Neuroscience and Biobehavioral Reviews 55 (2015) 247–267

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review

Reinforcement learning in depression: A review of computational research

Chong Chen [a,*], Taiki Takahashi [b], Shin Nakagawa [a], Takeshi Inoue [a,1], Ichiro Kusumi [a]

[a] Department of Psychiatry, Hokkaido University Graduate School of Medicine, Sapporo 060-8638, Japan
[b] Department of Behavioral Science/Center for Experimental Research in Social Sciences, Hokkaido University, Sapporo 060-0810, Japan

ARTICLE INFO

Article history:
Received 23 October 2014
Received in revised form 20 April 2015
Accepted 4 May 2015
Available online 12 May 2015

Keywords:
Anhedonia
Computational psychiatry
Depression
Dopamine
Incentive salience
Learning rate
'Liking'
Model-free
Model-based
Motivation
Prediction error
Reinforcement learning
Reward sensitivity
Stress
'Wanting'

ABSTRACT

Despite being considered primarily a mood disorder, major depressive disorder (MDD) is characterized by cognitive and decision making deficits. Recent research has employed computational models of reinforcement learning (RL) to address these deficits. The computational approach has the advantage in making explicit predictions about learning and behavior, specifying the process parameters of RL, differentiating between model-free and model-based RL, and the computational model-based functional magnetic resonance imaging and electroencephalography. With these merits there has been an emerging field of computational psychiatry and here we review specific studies that focused on MDD.

Considerable evidence suggests that MDD is associated with impaired brain signals of reward prediction error and expected value ('wanting'), decreased reward sensitivity ('liking') and/or learning (be it model-free or model-based), etc., although the causality remains unclear. These parameters may serve as valuable intermediate phenotypes of MDD, linking general clinical symptoms to underlying molecular dysfunctions. We believe future computational research at clinical, systems, and cellular/molecular/genetic levels will propel us toward a better understanding of the disease.

Contents

* Corresponding author at: Department of Psychiatry, Hokkaido University Graduate School of Medicine, Kita 15, Nishi 7, Kita-ku, Sapporo 060-8638, Japan. Tel.: +81 11 706 5160; fax: +81 11 706 5081.
E-mail address: cchen@med.hokudai.ac.jp (C. Chen).
[1] Present address: Department of Psychiatry, Tokyo Medical School, Tokyo 160-0023, Japan.

http://dx.doi.org/10.1016/j.neubiorev.2015.05.005
0149-7634/

1. Depression and reinforcement learning

Due to its hallmark feature of anhedonia, the inability to experience reward, major depressive disorder (MDD) has been considered primarily a mood disorder (Beck and Alford, 2009; American Psychiatric Association, 2014). However, deficits in cognitive function (McDermott and Ebmeier, 2009; Gotlib and Joormann, 2010; Lee et al., 2012b; Snyder, 2013; Belzung et al., 2014b) and decision making (Treadway and Zald, 2011; Paulus and Yu, 2012; Must et al., 2013) also occur in MDD, which reduces response to treatment, increases the risk of relapse and recurrence, and impairs social adaptation. Recently several lines of research have suggested dysfunctional reinforcement learning (RL) in MDD. RL, the process of maximizing reward and minimizing loss by modifying the behavior as a consequence of experience with the environment, plays a central role in decision making (Sutton and Barto, 1998; Schultz, 2006; Doll et al., 2012; Glimcher and Fehr, 2013).

Firstly, patients with MDD perform poorly on the Iowa Gambling Task which relies on RL (Must et al., 2013). In this task, subjects are asked to choose cards from four decks followed by monetary reward or punishment. Two of the four decks are disadvantageous because even though picking from them may be followed by a big gain, a bigger punishment occasionally occurs which eventually leads to inferior outcomes. Subjects have to learn this underlying rule by trial and error experiences. Research has consistently found that MDD patients tend to choose from the disadvantageous decks (Must et al., 2013).

Secondly, research using the signal-detection task also suggests impaired RL in MDD (Ragland et al., 2009; Pizzagalli, 2014). Similarly subjects have to learn a hidden rule in a probabilistic reward task and signal-detection analysis has shown that healthy volunteers would eventually demonstrate a bias toward frequently rewarded choices. The bias is likely based on implicit learning as it is rather difficult to infer the hidden rule explicitly (Ragland et al., 2009). On the other hand, MDD patients generally fail to show this bias and the impairment is correlated with anhedonia symptoms (Ragland et al., 2009; Pizzagalli, 2014).

Thirdly, consistent with these behavioral dysfunctions, in reward related tasks fMRI research has consistently found a low striatal (but see Knutson et al., 2008. Experimental design may have contributed to this inconsistency, which will be discussed later in Section 3.4) and high medial prefrontal cortex (PFC, although this is still somewhat controversial) response during monetary anticipation and/or outcomes stages in MDD (Steele et al., 2007; Kumar et al., 2008; Gradin et al., 2011. See a review by Forbes and Dahl, 2012; Zhang et al., 2013; Kerestes et al., 2014; Pizzagalli, 2014). Notably, striatum and medial PFC are both implicated in RL (Daw et al., 2011; Doll et al., 2012; Garrison et al., 2013; Chase et al., 2015).

Therefore a close examination of RL in depression may provide promising new insights into the underlying behavioral, cognitive and neural pathophysiology of the disease.

2. The computational approach and its merits

In the past decades, arising from two fast-advancing (but overlapping) fields of computational neuroscience (Sutton and Barto, 1998; Schultz, 2006; Niv, 2009; Doll et al., 2012) and neuroeconomics (Hasler, 2012; Sharp et al., 2012; Glimcher and Fehr, 2013), the computational theory of RL has been a major framework accounting for decision making. Under this framework, subjects choose actions according to mathematical value functions, which define the expected value of each action. Value functions can be updated through historical trial and error experience (i.e. prediction errors (PEs), the difference between received and expected values) or by prospective planning based on internal cognitive maps or learned models of the environment. The former is known as model-free RL and is habitual and slow to change. The latter is known as model-based RL, is goal-directed, and allows subjects to update value functions more flexibly (Daw et al., 2005; Dayan, 2009; Doll et al., 2012; Daw and O'Doherty, 2013; Dolan and Dayan, 2013; Nakahara, 2014). In model-based RL, subjects may simulate the consequences of potential actions that they may choose and then use the hypothetical outcomes to update their value functions. Therefore it is also known as counterfactual thinking (Lee et al., 2012a). There are at least four advantages of this computational approach.

Firstly, the mathematical analysis of RL allows explicit testable predictions about learning and behavior, thus linking learning and behavior to psychological processes and providing a useful normative framework to study them (Niv, 2009; Wiecki et al., 2015).

Secondly, the computational approach specifies the process parameters of RL, such as PE, learning rate, reward sensitivity, and memory of previous reinforcement (which will be discussed in detail below). For instance, in the field of depression, this reduces the broadly defined symptom of anhedonia to more specific and refined constructs (Treadway and Zald, 2011; Der-Avakian and Markou, 2012. See our later discussion of 'wanting', 'liking' and learning in Section 3.4). This moves the field from general symptom-based descriptions to direct and precise identification of impaired latent functions (Wiecki et al., 2015).

Thirdly, the recent differentiation of model-free and model-based RL brings the two most important but separated fields, cognition and reward learning/decision making together. This enables novel synthesis and may generate valuable insights regarding learning and behavior. The differentiation also perfectly resembles popular dual-process theories of decision making, learning and memory, such as the models of automatic and controlled processing (Shiffrin and Schneider, 1977), feeling and thinking (Zajonc, 1982), experiential and cognitive processing (Epstein, 1994), system 1 and system 2 (Kahneman, 2011), habitual and goal-directed control (Dickinson and Charnock, 1985; Balleine, 2005; Balleine and O'Doherty, 2010) and nondeclarative and declarative memory (Squire and Zola, 1996) (for an insightful review, see Dayan, 2009). Thus the computational approach provides a powerful and parsimonious tool to address these theories.

Fourthly, computational modeling makes model-based functional magnetic resonance imaging (fMRI) (O'Doherty et al., 2007. See also Huettel, 2012) and electroencephalography (EEG) (Larsen and O'Doherty, 2014) possible, which moves observations from the behavioral and cognitive level down to the underlying neural level. Briefly speaking, in model-based fMRI and EEG, a computational model is first fit to observed behaviors, and the best-fitting model is regressed against the fMRI data of blood-oxygen-level dependent (BOLD) or EEG signal changes over time. In this way,

statistical correlation of the cognitive process and BOLD or EEG signal change could be detected. This allows precise understanding of the neural computation of the cognitive process, which goes beyond mere 'activation' or 'response' of brain areas reported by traditional neuroimaging and electrophysiological studies (O'Doherty et al., 2007). Further, the combined model-based fMRI-EEG allows more advanced simultaneous analysis of spatio-temporal dynamics of the computation (Larsen and O'Doherty, 2014).

Thanks to these merits, there has been an emerging field of computational psychiatry (Huys et al., 2011; Maia and Frank, 2011; Montague et al., 2012; Lee, 2013; Stephan and Mathys, 2014; Wiecki et al., 2015). Computational psychiatry has been on its way to address the fundamental challenges of psychiatric research especially the key issue of a lack of intermediate phenotypes (Gottesman and Gould, 2003; Rasetti and Weinberger, 2011; Bogdan et al., 2013). Intermediate phenotype connects clinical symptoms and behavioral observations to underlying molecular dysfunctions (Gottesman and Gould, 2003; Rasetti and Weinberger, 2011; Bogdan et al., 2013. See also Montague et al., 2012; Wiecki et al., 2015), a lack of which poses a serious obstacle for the drug discovery in psychiatry (Papassotiropoulos and de Quervaine, 2015).

Despite dozens of reviews on computational psychiatry, to date no single focused review on computational research of depression is available. The purpose of the present article is to fill in this gap and review computational studies that have focused on depression. In the following, we first provide a brief introduction of the computational framework of model-free RL (for a more detailed introduction, please see Sutton and Barto, 1998; Niv, 2009; Glimcher, 2011; Daw, 2013; Daw and Tobler, 2013) and review studies that have employed this framework. We then introduce the recent differentiation between model-free and model-based RL (for a more detailed account, see Gläscher et al., 2010; Daw et al., 2011; Doll et al., 2012; Daw and O'Doherty, 2013; Otto et al., 2013; Radenbach et al., 2015) and review related studies. Finally, we suggest several directions for future lines of investigation.

3. Research with model-free RL

3.1. Theoretical framework

As a schematic introduction, in a probabilistic learning task, subjects are asked to choose between stimulus A and B. Unknown to them, choosing A has a high probability of being rewarded and low probability of receiving no reward (omission) or punishment, while choosing B results in the opposite. After subjects have learned the rule, receiving the less frequently conditioned feedback or reversing the contingency would generate PEs. Specifically, PE can be calculated for each subject by a Rescorla–Wegner (RW) model as:

$$\mathrm{PE} = R(t) - V(t) \quad (1)$$

in which R(t) is the received value in trial t, and $V(t)$ is the expected value of the choice in trial t. In model-free RL, subjects rely primarily on historical PEs to update the expected value of choosing the stimuli in a forthcoming trial $t+1$,

$$V(t+1) = V(t) + \in * \mathrm{PE} \quad (2)$$

whereas $\in$ is learning rate. Learning rate is the core variable in model-free RL, and it reflects the efficiency of utilizing PE information to update the value function.

