001.
- Di Mascio M, Di Giovanni G, Di Matteo V, Prisco S, Esposito E. Selective serotonin reuptake inhibitors reduce the spontaneous activity of dopaminergic neurons in the ventral tegmental area. *Brain Res Bull* 1998; 46: 547–54.
- Dunlop BW, Nemeroff CB. The role of dopamine in the pathophysiology of depression. *Arch Gen Psychiatry* 2007; 64: 327–37.
- Ebert D, Feistel H, Loew T, Pirner A. Dopamine and depression—striatal dopamine D2 receptor SPECT before and after antidepressant therapy. *Psychopharmacology (Berl)* 1996; 126: 91–4.
- Ebmeier KP, Donaghey C, Steele JD. Recent developments and current controversies in depression. *Lancet* 2006; 367: 153–67.
- Elliott R, Sahakian BJ, Herrod JJ, Robbins TW, Paykel ES. Abnormal response to negative feedback in unipolar depression: evidence for a diagnosis specific impairment. *J Neurol Neurosurg Psychiatry* 1997; 63: 74–82.
- Elliott R, Sahakian BJ, Michael A, Paykel ES, Dolan RJ. Abnormal neural response to feedback on planning and guessing tasks in patients with unipolar depression. *Psychol Med* 1998; 28: 559–71.
- Esposito E. Serotonin-dopamine interaction as a focus of novel antidepressant drugs. *Curr Drug Targets* 2006; 7: 177–85.
- Friston K. Introduction to statistical parametric mapping. In: Frackowiak RSJ, editor. *Human brain function*. London: Academic Press; 2004.
- Gershon AA, Vishne T, Grunhaus L. Dopamine D2-like receptors and the antidepressant response. *Biol Psychiatry* 2007; 61: 145–53.
- Graeff FG. Role of 5-HT in defensive behavior and anxiety. *Rev Neurosci* 1993; 4: 181–211.
- Hamilton M. Rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960; 23: 56–62.
- Kapur S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry* 2003; 160: 13–23.
- McBride WJ, Murphy JM, Ikemoto S. Localization of brain reinforcement mechanisms: intracranial self-administration and intracranial place-conditioning studies. *Behav Brain Res* 1999; 101: 129–52.
- McClure SM, Berns GS, Montague PR. Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 2003a; 38: 339–46.
- McClure SM, Daw ND, Montague PR. A computational substrate for incentive salience. *Trends Neurosci* 2003b; 26: 423–8.
- McIntosh C, Ritson B. Treating depression complicated by substance misuse. *Adv Psychiatr Treat* 2001; 7: 357–64.
- Menon M, Jensen J, Vitcu I, Graff-Guerrero A, Crawley A, Smith MA, et al. Temporal difference modeling of the blood-oxygen level dependent response during aversive conditioning in humans: effects of dopaminergic modulation. *Biol Psychiatry* 2007; 62: 765–72.
- Montague PR, Dayan P, Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 1996; 16: 1936–47.
- Murray GK, Corlett PR, Clark L, Pessiglione M, Blackwell AD, Honey G, et al. Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol Psychiatry* 2007; 12: 46–77.
- Nelson HE, Wilson JR. *The revised national adult reading test – test manual*. Winsor: NFER-Wilson; 1991.
- Niv Y, Daw ND, Joel D, Dayan P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 2007; 191: 507–20.
- 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 J, 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.
- 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, Jahn AL, O’Shea JP. Toward an objective characterization of an anhedonic phenotype: a signal-detection approach. *Biol Psychiatry* 2005; 57: 319–27.
- Prisco S, Esposito E. Differential effects of acute and chronic fluoxetine administration on the spontaneous activity of dopaminergic neurones in the ventral tegmental area. *Br J Pharmacol* 1995; 116: 1923–31.
- Robbins TW, Everitt BJ. A role for mesencephalic dopamine in activation: commentary on Berridge (2006). *Psychopharmacology (Berl)* 2007; 191: 433–7.
- Satel SL, Nelson JC. Stimulants in the treatment of depression: a critical overview. *J Clin Psychiatry* 1989; 50: 241–9.
- Schultz W. Getting formal with dopamine and reward. *Neuron* 2002; 36: 241–63.
- Schultz W, Dickinson A. Neuronal coding of prediction errors. *Annu Rev Neurosci* 2000; 23: 473–500.
- Sekine Y, Suzuki K, Ramachandran PV, Blackburn TP, Ashby CR. Acute and repeated administration of fluoxetine, citalopram and paroxetine significantly alters activity of midbrain dopamine neurons in rats: an in vivo electrophysiological study. *Synapse* 2007; 61: 72–7.
- Seymour B, O’Doherty J P, Koltzenburg M, Wiech K, Frackowiak R, Friston K, et al. Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat Neurosci* 2005; 8: 1234–40.
- Seymour B, Singer T, Dolan R. The neurobiology of punishment. *Nat Rev Neurosci* 2007; 8: 300–11.
- Smith AJ, Li M, Becker S, Kapur S. Linking animal models of psychosis to computational models of dopamine function. *Neuropsychopharmacology* 2007; 32: 54–66.
- 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.
- Solomon RL. An opponent-process theory of acquired motivation: the affective dynamics of addiction. In: Maser JD, Seligman MEP, editors. *Psychopathology: experimental models*. San Francisco: W.H. Freeman; 1977. p. 66–103.
- Spielberger CD. *State-trait anxiety inventory for adults, STAI Form Y-1*. Palo-Alto: Mind Garden; 1983.
- Steele JD, Christmas D, Eljamel MS, Matthews K. Anterior cingulotomy for major depression: clinical outcome and relationship to lesion characteristics. *Biol Psychiatry* 2008; 63: 670–7.
- Steele JD, Currie J, Lawrie SM, Reid I. Prefrontal cortical functional abnormality in major depressive disorder: a stereotactic meta-analysis. *J Affect Disord* 2007a; 101: 1–11.
- Steele JD, Kumar P, Ebmeier KP. Blunted response to feedback information in depressive illness. *Brain* 2007b; 130: 2367–74.
- Steele JD, Meyer M, Ebmeier KP. Neural predictive error signal correlates with depressive illness severity in a game paradigm. *NeuroImage* 2004; 23: 269–80.
- Taylor MJ, Freemantle N, Geddes JR, Bhagwagar Z. Early onset of selective serotonin reuptake inhibitor antidepressant action: systematic review and meta-analysis. *Arch Gen Psychiatry* 2006; 63: 1217–23.
- Tobler PN, O’Doherty JP, Dolan RJ, Schultz W. Human neural learning depends on reward prediction errors in the blocking paradigm. *J Neurophysiol* 2006; 95: 301–10.