A more advanced model is temporal difference (TD) learning which takes future reward into consideration, such that PE is calculated as:

$$\mathrm{PE} = R(t) + \gamma V(t+1) - V(t) \quad (3)$$

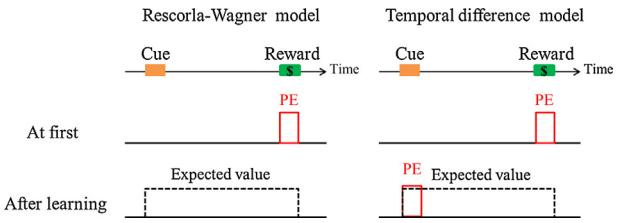


Fig. 1. Schematic comparison of time course of the prediction error and expected value signals in the Rescorla–Wagner (RW) model and the temporal difference (TD) model, respectively. At first an unexpected reward generates a prediction error (PE), in both the RW and TD model. As the task progresses, subjects will learn the cue-reward contingency gradually. After learning, the cue will result in an expected value signal, in both the RW and TD model. However, according to the RW model, there is no PE signal here (i.e. learning has already finished) since PE only occurs after receiving unexpected reward and here the reward has been predicted. In contrast, according to the TD model, the occurrence of the cue itself will generate a PE signal whose magnitude reflects the future predicted reward. For more detailed descriptions, please see Niv and Schoenbaum (2008), Niv (2009).

whereas γ is a discount factor or the weight given to future reward and $V(t+1)$ is the expected value of reward at trial (or state) $t+1$, which itself is only an estimate at trial t. Please see Fig. 1 for a schematic comparison of time course of the PE and expected value signals predicted by the RW model and the TD model, respectively.

Recent research suggests that PE signals are primarily calculated in the midbrain (ventral tegmental area, or VTA, and substantia nigra, or SN) dopamine neurons which then broadcast to the rest of the brain (see Fig. 2A for a schematic illustration of dopamine projections; Sutton and Barto, 1998; Niv, 2009; Schultz, 2010; Glimcher, 2011. For detailed descriptions of these projections, please see Haber and Knutson, 2010; Sesack and Grace, 2010; Roeper, 2013; Russo and Nestler, 2013; Haber, 2014; Haber and Behrens, 2014; Yetnikoff et al., 2014). The phasic dopamine activity signals PE so that it is enhanced by positive PE but suppressed by negative PE (more accurately speaking, here PE refers to reward PE rather than aversive PE, see discussion below; Schultz, 2006; Niv, 2009). However, the phasic dopamine is not exclusive to reward and punishment might be informed by a different signal other than dopamine (see Berridge, 2007; Schultz, 2010). Further, dopamine responses to reward or delayed reward have been found to be more consistent with the predictions of the TD model due to its superiority in accounting for the temporal relationships of multiple stimuli and rewards within a learning trial, and temporal discounting (reward that is far away in time is viewed as less valuable; for a review please see Niv and Schoenbaum, 2008, Niv, 2009).

In the following we review recently published articles falling under this section. This includes six studies on patients with MDD (Chase et al., 2010; Dombrovski et al., 2010, 2013, 2014; Gradin et al., 2011; Kumar et al., 2008), two on healthy subjects under stress (Friedel et al., 2014; Robinson et al., 2013) and one meta-analytic computational synthesis of six previous studies on patients with MDD or healthy subjects with depressive symptoms or under stress (Huys et al., 2013). The reason why we included studies on stress is that it is one of the most important risk factors for MDD so these studies may reveal the etiology and pathophysiology of MDD (de Kloet et al., 2005; Hammen, 2005; Maier and Watkins, 2005; Beck and Alford, 2009; Lupien et al., 2009; Bogdan et al., 2013).

To review these studies, a delineation between omission and punishment as negative feedback in the task is necessary since the former leads to negative 'reward PE' (RPE) while the latter leads to positive 'aversive PE' (APE). It is observed that omission of a reward suppresses dopamine neurons but delivery of a punishment activates dopamine neurons (Schultz, 2006, 2010; Niv, 2009). The phasic dopamine response to reward and omission of reward fits well with prediction by the reinforcement learning theory (Schultz, 2006, 2010; Niv, 2009; Glimcher, 2011). Thus recently omission of

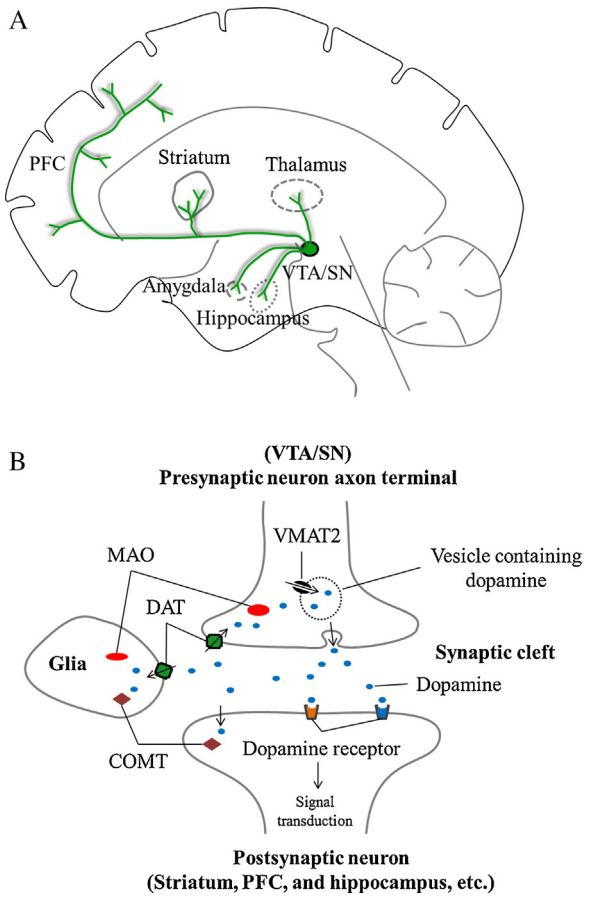
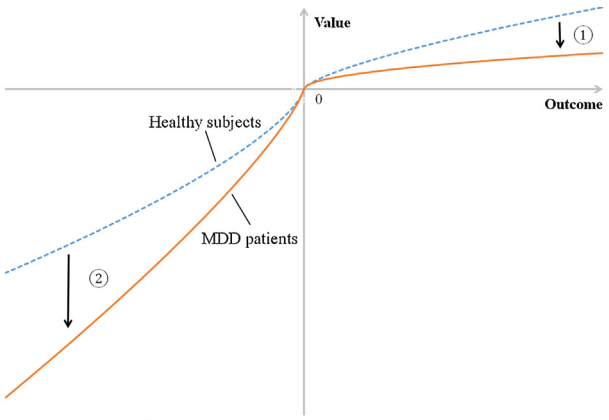


Fig. 2. A schematic illustration of dopamine projections (A) and dopamine neurotransmission (B). (A) Dopamine nuclei within the midbrain (VTA/SN) projects to the rest of the brain, including the striatum, PFC, hippocampus, amygdala, and thalamus, etc. Note that this is a simplified illustration in a way that does not differentiate the VTA from the SN projections. For detailed description of these projections, please see references in the text. (B) Dopamine is synthesized in the presynaptic neuron cytoplasm (for dopamine metabolism, see Meiser et al., 2013) and stored into the synaptic vesicles by the vesicular monoamine transporter 2 (VMAT2, Schuldiner et al., 1995; Liu and Edwards, 1997; Chaudhry et al., 2008). Upon neuronal excitation (action potential) dopamine is released into the synaptic cleft (for factors that influence dopamine release, Rice et al., 2011; Aggarwal et al., 2012). Once in the synapse, dopamine binds to and activates the postsynaptic dopamine receptors, which then trigger intracellular signal transduction. Dopamine receptors are subdivided into D1-like (D1 and D5) and D2-like (D2, D3, and D4) receptors according to their structural, pharmacologic and signaling properties (Beaulieu and Gainetdinov, 2011; Gerfen and Surmeier, 2011; Tritsch and Sabatini, 2012; Perreault et al., 2014; Savica and Benarroch, 2014). After an action potential, dopamine quickly becomes unbound from its receptors. It is then (1) reuptake back into the presynaptic neuron via dopamine transporter (DAT, Bannon, 2005; Eriksen et al., 2010) and subsequently repackaged into vesicles by VMAT2 or degraded by the enzyme monoamine oxidase (MAO, which is localized to the mitochondrial outer membrane and has two isoforms, MAO-A and MAO-B. For an overview of MAO and its role in depression, see Duncan et al., 2012; Finberg, 2014), or (2) uptake by glia cells via DAT and then degraded by either MAO or catechol-O-methyltransferase (COMT, Männistö and Kaakkola, 1999; Karhunen et al., 1995; Schott et al., 2010), or (3) uptake into the postsynaptic neuron by a DAT-independent but yet unknown mechanism and then degraded by COMT (Karhunen et al., 1995; Schott et al., 2010). Notably, the expression of DAT in the PFC is low and dopamine is reuptake by both DAT and norepinephrine transporter (Sesack et al., 1998; Wayment et al., 2001). In addition, D2-like receptors (autoreceptors) are also expressed on the midbrain dopamine neurons and on their axon terminals in projection areas. As feedback regulators, activation of these receptors decreases the excitability of dopamine neurons and the release of dopamine (Ford, 2014).

reward has been frequently treated in the reward category and used to generate negative RPEs. This treatment is also consistent with the findings in probability discounting of uncertain reward and punishment (Rachlin et al., 1991; Takahashi et al., 2013) in which uncertain non-gain (omission of uncertain reward) is processed as probability



MDD is characterized by ① hyposensitivity to reward and ② hypersensitivty to punishment.

Fig. 3. Schematic value functions. According to prospect theory (Kahneman and Tversky, 1979; Kahneman, 2011), humans are more sensitive to an increase in loss than gain ('loss aversion'), and are risk aversive in gain but risk-taking in loss. Algorithmically, a value function $v(x)$ can be described as,

$$v(x) = x^a \quad (\text{for } x > 0, \text{ gain})$$

$$v(x) = -\lambda(-x)^a \quad (\text{for } x < 0, \text{ loss})$$

where λ indicates loss aversion (>1) and a (<1) indicates risk-aversion in gain and risk-proneness in loss. In other words, value function is concave in gain but convex in loss. As MDD is associated with reduced reward sensitivity (hyposensitivity to reward) and enhanced sensitivity to punishment (hypersensitivity to punishment, see the test for references), we speculate that MDD patients may have larger λ values and may be more risk aversive in gain while more risk-taking in loss than healthy subjects.

discounting of gain, while uncertain loss is processed as probability discounting of loss, by human and animal subjects. Research has shown that reward (gain) and aversiveness (punishment/loss) have different value functions (Kahneman and Tversky, 1979; Takahashi et al., 2008a,b; Kahneman, 2011; Han and Takahashi, 2012) and that their neural basis are different (Boureau and Dayan, 2011; Garrison et al., 2013). Human valuation has been known to differ between gain and loss. In prospect theory (Kahneman and Tversky, 1979; Kahneman, 2011), such gain–loss asymmetry is characterized with the value function (Fig. 3). This describes that humans as more sensitive to an increase in loss than gain, and engage in more risk-taking behaviors in loss than gain (i.e., value function is concave in gain but convex in loss). Specifically, as depression is associated with reduced reward sensitivity (hyposensitivity to reward, Steele et al., 2007; Kumar et al., 2008; Gradin et al., 2011. See a review by Forbes and Dahl, 2012; Zhang et al., 2013; Kerestes et al., 2014; Pizzagalli, 2014) but enhanced punishment sensitivity (hypersensitivity or oversensitivity to punishment or error, Steele et al., 2007; Dombrovski et al., 2013, 2014. See reviews by Olvet and Hajcak, 2008; Clark et al., 2009; Vaidyanathan et al., 2012), we speculate that MDD patients may have an exaggerated tendency of this gain–loss asymmetry in their valuation. That is, they may have different features of RPE and APE than healthy controls.

3.2. RPE studies

Kumar and colleagues (Kumar et al., 2008; Gradin et al., 2011) conducted studies that used drops of water as positive feedback (rewards) and absence of drops of water as negative feedback (omission of rewards). Specifically, subjects were asked to not drink fluids the night before experiment to ensure they were thirsty. Kumar and colleagues examined RPE signals with the TD model-based fMRI.

Kumar et al. (2008) found that compared to healthy controls (age 42.0 ± 12.8 years), MDD patients (age 45.3 ± 12.3 years) who are

unresponsive to antidepressant treatment (Beck depression inventory (BDI) or Hamilton depression score > 21 despite more than three months of continuous antidepressant treatment) demonstrated reduced RPE signals in the ventral striatum and dorsal anterior cingulate cortex (ACC) and increased RPE signals in the VTA, rostral ACC, retrosplenial cortex and hippocampus. However, the increased signals in the rostral ACC, retrosplenial cortex, and hippocampus were due to lack of deactivation in the patients. That is, the RPE signals in these areas were actually attenuated and only the RPE signal in VTA increased. Surprisingly, in healthy controls (age 41.7 ± 12 years) three days of SSRI citalopram medication also induced similar attenuated RPE signals. The explanation of the latter finding remains unsure but it does suggest the involvement of 5-HT in RL (see a discussion by Eshel and Roiser, 2010). Nevertheless, when compared to medicated controls, MDD patients had blunted RPE signals in the rostral ACC and increased RPE signals in the VTA. Finally, the severity of MDD was associated with RPE signals, such that more severe MDD was associated with weaker hippocampal and rostral ACC signals but stronger VTA and amygdala RPE signals.

Note that in this study, the authors did not estimate the learning rate for each subject but instead used a fixed learning rate based on previous research for all subjects ($\in = 0.1$, the results were also stable when $\in = 0.4$). The former is known as 'individual learning' whereas the latter is referred to as 'fixed learning' (Cohen, 2007; Chase et al., 2015). Whereas individual learning is better at accommodating subjects' behaviors, fixed learning reduces noise and may improve reliability but at the expense of losing individual learning rate data (Cohen, 2007; Chase et al., 2015).

In their later study, Gradin et al. (2011) employed the individual learning strategy. They found no difference in the learning rate between MDD patients (age 45.27 ± 12.32 years) and healthy controls (age 40.64 ± 11.87 years). With the TD model-based fMRI, they found reduced RPE signals in the striatum, thalamus, midbrain and right hippocampus of patients. Further, the signal reduction in the striatum and midbrain correlated with increased anhedonia (derived from BDI). The authors also estimated the expected value signal. They found that MDD patients showed reduced expected value signal in the right hippocampus and parahippocampus gyrus, although the reduction was not correlated to severity of depression or anhedonia. We will discuss this expected value signal later in Section 3.4.

Taken together, these two RPE studies consistently reported reduced brain signals of RPE in the hippocampus, ACC, striatum, and thalamus in depressed patients. Some of these attenuated signals were associated with the severity of MDD or anhedonia. However, whereas Kumar et al. (2008) found increased RPE signals in the VTA in patients, Gradin et al. (2011) found decreased RPE signals in the VTA. Further, in both studies these RPE signals were correlated with increased severity of anhedonia symptom. How does one explain these seemingly contradictory findings?

On the one hand, the findings by Gradin et al. (2011) suggest that the RPE signals in the mesolimbic system are overall reduced. As mentioned above, the RPE signals are primarily calculated in the midbrain (VTA/SN) dopamine neurons which then broadcast to the rest of the brain (Fig. 2A). Therefore, the findings by Gradin et al. (2011) may have highlighted the role of the hypoactive midbrain (the source of RPE signals) in MDD. In support of this, research shows that whereas lesions of dopamine neurons in the VTA or SN (Winter et al., 2007; Favier et al., 2014; Santiago et al., 2014) or optogenetic inhibition of VTA dopamine neurons (Tye et al., 2013) enhance depressive-like behavior in rodents, deep brain stimulation of the VTA (Friedman et al., 2009, 2012; Gazit et al., 2015) or optogenetic activation of VTA dopamine neurons (Tye et al., 2013) alleviate depressive behavior in animal models of depression. Several antidepressants, such as ketamine, escitalopram, and desipramine, have also been shown to correct reduced VTA dopamine neuron firing while normalizing depressive-like behavior in animal models of depression (Friedman et al., 2008; Schilström et al., 2011; Belujon and Grace, 2014). Together these findings argue for the fundamentally critical role of the VTA/SN and the dopamine system in the etiology, pathophysiology and treatment of depression and anhedonia (Kapur and Mann, 1992; Naranjo et al., 2001; Nestler and Carlezon, 2006; Gershon et al., 2007; Dunlop and Nemeroff, 2007, 2010; Russo and Nestler, 2013; Pizzagalli, 2014; Polter and Kauer, 2014).

In line with the notion that overall reduced dopamine plays a role in depression, maladaptive coping with stress (such as learned helplessness) has been shown to be associated with reduced mesolimbic dopamine release (Cabib and Puglisi-Allegra, 2012). Similarly, people who are genetically predisposed to lower dopaminergic neurotransmission are more likely to have higher levels of depression (Pearson-Fuhrhop et al., 2014). Further, cellular and molecular studies suggest that depression is associated with (1) reduced homovanillic acid (a major dopamine metabolite) in central nervous system in depression (see a brief review by Dunlop and Nemeroff, 2007; Pizzagalli, 2014. See Fig. 2B for a brief introduction of the dopamine neurotransmission); (2) elevated monoamine oxidase-A (MAO-A, an enzyme that metabolizes monoamines) throughout the brain in patients with MDD, notably the dopamine system (i.e. midbrain, striatum, prefrontal cortex, and hippocampus; Meyer et al., 2006; Chiuccariello et al., 2014; Sacher et al., 2015); (3) reduced vesicular monoamine transporter 2 (VMAT2, which packages dopamine into vesicles presynaptically to be ready for release upon receiving an action potential) in the VTA, SN and striatum (Schwartz et al., 2003); (4) hyperactivity of catechol-O-methyltransferase (COMT), one of the main enzymes responsible for catecholamine degradation (Opmeer et al., 2010; Kocabas, 2012; Antypa et al., 2013).

On the other hand, the results by Kumar et al. (2008) that RPE signals in VTA were increased in MDD and were correlated with the severity of MDD, makes the picture more complex. How can the source of RPE signals be enhanced while the receiving ends are attenuated? One possibility is that, perhaps the deficit is at the level of the receiving ends (e.g. ventral striatum, dorsal ACC, and hippocampus) whereas the dopamine activity in the source (VTA) up-regulates in a compensatory way (Dunlop and Nemeroff, 2007). This might especially be true in severe depression (Dunlop and Nemeroff, 2007; Kumar et al., 2008). Indeed, increasing evidence supports this receiving-ends-deficits account of depression. Deep brain stimulation of the striatum has demonstrated its effectiveness and efficacy in treating depression both clinically (Schlaepfer et al., 2008; Aouizerate et al., 2009; Malone et al., 2009; Bewernick et al., 2010, 2012) and preclinically (Gersner et al., 2010; Schmuckermair et al., 2013; Hamani et al., 2014). At cellular level, the deficits in striatum might be related to reduced VMAT2, elevated MAO-A (reviewed above), increased DAT (Laasonen-Balk et al., 1999; Brunswick et al., 2003; Yang et al., 2008; Amsterdam et al., 2012; Hsiao et al., 2013), decreased D2 receptor binding (Willner et al., 1991; Papp et al., 1994), and/or deficient D3 receptor (Moraga-Amaro et al., 2014). Bupropion, which selectively blocks DAT, is effective at treating depression in patients, likely by decreasing striatal DAT binding (Hsiao et al., 2013). Further, antidepressant treatments increase the density of D2/D3 binding sites in the nucleus accumbens (ventral striatum; Ainsworth et al., 1998. But see Hirvonen et al., 2011), whereas the antidepressant-like effect of ketamine and MK-801 is prevented by the dopamine D2/D3 receptor antagonist (Li et al., 2015). The dopamine D2/D3 receptor agonist pramipexole alone or as an additive to traditional antidepressants, is effective in treating depression or drug-resistant depression in patients (Cassano et al., 2004; Barone et al., 2006; Inoue et al., 2010; Cusin et al., 2013. But see Leentjens, 2011;

Franco-Chaves et al., 2013), and in reversing the motivational deficit induced by SN dopamine depletion (Favier et al., 2014) and preventing depression-like behavior (Maj and Rogóz, 1999; Li et al., 2015) in rats. Actually, the increased expression or binding of the dopamine D2-like receptor (Gershon et al., 2007) especially D3 receptor (Lammers et al., 2000) has been proposed to be a common pathway of chronic antidepressant treatments. Similar findings have been reported for the prefrontal cortex and hippocampus. For instance, greater severity of depression was associated with elevated MAO-A density in the PFC and ACC (Chiuccariello et al., 2014; Sacher et al., 2015). As well, chronic isolation stress caused reduction of the VMAT2 protein in the rat hippocampus (Spasojevic et al., 2012). Further, social defeat stress increased DAT binding in the medial PFC (Novick et al., 2011) and up-regulated dopamine β-hydroxylase protein (which converts dopamine to norepinephrine) levels in the frontal cortex and hippocampus (Fan et al., 2013). Of note, the latter was prevented by treatment with the antidepressant desipramine, an inhibitor of the norepinephrine transporter which prevents dopamine re-uptake into noradrenergic neurons (Fan et al., 2013). Further, chronic restraint stress reduces D1 receptor binding in the medial PFC which is normalized after chronic recovery (Goldwater et al., 2009), whereas pramipexole increases the tonic activation of postsynaptic D2 receptor in PFC in rats (Chernoloz et al., 2012).

The findings by Kumar et al. (2008) and our proposed receiving-end-deficits account of depression can also be understood from the perspective of neural functional connectivity. We can speculate that the functional connectivity between the VTA/SN and the deficit receiving-end(s) might be reduced in MDD. Indeed, a recent study reported that, with higher level of baseline anhedonia, non-responders to repetitive transcranial magnetic stimulation (rTMS) of the dorsomedial PFC demonstrated reduced baseline functional connectivity between the VTA and medial PFC when compared to responders (Downar et al., 2014).

Of course, the compensatory mechanisms can also occur at the receiving-ends outside of the VTA. Similar compensatory mechanisms to reduced dopamine release include increased D1 receptor density in the PFC in chronically stressed rats (Mizoguchi et al., 2000) and decreased DAT levels and increased D2 receptor binding in the striatum in MDD patients (see a review by Dunlop and Nemeroff, 2007, 2010; Pizzagalli, 2014).

Finally, in addition to the compensatory mechanism theory, clinical heterogeneity might also contribute to the inconsistency of VTA activation. For instance, social defeat and chronic restraint increase VTA dopamine neuron firing while chronic cold stress and inescapable, uncontrollable footshock decrease it (Valenti et al., 2012; Belujon and Grace, 2014; Friedman et al., 2014. For a review and discussion of this opposite effect, please see Russo and Nestler, 2013; Polter and Kauer, 2014).

Whatever the underlying mechanism may be, what is obvious from these studies is the essential role of the dopamine system in depression and the usefulness of the RL approach in differentiating the subtypes and revealing the underlying pathophysiology of MDD. Specifically, the abnormal RPE signals might be especially helpful in identifying the origin of the dysfunctional brain network in MDD and in revealing treatment target(s).

3.3. RPE and APE studies

Chase et al. (2010) and Dombrovski et al. (2010, 2013, 2014) used 'Correct' and 'Incorrect/Wrong' as the feedback, which have shown similar behavioral and neural effects as symbolic or monetary reward and punishment.

Using a probabilistic reversal learning task, Dombrovski et al. (2010) found that suicide attempters (age 66.8 ± 7.8 years) were impaired in the reversal but not acquisition stage when compared to non-depressed controls (age 65.6 ± 4.6 years), non-suicidal depressed patients (age 70.0 ± 7.4 years), or patients with suicidal ideation (age 68.8 ± 5.6 years). In contrast, Dombrovski et al. (2013) found that MDD patients all performed poorly in the reversal stage compared to controls (age 70.7 ± 8.7 years), with no difference between suicide attempters (age 65.9 ± 6.3 years) and non-suicidal patients (age 66.7 ± 5.7 years). However, non-suicidal depressed patients were more likely to switch after misleading negative feedback than controls (as they were likely oversensitive to the punishment). Finally, excluding six participants missing measures of executive control from the sample of Dombrovski et al. (2013), the researchers re-analyzed their data and reported that depression predicted lose-switches in the task (i.e. oversensitivity of punishment), while poor executive control was associated with a failure to persist with reward actions (although executive control but not depression predicted the overall correct responses in the task; Dombrovski et al., 2014).

Since subjects can achieve learning in reversal learning task by either trial-and-error experience (model-free RL) or by inferring the reward rule (model-based RL) (Daw et al., 2005; Costa et al., 2015), the impairment in reversal learning observed by Dombrovski and colleagues suggests dysfunctional model-free RL, model-based RL, or both. Further, the association between depression and oversensitivity to punishment found by Dombrovski et al. (2013, 2014) is consistent with a large literature on the impaired control over negative feedback in depressed individuals, behaviorally and/or neurally (Steele et al., 2007; Mueller et al., 2015. See reviews by Olvet and Hajcak, 2008; Clark et al., 2009; Vaidyanathan et al., 2012). The neural response frequently visited is known as feedback-related negativity (FRN) or error-related negativity (ERN), a component of the event-related brain potential that closely follows error commission in reaction-time tasks (Holroyd and Coles, 2002; Nieuwenhuis et al., 2004). It is proposed that the FRN or ERN resembles the RPE and/or APE signals in RL framework (Holroyd and Coles, 2002; Nieuwenhuis et al., 2004; Walsh and Anderson, 2012).

With regard to computational analysis, Chase et al. (2010) used the RW model. They observed that the rewarding learning rate and punishment learning rate were both reduced in MDD patients (age 46.22 ± 2.25 years) compared to healthy controls (age 47.74 ± 2.14 years) in the initial training phase but not in the subsequent test phase. Similarly, they also found a negative correlation between learning rate and anhedonia (measured by the Snaith–Hamilton anhedonia score, high score reflects low anhedonia) for all subjects in the training phase, and the results were not significantly different between reward learning rate and punishment learning rate. Unlike the RW model used by Chase et al. (2010) however, Dombrovski and colleagues (Dombrovski et al., 2010, 2013, 2014) also incorporated a 'memory' parameter to predict the expected value (as compared to formula (2)):

$$V(t+1) = m * V(t) + \in * \mathrm{PE} \quad (4)$$

whereas m is memory, or the degree to which subject is influenced by the prior reinforcement history rather than the last trial on that choice. This reflects one of the on-going efforts to bridge the approaches of memory and RL (Landauer, 1969; Shimp, 1976; Williams, 1991; Ragland et al., 2009; Otto et al., 2013).

With the new formula, Dombrovski et al. (2010) found that the three depressed groups tended to have a lower punishment learning rate ($p = 0.066$) than non-depressed controls, although the reward learning rate did not differ across groups. Further, suicide attempters had lower memory than non-depressed controls. In contrast, Dombrovski et al. (2013) detected a decreased reward learning rate among suicide attempters ($p = 0.048$), while the punishment learning rate and memory did not differ across groups.

Recall that in Section 3.2, Gradin et al. (2011) did not find reduced learning rate in middle-aged MDD patients. Here Chase

et al. (2010) found equally reduced reward and punishment learning rate in middle-aged MDD patients in the initial training phase. While Dombrovski et al. (2010) found a trend toward reduced punishment learning rate in older depressed patients, Dombrovski et al. (2013) found reduced reward learning rate in older patients with suicide behaviors. The reinforcer type may have contributed to this inconsistency. Gradin et al. (2011) used drops of water and absence of drops of water for thirsty subjects as the reinforcer, while Chase et al. (2010) and Dombrovski et al. (2010, 2013) used 'Correct' and 'Incorrect/Wrong'. The former is more actual and immediate and belongs to the RPE (reward and omission of reward) paradigm. The latter is more cognitive and abstract and belongs to the RPE and APE (reward and punishment) paradigm. It is possible that depressed patients may be much worse at learning from cognitive and abstract reinforcers given their cognitive deficits, and at learning from reward in the face of punishment. This hypothesis remains subject to further investigation. Besides, depression severity and number of probability reversals in the trial may also have affected the result. Patients in Dombrovski et al. (2013) seem to be less severely depressed (Hamilton depression score: non-suicidal depressed 11.1 ± 6.2; suicide attempter 12.9 ± 8.1) than those in Dombrovski et al. (2010; Hamilton depression score: non-suicidal depressed 18.5 ± 3.7; suicide ideator 20.2 ± 4.7; suicide attempter 19.2 ± 4.7), Kumar et al. (2008; Hamilton depression score: 23.2 ± 5.3; BDI: 22.9 ± 8.2), Gradin et al. (2011; Hamilton depression score: 23.2 ± 4.3; BDI: 22.9 ± 8.2), and Chase et al. (2010; BDI: 26.78 ± 1.79). Whereas Dombrovski et al. (2010) included one acquisition stage and one reversal stage, others generally had several reversals. For instance, Dombrovski et al. (2013) included 300 trials in which the contingency changed every 25 trials. These findings suggest that impaired learning rate may not be reliably detected in MDD patients, depending on the specific research paradigm (e.g. reinforce type, number of probability reversals), patient characteristic (e.g. age, severity of depression, and history of suicide behaviors), and situation (e.g. training vs. test phase), etc.

Further, the reduced memory of previous reinforcement history in suicide attempters found by Dombrovski et al. (2010) merits attention. It suggests that suicide attempters are excessively influenced by the last single episode compared to all previous reinforcement history. That is, they may frequently focus on the current negative outcomes irrespective of previous and future rewards. This is consistent with impaired reward-related memory (Gotlib and Joormann, 2010; Russo and Nestler, 2013; Belzung et al., 2014b; Dillon et al., 2014) and dysfunctional cognitive style of rumination (Belzung et al., 2014b; Watkins and Nolen-Hoeksema, 2014) frequently observed in MDD. It again suggests that among MDD patients those with suicide behaviors in particular may have dysfunctional RL. This is in accordance with a recent meta-analysis which found that suicide attempters perform much worse than mood disorder patient controls and healthy controls in several decision making and cognitive tests, including the IGT, category verbal fluency, and the Stroop interference test (Richard-Devantoy et al., 2014). These dysfunctions seem to be rooted in the impaired neural control over reaction to salient negative stimuli, which has been linked to structural and functional deficits of the brain, e.g. reduced volume of caudate nucleus and increased activation of ACC (van Heeringen et al., 2014).

Finally, Dombrovski et al. (2013) also examined the RPE signals with a model-based fMRI. They found reduced positive RPE signals in MDD patients in the corticostriatothalamic networks, including the right thalamus, bilateral superior temporal gyrus, bilateral operculoinsular cortex, bilateral postcentral gyrus, and bilateral supplementary motor area. This observation held true when the authors excluded the parameter memory in their computational modeling (i.e. formula (2)). These findings were robust to co-occurring substance use and anxiety disorders, and antidepressant treatment, etc. The reduced positive RPE signals observed here is in agreement with those found by the RPE studies in Section 3. Further, Dombrovski et al. (2014) showed that poor executive control and depression additively explained the reduced brain RPE signals in the corticostriatothalamic networks, controlling for age, gender and education. This is consistent with the essential role of executive control and cognition in model-based RL (Chen, 2014), which will be discussed in detail below in Section 4.2.

Besides, Dombrovski et al. (2013) also analyzed the expected value signal. The neural responses to expected value were qualitatively similar between non-suicidal depressed patients and healthy controls. However, suicide attempters failed to display the paralimbic responses to high expected value. Specifically, a history of suicide attempts was associated with a reduced expected value signal in the ventromedial PFC. Again we will return to these results on expected value signal in Section 3.4.

Unfortunately however, Dombrovski et al. (2013) failed to predict BOLD responses to punishment with the model-based fMRI, confirming that the neural substrates of RPE and APE may be indeed different. It suggests the limitation of the current computational models in addressing APE in MDD and future research should identify more specific algorithms for modeling APE.

In addition of these RPE and APE studies on MDD, Robinson et al. (2013) and Friedel et al. (2014) examined RPE and APE in healthy subjects under stress. Both studies employed a probabilistic learning task with a happy or smile face as positive feedback (reward) and a fearful or frowning face as negative feedback (punishment) and a RW model for computational modeling. Robinson et al. (2013) found that an acute stress of foot shock threat enhanced the ventral striatum signals of APE but not RPE. This is in agreement with the frequent involvement of the ventral striatum in signaling APE (Delgado et al., 2008; Schiller et al., 2008; see review by Garrison et al., 2013. However, ventral striatum itself may not be the origin of APE, see Roy et al., 2014) and the facilitation of aversion learning by stress (Maier and Watkins, 2005). It is also in accordance with growing evidence from the ERN approach that acute stress especially unpredictable stress increases the ERN (Hajcak et al., 2005; Riesel et al., 2012; Jackson et al., 2015). In contrast, Friedel et al. (2014) examined chronic stress as defined by the number of stressful life events happened in the past two years, and investigated its effect on brain signals of overall PE without differentiating the RPE from APE. They showed that chronic stress was positively correlated with BOLD PE signal in the left ventral striatum, controlling for age. Moreover, they found a positive association between fluid IQ and BOLD PE signal in the bilateral ventral striatum, and that this association was enhanced in subjects reporting more chronic stress. As Friedel et al. (2014) did not separate RPE from APE, it is difficult to interpret their findings except that chronic stress influences brain signals of PE and that the fluid IQ also plays a role in it. It remains for future investigations to clarify potential hidden associations.

To summarize, the research we reviewed in Sections 3.2 and 3.3 with the model-free approach seems to support three preliminary conclusions. Firstly, depression is associated with impaired brain signals of RPE, which may help locate the deficit brain area(s). Secondly, whether learning rate is reduced in depression remains relatively inconsistent (see Table 1), although it may be impaired under some research paradigms, in particular patients or under specific situations. Thirdly, memory of previous reinforcement may be impaired in patients with suicide behaviors.

3.4. *Learning rate or reward sensitivity? The dissociation of learning, 'liking' and 'wanting'*

However, an important limitation of the above studies is that they may have failed to specify another critical factor, the reward

Table 1
Intermediate phenotypes of depression suggested by computational research.

	Intermediate phenotypes		Neural substrates[b]	Changes in depression
Neural	RPE signals		Phasic dopamine in the VTA/SN and its projection areas, e.g. striatum, PFC, hippocampus	(1) VTA/SN: ↑ (Kumar et al., 2008) ↓ (Gradin et al., 2011) (2) Striatum, ACC, hippocampus, thalamus, etc.: ↓ (Kumar et al., 2008; Gradin et al., 2011; Dombrovski et al., 2013) (3) Amygdala: ↑ (Kumar et al., 2008)
Behavioral	'Wanting'	Expected value[a]	Medial PFC, hippocampus	(1) Hippocampus: ↓ (Gradin et al., 2011) (2) Ventromedial PFC: ↓ (Suicide attempters, Dombrovski et al., 2013)
	'Liking'	Reward sensitivity	Opioid/endocannabinoid 'hedonic hotspots' in nucleus accumbens, ventral pallidum, etc.	↓ (Meta-analysis, Huys et al., 2013)
	Learning	Model-free: learning rate	Ventral striatum, putamen (dorsolateral striatum)	(1) Reward learning rate: ↓ (Chase et al., 2010; suicide attempters Dombrovski et al., 2013); — (Dombrovski et al., 2010; Gradin et al., 2011. Meta-analysis, Huys et al., 2013) (2) Punishment learning rate: ↓ (Chase et al., 2010; Dombrovski et al., 2010); — (Dombrovski et al., 2013) (3) ↑ Maddox et al. (2014)
		Model-based	Medial PFC, hippocampus, ventral striatum, caudate (dorsomedial striatum)	↓ (Blanco et al., 2013; Maddox et al., 2014)
	Others	Memory of previous reinforcement		↓ (Suicide attempters, Dombrovski et al., 2010) — (Dombrovski et al., 2013)

[a] This is cognitive 'wanting'. Other measures of 'wanting'/motivation include incentive salience (incentive 'wanting'), effort willing to engage to pursue the reward, etc.
[b] For references, please see the text.

sensitivity (Huys et al., 2013), neither theoretically nor algorithmically. With a closer look at formula (1) we described in Section 3.1, PE is determined by received value R and expected value V, the latter being affected by learning rate. Since in the modeling R is generally chosen as 1 for delivery of reward and 0 for omission, the value subjects assign to R (i.e. reward sensitivity or 'liking') is not explicitly specified.

One prominent theory proposes that 'liking' is dissociable from 'wanting' and learning (Berridge and Robinson, 1998; Berridge, 2007, 2012; Berridge et al., 2009). 'Liking' represents a neurophysiological response to hedonic stimuli and is consummatory. 'Wanting' refers to the incentive salience that motivates the individual approach rewards and is anticipatory. 'Wanting' usually corresponds to the expected value V (learned prediction about the associated reward outcome or cognitive 'wanting', see formula (2)), but under some situations such as physiological state changes (e.g. appetite, satiety or drug-state changes) the cue-triggered incentive 'wanting' can dramatically depart from cognitive 'wanting' (Berridge et al., 2009; Berridge and O'Doherty, 2014). On the other hand, learning is defined as the acquisition of contingencies between rewards and cues to construct expected value functions, which can be achieved by either model-free or model-based RL. Unlike the involvement of mesolimbic dopamine system in 'wanting' and learning, 'liking' seems to depend on opioid/endocannabinoid 'hedonic hotspots' in a few discrete brain structures, such as the nucleus accumbens and ventral pallidum (for recent reviews see Richard et al., 2013; Castro and Berridge, 2014. For a description of ongoing debates, see Berridge and O'Doherty, 2014).

A new formula incorporating the reward sensitivity for calculating PE would be,

$$\mathrm{PE} = \rho * R(t) - V(t) \quad (5)$$

whereas ρ is reward sensitivity. Therefore, the reduced RPE signals in MDD may result from either decreased reward sensitivity or decreased learning rate. Previous research may have not separated them. However, firstly, as shown by Huys et al. (2013), learning rate and reward sensitivity have partially replaceable roles and are anti-correlated with each other. Secondly, reward sensitivity can be substituted by another internal temperature parameter controlling exploration/exploitation, which has been often incorporated into the computational modeling. These two issues make the computational method used by Huys et al. (2013) somewhat complicated and potentially confusing (for detailed methodological discussions please see Huys et al., 2013). Nevertheless, separating reward sensitivity from learning rate in the computational modeling is theoretically pioneering and merits special attention.

Employing the new formula, Huys et al. (2013) re-analyzed the behavioral data in six of their previous RPE studies under signal-detection approach (subjects including MDD patients, bipolar disorders patients, and healthy volunteers, etc.). They observed a significant negative association between anhedonia and reward sensitivity. Surprisingly, despite a negative correlation between learning rate and reward sensitivity ($r = -0.41$), anhedonia was not associated with learning rate. A group-level comparison also showed that subjects with MDD or high anhedonia had reduced reward sensitivity but not learning rate. Therefore the authors concluded that anhedonia and MDD affect primarily reward sensitivity rather than learning rate. However, they found that stress reduced both reward sensitivity and learning rate in healthy subjects.

Consequently based on Huys et al. (2013) and studies that identified consistent dysfunctional nucleus accumbens (involved in both model-free and model-based RL) but inconsistent dysfunctional midbrain (involved more in model-free RL) in MDD (e.g. Kumar et al., 2008 vs. Gradin et al., 2011), Huys et al. (2015) argued that the learning problem in depression may rest more on deficits in the model-based rather than model-free RL. This argument is potentially in agreement with emotion regulation literature that has shown that voluntary rather than automatic regulation process is impaired in MDD (Rive et al., 2013). It may also explain the

inconsistent results on learning rate we reviewed in Sections 3.2 and 3.3 (see Table 1). Nevertheless, it seems still immature to draw any conclusive statements regarding this issue. Whether model-free RL is unaffected in MDD, or it is only impaired under some research paradigms, in particular patients or under specific situations remain to be clarified.

On the other hand, the reduced reward sensitivity ('liking') in depression is consistent with the reduced striatal reactivity to reward outcomes (Steele et al., 2007; Kumar et al., 2008; Gradin et al., 2011. See a review by Forbes and Dahl, 2012; Zhang et al., 2013; Kerestes et al., 2014; Pizzagalli, 2014), which might be partly accounted for by the deficit opioid/endocannabinoid system. Early evidence suggested reduced endogenous opioid tone in depressed people (Burnett et al., 1999). Later, Zubieta et al. (2003) reported that in healthy human volunteers, sadness was characterized by reduced mu-opioid neurotransmission in the hedonic hot spot ventral pallidum, the reduction of which was correlated with increases in negative affect and decreases in positive affect. More recently, Kennedy et al. (2006) reported dysfunctional mu-opioid transmission in both the nucleus accumbens and ventral pallidum in depressed patients. Further, the mu-opioid receptor agonists, delta-opioid receptor agonists and kappa-opioid receptor antagonists show promising anti-stress and antidepressant effects (Bershad et al., 2015. See reviews by Lutz and Kieffer, 2013; Van't Veer and Carlezon, 2013; Bali et al., 2015 for related findings). In the meantime, chronic unpredictable stress, a rat model of depression, decreases cannabinoid CB1 receptor binding density in the ventral striatum, which was attenuated by concurrent antidepressant treatment (Hill et al., 2008). Mice lacking CB1 receptor has been considered a validated animal model for depression (Valverde and Torrens, 2012; Burokas et al., 2014), while mice overexpressing CB2 receptor show decreased depressive-like behavior (García-Gutiérrez et al., 2010). Acute administration of a CB1 receptor agonist (Kruk-Slomka et al., 2015), CB2 receptor agonist (Bahi et al., 2014; Kruk-Slomka et al., 2015) or antagonist (Kruk-Slomka et al., 2015) also causes antidepressant-like effect in mice. In humans, variation in the endocannabinoid genes is found to (1) be related to depression (Onaivi et al., 2008), (2) moderate the effects of childhood physical abuse (Agrawal et al., 2012) and recent life events (Juhasz et al., 2009) on anhedonia and depression, and (3) affect the pallidum reactivity to happy faces and responsiveness to antidepressant treatment (Domschke et al., 2008. See reviews by Vinod and Hungund, 2006; Hill et al., 2009; Gorzalka and Hill, 2011 for research on endocannabinoid and depression). Given the central role of opioid/endocannabinoid in hedonic 'liking' (Richard et al., 2013; Castro and Berridge, 2014), together these findings suggest underlying molecular pathophysiology of the attenuated 'liking' in depression.

Finally, with regard to 'wanting', earlier in Sections 3.2 and 3.3 we have described findings on brain signals of expected value by Gradin et al. (2011) and Dombrovski et al. (2013). Recall that Gradin et al. (2011) found reduced expected value signal in the right hippocampus and parahippocampus gyrus in MDD patients. Dombrovski et al. (2013) reported qualitatively similar expected value signals between non-suicidal depressed patients and healthy controls, but disappeared paralimbic signals of high expected value in suicide attempters. Specifically, a history of suicide attempts was associated with a reduced expected value signal in the ventromedial PFC. These results suggest deficient 'wanting' in patients with depression and in particular suicide behaviors. That is, they show reduced motivation to pursue rewards, which may contribute to the onset and maintenance of the disease. At least three lines of evidence support this hypothesis. Firstly, as we have mentioned at the beginning of this review, fMRI research has consistently found a low striatal and/or high medial PFC response during monetary anticipation in MDD (Steele et al., 2007. See a review by Forbes and Dahl, 2012; Zhang et al., 2013; Kerestes et al., 2014; Pizzagalli, 2014). A recent meta-analysis confirmed the role of ventromedial PFC in coding expected value (Chase et al., 2015; for an elegant review, see also Rangel and Clithero, 2013). Secondly, Treadway et al. (2009) identified an inverse association between anhedonia and willingness to expend effort for rewards, an objective measure of reward motivation. Treadway et al. (2012a) further found that MDD patients were less willing than controls to expend effort for rewards, and the impairment was more severe in patients with longer duration of the current episode. Not surprisingly, the willingness to expend more effort for larger rewards was correlated with individual differences in dopamine functioning in the striatum and ventromedial PFC, particularly when the probability of reward receipt was low (Treadway et al., 2012b). Thirdly, the reduced 'wanting' is also in line with the behavioral approach deficits in depression (Trew, 2011). Behavioral approach or behavioral activation is defined as desires to approach rewards or positive outcomes. It is in contrast to the construct of behavioral avoidance or behavioral inhibition, which refers to avoiding punishment or negative outcomes (Trew, 2011). Research shows that approach motivation predicts depression development and recovery (Trew, 2011), and that behavioral activation treatment which aims to enhance the approach motives and goals and initiation of actions is effective at relieving depression (Mazzucchelli et al., 2009; Kanter et al., 2010; Dimidjian et al., 2011; Trew, 2011). Thus, 'wanting' may be impaired in depression. Recently Baskin-Sommers and Foti (2015) argued that 'wanting' is intact in MDD, mainly based on findings by Knutson et al. (2008) and Pizzagalli et al. (2009). Both studies used a monetary incentive delay task, in which subjects were asked to earn rewards by quickly responding to a cue indicating a potential outcome and learning contingencies. Two variable delays (e.g. 3–8 s) exist, one between the cue and the stimuli to which subjects respond, the other between subjects' responding and the feedback outcome. This was designed to 'leverage the spatial and temporal resolution of functional magnetic resonance imaging (i.e., millimeters and seconds) to 'localize' affective responses deep in the brain' (pp. 418, Knutson and Heinz, 2015). In these two studies, MDD patients displayed robust nucleus accumbens activation during anticipation of rewards and showed normal reaction time similar to those of health controls (Knutson et al., 2008; Pizzagalli et al., 2009). Therefore Baskin-Sommers and Foti (2015) argued that 'wanting' is intact in MDD, at least when patients are explicitly instructed to earn rewards. However, we agree with Treadway and Zald (2011) that the anticipation phase of this task may not be a pure, reliable measure of motivation, as it may capture the neural responses of motor preparation. Thus in agreement with Treadway and Zald (2011, 2013), we argue that 'wanting' might be impaired in MDD.

Do impairments arise in all of these processes, learning, 'liking' and 'wanting' in MDD? Not necessarily. For instance, in a recent study by Sherdell et al. (2012), MDD patients and healthy controls reported similar consummatory liking of humorous cartoons. However, levels of liking predicted motivation to expend effort for the rewards in healthy controls but not in MDD patients, suggesting dissociated 'liking' and 'wanting' in MDD. Thus, impairments may arise from problems in some or all of these processes and clarification of this can facilitate more precise differentiation of subtypes of MDD and improve diagnoses and treatments.

4. Research with model-based RL

4.1. Theoretical framework

Following the recent advance in theoretical differentiation between model-free and model-based RL, Daw and colleagues have

designed a two-step task and computational algorithms to directly simulate this differentiation (Gläscher et al., 2010; Daw et al., 2011; Otto et al., 2013). In this task, subjects are asked to choose between two stimuli (action a_A vs. a_B) at step-one (s_1) that, with different probabilities (e.g. 0.7 with a_A vs. 0.3 with a_B) lead to one of the two different subsequent step-two states (s_{2a} vs. s_{2b}). At step-two they have to make another choice. The reward probability of this second choice changes over time and subjects have to adjust their choices trial by trial. Consider a step-one choice (a_B) associated with a lower probability (0.3) of transition to a step-two state, and the subsequent step-two choice is rewarded. Under a pure model-free RL that depends primarily on reinforcement history, subjects would repeat the same step-one choice (a_B) since it is ultimately rewarded. In contrast, model-based RL predicts a decreased tendency to repeat the same step-one choice (a_B) because it is the other choice (a_A) that has a higher probability (0.7) of leading to the rewarded step-two state. In general, subjects with model-based RL first learn the transition function, i.e. which step-one action maps to which step-two state with what probability and then learn the reward value for each state, after which they calculate the cumulative state-choice values by iterative expectation. Further, the relative influence of model-free vs. model-based RL can also be estimated. For details about the algorithms, please refer to Daw et al. (2011), Otto et al. (2013), Radenbach et al. (2015).

Utilizing this framework and other similar logics, recent neuroimaging research has demonstrated that, whereas model-free RL involves the ventral striatum (Daw et al., 2011) and the dorsolateral striatum (Liljeholm and O'Doherty, 2012; for a schematic anatomy of the striatum and current theories of the function of its subregions, see Chen et al., 2014), model-based RL involves the ventral striatum (Daw et al., 2011), the dorsomedial striatum (Liljeholm and O'Doherty, 2012), the lateral/dorsolateral PFC (Gläscher et al., 2010; Smittenaar et al., 2013), the medial PFC (Daw et al., 2011; O'Doherty, 2011; Seo et al., 2014), the orbitofrontal cortex (OFC; McDannald et al., 2014; Wilson et al., 2014), and the hippocampus (Bornstein and Daw, 2012; McKenzie et al., 2014; for a recent review see Daw and O'Doherty, 2013; Nakahara, 2014) etc.

4.2. Model-based RL under stress and depressive symptoms

Recall our reviewed studies above in Section 3, despite acquiring the framework of model-free RL, some of the reduced RPE signals do reflect impaired model-based RL, for instance, ACC and hippocampus. In theory, the learned helpless model of depression (Maier and Seligman, 1976; Maier and Watkins, 2005) predicts that under depression the prior belief that the situation is uncontrollable will discourage subjects from exploring and using model-based RL (Montague et al., 2012). Further, it has been proposed that cognitive function is closely associated to model-based RL (Chen, 2014). It is likely that the cognitive deficits in MDD (McDermott and Ebmeier, 2009; Gotlib and Joormann, 2010; Lee et al., 2012b; Snyder, 2013; Belzung et al., 2014b) as well as its underlying dysfunctional lateral and medial PFC, OFC and hippocampus (Clark et al., 2009; Russo and Nestler, 2013; Pizzagalli, 2014) would predispose patients to diminished model-based RL. Further considerable evidence from clinical and preclinical research has shown that rTMS of the dorsolateral PFC (see meta-analyses by Berlim et al., 2013, 2014; Gaynes et al., 2014) or deep brain stimulation (clinical: Mayberg et al., 2005; Lozano et al., 2008, 2012; McNeely et al., 2008; Neimat et al., 2008; Kennedy et al., 2011; Holtzheimer et al., 2012; Puigdemont et al., 2012; Riva-Posse et al., 2014. Preclinical: Hamani et al., 2010, 2012, 2014; Rea et al., 2014; Veerakumar et al., 2014) or optogenetic stimulation (Covington et al., 2010) of the medial PFC relieves depression. As it has been shown that similar stimulations may enhance cognition and memory (Bergfeld et al., 2013; Brunoni and Vanderhasselt, 2014), we speculate that cognition and memory or more specifically the model-based RL (Chen, 2014) might mediate the above therapeutic effect on depression by these stimulations.

Actually, research has already indicated impaired counterfactual thinking (one form of model-based RL; Lee et al., 2012a) in subjects with depressive symptoms. When asked to imagine future negative events and make counterfactuals about those events, compared to non-depressed subjects, subjects with mild-to-moderate depressive symptoms (BDI-II 10–23) generated more proportion of controllable counterfactuals (which could have been done to change the situation, e.g., "If only I had studied more..."). In contrast, severely depressed subjects (BDI-II 24+) generated counterfactuals that were less controllable (which could not have been done, e.g., "If only I were naturally brilliant..."), less reasonable (unlikely to be the cause of the outcome), and more characterological (e.g., "if only" a trait aspect of the self had been different) in nature (Markman and Miller, 2006. See also Markman and Weary, 1996 and discussion by Howlett and Paulus, 2013). In a later study, Quelhas et al. (2008) studied counterfactual thinking in a real-life situation in which subjects experienced a negative outcome (failure in an academic test). They found that depressed subjects (BDI-II > 17) were less likely to use counterfactual thinking than non-depressed subjects (BDI-II 0–5). Moreover, among those who did use counterfactual thinking, depressed subjects were less likely to benefit from it cognitively (e.g. feel more prepared for future events) and behaviorally (e.g. intention to change in the coming week). These findings together indicate dysfunctional counterfactual thinking therefore model-based RL in depression, consistent with impaired control over negative outcomes in depression we discussed in Section 3.3.

The hypothesis that the model-based RL may be attenuated has yet to be tested by computational research in depression but it has been studied in healthy subjects under stressful situations (Otto et al., 2013; Radenbach et al., 2015). Within the above framework, Otto et al. (2013) showed that under a stressful Cold pressor test, subjects reduced the employment of model-based but not model-free RL in a stress response-dependent way. That is, subjects with higher cortisol response showed more severely reduced use of model-based RL. Further, this reduction was moderated by working memory such that subjects with lower working memory were more sensitive to this stress effect. Radenbach et al. (2015) replicated the stress effect on model-based RL by Otto et al. (2013) with a potent psychosocial stressor (Trier social stress test). Subjects' cortisol response was negatively correlated with reduced use of the model-based RL. Besides, Radenbach et al. (2015) found that subjective ratings of arousal and chronic life stress in the past 12 months also moderated the stress effect. Those with higher subjective rating of arousal showed increased use of the model-based RL, while individuals with higher levels of chronic life stress in the past 12 months demonstrated more severely impaired model-based RL after acute stress induction, indicating the interplay between chronic and acute stress. These findings provide computational support for the observation that stress impairs the PFC structure and function (Arnsten, 2009; McEwen and Morrison, 2013) and shifts goal-directed to habitual control (Schwabe and Wolf, 2011; Schwabe, 2013). They also go beyond and further reveal important moderating factors that contribute to the individual vulnerability to stress, suggesting that stress induced impairment in model-based RL may play a role in the onset and maintenance of depression.

In addition, with a similar logic albeit different task, Blanco et al. (2013) and Maddox et al. (2014) examined model-free vs. model-based RL in healthy subjects with depressive symptoms (scored 16 or greater in the Center for Epidemiologic Studies Depression Scale). Blanco et al. (2013) employed a probabilistic Leapfrog task. Of two choices in this task, one brings a higher reward. However on any trial the inferior choice may increase its reward thus leapfrog over the other choice. Subjects could stick to the current

choice with higher reward (exploiting) or switch to see if the other choice has leapfrogged (exploring). Model-free RL explores with equal probability on each trial, whereas model-based RL explores more after each consecutive exploitive choice, since only the latter tracks uncertainty with a world model. By fitting computational models, Blanco et al. (2013) found that depressive subjects used model-based RL less than non-depressives (even on the model-based RL optimal trials), and that within depressive subjects, those have higher depressive scores were better fit by the model-free RL model.

Maddox et al. (2014) borrowed the dual-process paradigm in auditory category learning (ACL). Echoing the distinction between model-free and model-based RL, it is postulated that ACL can be achieved by a non-verbalizable reflexive system or a verbalizable reflective system. Maddox et al. (2014) found that depressive subjects showed a deficit in reflective-optimal ACL, but an advantage in reflexive-optimal ACL. Further, using a Mandarin tones ACL task and computational models of Striatal Pattern Classifier and reflective category learning system, Maddox et al. (2014) showed that the advantage in depressives was due to their faster, more accurate (surprisingly) and more frequent use of reflexive strategies.

Consistent with the theoretical prediction, these studies together suggest that stress and depressive symptoms attenuate model-based RL while favor model-free RL. It remains for future study to see if these findings hold true for MDD patients, whether this reduced model-based RL constitutes the underlying cause of MDD, and address the inquiry raised by Huys et al. (2013, 2015) that model-based rather than model-free RL is affected in MDD.

5. Conclusions and future directions

Although still in its infant stage, the computational approach has contributed substantially to our understanding of the etiology and pathophysiology of MDD. Going beyond previous studies with the Iowa Gambling Task (Must et al., 2013), the signal-detection task (Ragland et al., 2009; Pizzagalli, 2014), and the monetary incentive delay task or other pseudorandom reward tasks (Forbes and Dahl, 2012; Zhang et al., 2013) that have suggested impaired RL in MDD, computational research is gaining more insights into the underlying mechanism and specific processes of this impairment. As summarized in Table 1, research from the perspective of model-free RL suggests that MDD is characterized by impaired brain signals of RPE and expected value ('wanting'), decreased reward sensitivity ('liking'), learning rate (although somewhat inconsistent), and/or memory of previous reinforcement. Research employing the model-based approach found attenuated model-based RL in subjects with depressive symptoms and in subjects under stress. These dysfunctional features demonstrated by computational research hold enormous promise to serve as valuable intermediate phenotypes of depression (Gottesman and Gould, 2003; Rasetti and Weinberger, 2011; Bogdan et al., 2013). These intermediate phenotypes are particularly useful as they can link the general symptom-based descriptions to precise underlying molecular dysfunctions, and fill the explanatory gap between them by bridging clinical and preclinical discoveries (Montague et al., 2012). They may also improve the stagnant antidepressants development (Papassotiropoulos and de Quervaine, 2015).

We have already mentioned several possible directions for further research throughout this review. Here we would like to suggest other five lines of future inquiry at the level of computational methodology, clinical, systems, cellular/molecular/genetic and animal research.

5.1. Computational methodology

Most of our reviewed articles have employed the RW approach, either the original (Chase et al., 2010; Robinson et al., 2013; Friedel et al., 2014) or the modified (Dombrovski et al., 2010) RW model, or the RW model-based fMRI (Dombrovski et al., 2013, 2014). Only two have employed the TD model-based fMRI (Kumar et al., 2008; Gradin et al., 2011). Given the superiority of the TD model to the RW model (Niv and Schoenbaum, 2008; Niv, 2009) and the merit of the model-based fMRI (O'Doherty et al., 2007) as well as the model-based fMRI-EEG (Larsen and O'Doherty, 2014), future studies should try to utilize these advanced tools.

More importantly, although the two studies we reviewed with TD model (Kumar et al., 2008; Gradin et al., 2011) have incorporated the TD discount factor γ, both of them used a fixed parameter ($\gamma = 1$) therefore the individual data was actually not analyzed. Further, in the current RL literature, γ is typically perceived as time-independent. That is, an individual has a stable discount factor along the time scale. It is only recently that the temporal discounting or the "hyperbolic" (time-inconsistent, the discount factor is an increasing function of delay, see Takahashi, 2009 for a review) is incorporated into the computational framework (Alexander and Brown, 2010). Our previous study (Takahashi et al., 2008b) showed that the discount factor in patients with MDD is more time-dependent than healthy subjects (see Fig. 4 for temporal discounting curves for gain and loss in healthy subjects and MDD). They devalue rewards in the near future but value those in the distant future. Also, patients are more sensitive to delayed punishment (loss) in the distant future, in comparison to healthy subjects (Takahashi et al., 2008b). This indicates that patients with MDD may have exaggerated anxiety for potential bad events in the distant future. Further, patients with suicide behaviors may have specific features of temporal discounting (Dombrovski et al., 2011; Takahashi, 2011b). As we have shown that psychological time differs between waiting for delayed gain and loss (Han and Takahashi, 2012), it can be speculated that depressed patients have marked differences in psychological time between delayed gain and loss (Fig. 5). They may perceive gain as much farther while loss as nearer than healthy subjects. Prospective studies can employ the "hyperbolic" computational framework to advance our understanding of these psychophysical effects.

Further, although we mentioned that incentive 'wanting' can be parsed apart from cognitive 'wanting', throughout our discussion we treated these two as if they are the same. It has to be noted that the differentiation of incentive 'wanting' from cognitive 'wanting' is the essence of the incentive salience theory (Berridge and Robinson, 1998; Berridge, 2007, 2012; Berridge et al., 2009; Berridge and O'Doherty, 2014). Zhang et al. (2009) recently suggested a computational model that incorporates a dynamic physiological factor (kappa), which changes according to physiological states such as appetite, satiety or drug-state changes, and which thus modulates incentive 'wanting' generated from the learned value of a relevant conditioned stimulus for reward (cognitive 'wanting'). In Section 3.4 we have reviewed evidence suggesting reduced cognitive 'wanting', reduced willingness to exert effort for rewards, and deficient behavioral approach. It will be interesting for future research to examine incentive 'wanting' in depression using the Zhang et al. (2009) or similar computational model of incentive salience.

Finally, due to a lack of research, some important moderating or confounding factors are actually not addressed in the present review. For instance, regarding the probabilistic learning task, all of our reviewed studies employed the instrumental paradigm except Kumar et al. (2008) and Robinson et al. (2013), who used the Pavlovian paradigm. In Pavlovian learning, subjects predict the reward value by a conditioned cue. In instrumental learning, on the other hand, subjects have to predict which cue will be rewarded and

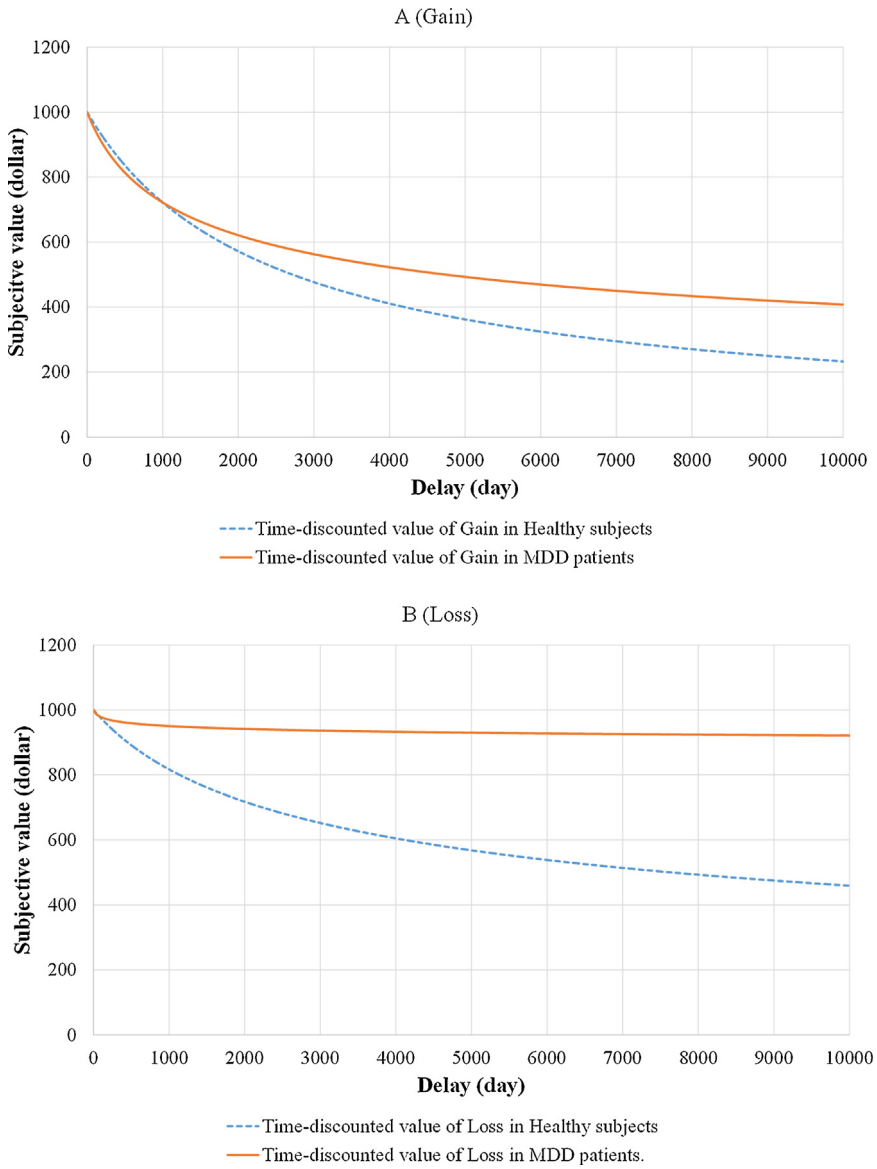


Fig. 4. Temporal discounting curves for gain (A) and loss (B) (from Takahashi et al., 2008b data). Temporal discounting is well-described with the following q-exponential time-discount model for gain and loss (Han and Takahashi, 2012).

$$v(x,t) = v(x,0)/(1+k(1-q)t)^{1/(1-q)}$$

where k indicates time-discount rate at delay = 0 and q (<1) indicates the deviation from exponential (i.e., time-consistent) temporal discounting ($q = 1$ corresponds to exponential discounting where time-discount rate is time-independent, while $q = 0$ corresponds to hyperbolic discounting). We can see that temporal discounting is faster (steeper temporal discount curve) in the near future in comparison to the distant future, indicating that time-discount rate is time-dependent (i.e., time-discount rate is a decreasing function of delay; i.e., $q < 1$). Also, temporal discounting is faster (steeper) for gain than loss (the sign effect in temporal discounting). Takahashi et al. (2008b) demonstrated that MDD patients are more time-inconsistent in temporal discounting of gain and loss, and the sign effect in temporal discounting was greater for MDD patients. MDD patients devalue rewards in the near future (delay < 1000 days) but value those in the distant future (delay > 1000 days) more than healthy subjects. In the meantime, they devalue delayed punishment (both in the near and distant future) less strongly than healthy subjects.

choose among cues. Pavlovian learning and instrumental learning seem to depend on different neural substrates (Liljeholm and O'Doherty, 2012; Daw and O'Doherty, 2013; Chen et al., 2014). Other factors such as the individual vs. fixed learning, reinforcer type, etc. (Chase et al., 2015) may also affect the behavioral and neural results of RL. Future research should try to analyze the effects of these factors.

5.2. Clinical considerations

We list several promising directions for research at the clinical level. Firstly, research has shown that several well documented factors that moderate the risk of depression, such as age, sex, and socioeconomic status (see a review by Gotlib and Hammen, 2014) also affect dopamine activity (Bäckman et al., 2010; Wahlstrom et al., 2010; Gillies et al., 2014; Sinclair et al., 2014) and/or PFC and hippocampal reactivity to stress (Lupien et al., 2009; McEwen and Gianaros, 2010; Brito and Noble, 2014; McEwen and Morrison, 2013). We have mentioned above in Section 3.2 that animal literature suggests that stress under various paradigms, such as acute vs. chronic, social isolation and restraint vs. cold and footshock, may affect dopamine neurotransmission in different ways (see also Pizzagalli, 2014). Future research of how these factors influence RL may shed light on the etiology of depression.

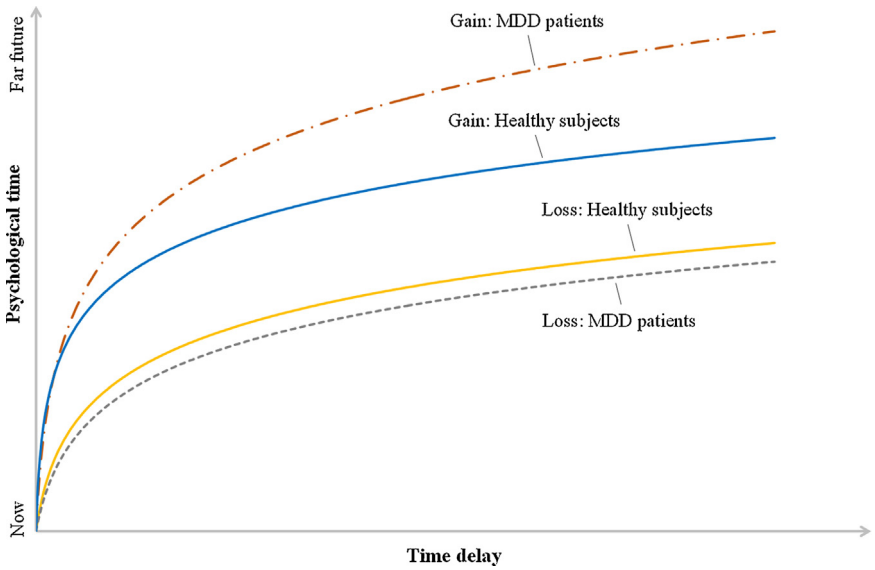


Fig. 5. Schematic psychological time for gain and loss. Psychological time (τ) is logarithmic in terms of physical time t:

$$\tau(t) = \alpha \log(1 + \beta t)$$

where α and β are free parameters (Han and Takahashi, 2012). Also, our previous theoretical study (Takahashi, 2011a) indicates that α and β are modulated by dopamine and serotonin, respectively. Logarithmic psychological time explains time-inconsistent temporal discounting for gain and loss (Takahashi, 2009; Han and Takahashi, 2012). This perception error in future time during temporal discounting may be exaggerated in MDD patients. Specifically, MDD patients may have prolonged psychological time for delayed gain (i.e., hyposensitivity to future reward) but reduced psychological time for delayed loss (hypersensitivity to future punishment), in comparison to healthy subjects.

Secondly, our reviewed studies have examined RL in the anhedonia or melancholic MDD, MDD unresponsive to treatment, and MDD with suicide ideations and attempts. RL in other subtypes of MDD, notably the atypical MDD remains to be addressed. Atypical MDD, or depression with atypical features, is characterized by a lack of anhedonia (patients show mood reactivity to actual or potential positive events), overeating, oversleeping interpersonal rejection sensitivity, and leaden paralysis (feelings of physical heaviness like lead) (American Psychiatric Association, 2014; though the appropriateness of this diagnosis itself is still controversial, see Lam and Stewart, 1996; Parker et al., 2002; Quitkin et al., 2003; Akiskal and Benazzi, 2005; Stewart et al., 2009; Thase, 2009). It shows preferable response to MAO inhibitors than tricyclic antidepressants (Lam and Stewart, 1996; Quitkin et al., 2003; Stewart, 2007). Given the influence of MAO inhibitors on dopamine, it will be stimulating for future research to investigate the likely specific link between RL and atypical MDD.

Moreover, the dopamine system has long been target for treating depression (Kapur and Mann, 1992; Nestler and Carlezon, 2006; Dunlop and Nemeroff, 2007, 2010; Gershon et al., 2007; Russo and Nestler, 2013). Besides above reviewed dopamine-related reagents in Section 3.2 (e.g. pramipexole), several atypical antipsychotics, e.g., aripiprazole, quetiapine, and olanzapine have consistently been demonstrated to be effective and regularly used adjunctively with antidepressants in treatment-resistant depression (Kato and Chang, 2013; Wright et al., 2013; McIntyre et al., 2014; Turner et al., 2014). Dopamine has been proposed to be an underlying pathway of this add-on effect (Rogóż, 2013). As another line of research, recently several dopamine-related drugs have shown to be effective in improving RL performance and 'wanting' (motivation). For instance, acute administration of amisulpride, a dopamine D2 receptor antagonist, enhanced the striatal PE coding and ventromedial PFC tracking of learned value in a later transfer phase of RL, which in turn predicted performance (Jocham et al., 2011). Acute administration of the dopamine precursor levodopa enhanced model-based over model-free RL (Wunderlich et al., 2012) and restored neural signals of RPE in aged people, improving their learning rate and task performance up to the level of young adults (Chowdhury et al., 2013). Moreover, administration of the dopamine agonist d-amphetamine enhanced the willingness to exert effort in healthy subjects, particularly when reward probability was low (Wardle et al., 2011). Similarly, the dopamine precursor levodopa enhanced the subjective expectation of pleasure when imaging future positive life events (Sharot et al., 2009) while reduced negative expectations and increased an optimism bias by impairing the ability to update belief in response to negative information (Sharot et al., 2012). Therefore it is interesting to investigate whether the above mentioned dopamine-related drugs exert their antidepressant effect through the mediation of RL, e.g., by normalizing RPE signals, improving learning rate, enhancing model-based RL and/or increasing 'wanting', etc. With regard to learning, taking this inquiry one step further, Roiser et al. (2012) argued that the cognitive mechanism, i.e. normalizing the negative information processing biases, is a common pathway of the therapeutic effect of both antidepressant drugs and psychological therapies. Whereas antidepressants achieve this by a bottom-up approach of gradually altering the brain's processing or learning of affective stimuli, psychological therapies resolve the high-order cognitive biases and train cognitive control directly in a top-down manner (Roiser et al., 2012). Following this proposal, we predict that antidepressants might work more through the model-free RL, while psychological therapies may more specifically enhance model-based RL. Future research is warranted to test this prediction. This not only improves our understanding of the therapeutic effect of antidepressants and psychological therapies, but also contributes to our current knowledge of the role of RL in depression. Besides dopamine, drug targeting the opioid/endocannabinoid system is another area of future antidepressants development, as they may normalize the attenuated 'liking' in depression.

Finally, several authors (e.g. Lee et al., 2012b; Roiser et al., 2012; Martinez-Aran and Vieta, 2015; Solé et al., 2015) have suggested cognitive function as a target for early detection and treatment of MDD. The present review further highlights RL as a more refined candidate undertaking this role. As encouraging preliminary data, Bress et al. (2013) showed that reduced feedback negativity, which encodes a PE (Holroyd and Coles, 2002; Nieuwenhuis et al., 2004; Walsh and Anderson, 2012), predicted the onset of MDD over two years in adolescents, controlling for baseline neuroticism and depressive symptoms. Using the signal-detection approach, Vrieze et al. (2013) further reported that diminished behavioral reward learning in MDD increased the odds of being unresponsive to 8 weeks of treatment (odds ratio 7.84) (see Forbes et al., 2010; Morgan et al., 2013 for the predictive value of the striatal and medial PFC activation in pseudorandom reward tasks, and see Pizzagalli, 2014; Proudfit, 2014 for a discussion of other related findings). Since our reviewed computational studies have all been correlational in nature, longitudinal studies are urgently needed to investigate the causality between impaired RL and depression, and to examine the potential usefulness of the intermediate phenotypes suggested by computational research (Table 1) as early detection biomarkers and therapeutic intervention targets for MDD. In addition, possibly distinct features of RL in different subtypes and etiology of MDD and the prediction that antidepressants work more through model-free RL whereas psychological therapies more specifically enhance model-based RL may provide a new possibility to detect which patients will benefit the most from what kind of treatment therefore achieve individualized therapeutic treatment (Stephan and Mathys, 2014; Wiecki et al., 2015).

5.3. Systems level

At the systems, brain area or network level, recent work has revealed a significant role of the lateral habenula, which sends projections to the VTA (Lammel et al., 2012), in coding 'negative value' signals and the hyperactivity of lateral habenula in depression (Lecca et al., 2014; Proulx et al., 2014). Similarly dorsal raphe nucleus, a key area involved in MDD (Maier and Watkins, 2005; Michelsen et al., 2007; Lowry et al., 2008; Deakin, 2013), has also been linked to negative PE (Berg et al., 2014). Computational model-based fMRI or model-based fMRI-EEG can be designed to detect the potential role of these areas in RL and MDD.

Moreover, Dombrovski et al. (2014) has observed disrupted functional connectivity between the striatum and PFC following unpredicted rewards in aged patients suffering from MDD. It has been found that the white matter tract strength between the posterior putamen and premotor cortex predicted performance of the model-free RL while that between the caudate and ventromedial PFC predicted performance of the model-based RL (de Wit et al., 2012). In another study, Chaudhury et al. (2013) showed that optogenetic stimulation of the VTA-nucleus accumbens projection was sufficient to produce depressive-like behaviors, while inhibition of it elicited opposite effects. In contrast, inhibition of the VTA-medial PFC projection elicited depressive-like effects while stimulation of it had no effect. The precise explanation still remains unsure but we believe future studies especially those with advanced techniques such as optogenetics (see a review or perspective by Belzung et al., 2014a; Tye, 2014) can improve our knowledge of the role of these networks in the onset and treatment of MDD (for excellent discussions of RL at the network or interaction level, see Frank and Claus, 2006; Roeper, 2013; Buschman and Miller, 2014; Chatham et al., 2014. For discussions of depression and stress at the network level, see Price and Drevets, 2010; Russo and Nestler, 2013; Walsh and Han, 2014).

5.4. Cellular, molecular and genetic studies

Cellular, molecular and genetic study of the neurochemical and neuroendocrinal abnormalities is another area of future research. Our understanding of the dopamine neurotransmission has progressed greatly in the past decades (Fig. 2B). As described above in Section 3.2, the impairment in dopamine neurotransmission in depression may arise from the process of synthesis, degradation, release, uptake, reuptake, and/or receptor binding (e.g. D2-like vs. D1-like) etc. Whether these different cellular and molecular mechanisms lead to distinct features of dysfunctional RL remain unknown.

Besides dopamine, human and animal research has established the abnormalities of serotonin (Maier and Watkins, 2005; Berton and Nestler, 2006; Dayan and Huys, 2009; Blier and El Mansari, 2013; Deakin, 2013), noradrenaline (Berton and Nestler, 2006), and the hypothalamic-pituitary-adrenal (HPA) axis (de Kloet et al., 2005; Hammen, 2005; Maier and Watkins, 2005; Lupien et al., 2009; Bogdan et al., 2013) in the pathophysiology of MDD and their role in the treatment. In the meantime, serotonin plays a major role in aversive processing (Doya, 2008; Dayan and Huys, 2009; Robbins and Arnsten, 2009; Boureau and Dayan, 2011) and may also be associated to reward processing (Ullsperger et al., 2014) such as moderating the use of goal-directed (model-based) vs. habitual (model-free) behavioral control (Worbe et al., 2015). Noradrenaline is involved in estimation of uncertainty and risk (Yu and Dayan, 2005; Doya, 2008; Preuschoff et al., 2011; Nassar et al., 2012; Payzan-LeNestour et al., 2013) and temporal discounting (Takahashi et al., 2007, 2008a, 2010). It also has profound influences on cognition (Chamberlain and Robbins, 2013). Stress and HPA axis activates the mesolimbic dopamine system (Koob and Kreek, 2007; Starcke and Brand, 2012; also see studies reviewed above) and affects temporal discounting (Takahashi, 2004; Takahashi et al., 2010).

Despite these findings, the independent and collective influence of these neurochemical and neuroendocrinal factors on the impairment of RL in MDD is far from clear. Future research should try to address this by means of cellular, molecular and genetic studies.

5.5. Animal research

There has been extensive animal literature on habitual and goal-directed behavior (Dickinson and Charnock, 1985; Balleine, 2005; Balleine and O'Doherty, 2010) and animal models of RL have long been available (Markou et al., 2013; McDannald et al., 2014; Wilson et al., 2014). However to our knowledge none of them has been combined with animal models of depression (Henn and Vollmayr, 2005; Berton and Nestler, 2006; Nestler and Hyman, 2010; Caldarone et al., 2015; see also Berton et al., 2012; Dzirasa and Covington, 2012 for insightful discussions). Initiation of animal research allows subsequent pioneering pharmacological intervention as well as extensive cellular, molecular and genetic investigation, thus it will be another illuminating direction for future research.

We believe that these areas of research will propel us toward a better understanding of the disease, promote its prevention and treatment and may ultimately bring a 'paradigm shift' to psychiatry.

Conflict of interest

The authors report no financial or other relationship that is relevant to the subject of this article. S.N. has received honoraria from GlaxoSmithKline, Eisai, Pfeizer, Daiichi-Sankyo, Meiji Seika Pharma, Ono Pharmaceutical and Eli Lilly, and has received research/grant support from Pfeizer, Eli Lilly, Eisai and Ono Pharmaceutical. T.I. has received honoraria from GlaxoSmithKline,

Pfeizer, Astellas, Eli Lilly, Mitsubishi Tanabe Pharma, Mochida Pharmaceutical, Otsuka Pharmaceutical, Meiji Seika Pharma, Asahi Kasei Pharma, Shionogi, Janssen Pharmaceutical, Takeda Pharmaceutical, MSD and Yoshitomi Pharmaceutical, has received research/grant support from Otsuka Pharmaceutical, Mitsubishi Tanabe Pharma and Eli Lilly, and is a member of the advisory boards of GlaxoSmithKline, Eli Lilly, Mochida Pharmaceutical and Mitsubishi Tanabe Pharma. I.K. has received honoraria from Astellas, Dainippon Sumitomo Pharma, Eisai, Eli Lilly, Janssen Pharmaceutical, Kyowa Hakko Kirin, Meiji Seika Pharma, MSD, Nippon Chemiphar, Novartis Pharma, Otsuka Pharmaceutical, Pfeizer, Tanabe Mitsubishi Pharma, Shionogi and Yoshitomi Pharmaceutical, and has received research/grant support from AbbVie GK, Asahi Kasei Pharma, Astellas, Boehringer Ingelhaim, Chugai Pharmaceutical, Daiichi Sankyo, Dainippon Sumitomo Pharma, Eisai, Eli Lilly, GlaxoSmithKline, Kyowa Hakko Kirin, Meiji Seika Pharma, MSD, Novartis Pharma, Ono Pharmaceutical, Otsuka Pharmaceutical, Pfeizer, Takeda Pharmaceutical, Tanabe Mitsubishi Pharma, Shionogi and Yoshitomi Pharmaceutical, and is a member of the advisory board of Dainippon Sumitomo Pharma and Tanabe Mitsubishi Pharma. The other authors declare that they have no actual or potential conflict of interest.

Acknowledgments

The authors thank three anonymous reviewers for their thoughtful comments and suggestions to improve the earlier manuscript. The authors are grateful to Peter Dayan, Kent C. Berridge, Minori Koga, and Yukiko Ogura for their valuable discussions and/or comments. The authors also thank Alexandre Y. Dombrovski for answering questions about their research. The research reported in this paper was supported by a grant from the Grant-in-Aid for Scientific Research (Global COE and on Innovative Areas, 23118001; Adolescent Mind and Self-Regulation) to T.T. from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Aggarwal, M., Hyland, B.I., Wickens, J.R., 2012. Neural control of dopamine neurotransmission: implications for reinforcement learning. Eur. J. Neurosci. 35 (7), 1115–1123.

Agrawal, A., Nelson, E.C., Littlefield, A.K., Bucholz, K.K., Degenhardt, L., Henders, A.K., Madden, P.A., Martin, N.G., Montgomery, G.W., Pergadia, M.L., Sher, K.J., Heath, A.C., Lynskey, M.T., 2012. Cannabinoid receptor genotype moderation of the effects of childhood physical abuse on anhedonia and depression. Arch. Gen. Psychiatry 69 (7), 732–740.

Ainsworth, K., Smith, S.E., Zetterström, T.S., Pei, Q., Franklin, M., Sharp, T., 1998. Effect of antidepressant drugs on dopamine D1 and D2 receptor expression and dopamine release in the nucleus accumbens of the rat. Psychopharmacology (Berlin) 140 (4), 470–477.

Akiskal, H.S., Benazzi, F., 2005. Atypical depression: a variant of bipolar II or a bridge between unipolar and bipolar II? J. Affect. Disord. 84 (2–3), 209–217.

Alexander, W.H., Brown, J.W., 2010. Hyperbolically discounted temporal difference learning. Neural Comput. 22 (6), 1511–1527.

American Psychiatric Association, 2014. The Fifth Edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5). APA, Washington, DC.

Amsterdam, J.D., Newberg, A.B., Soeller, I., Shults, J., 2012. Greater striatal dopamine transporter density may be associated with major depressive episode. J. Affect. Disord. 141 (2–3), 425–431.

Antypa, N., Drago, A., Serretti, A., 2013. The role of COMT gene variants in depression: bridging neuropsychological, behavioral and clinical phenotypes. Neurosci. Biobehav. Rev. 37 (8), 1597–1610.

Arnsten, A.F., 2009. Stress signalling pathways that impair prefrontal cortex structure and function. Nat. Rev. Neurosci. 10 (6), 410–422.

Aouizerate, B., Cuny, E., Bardinet, E., Yelnik, J., Martin-Guehl, C., Rotge, J.Y., Rougier, A., Bioulac, B., Tignol, J., Mallet, L., Burbaud, P., Guehl, D., 2009. Distinct striatal targets in treating obsessive-compulsive disorder and major depression. J. Neurosurg. 111 (4), 775–779.

Bäckman, L., Lindenberger, U., Li, S.C., Nyberg, L., 2010. Linking cognitive aging to alterations in dopamine neurotransmitter functioning: recent data and future avenues. Neurosci. Biobehav. Rev. 34 (5), 670–677.

Bahi, A., Al Mansouri, S., Al Memari, E., Al Ameri, M., Nurulain, S.M., Ojha, S., 2014. β-Caryophyllene, a CB2 receptor agonist produces multiple behavioral changes relevant to anxiety and depression in mice. Physiol. Behav. 135, 119–124.

Bali, A., Randhawa, P.K., Jaggi, A.S., 2015. Stress and opioids: role of opioids in modulating stress-related behavior and effect of stress on morphine conditioned place preference. Neurosci. Biobehav. Rev. 51C, 138–150.

Balleine, B.W., 2005. Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. Physiol. Behav. 86, 717–730.

Bannon, M.J., 2005. The dopamine transporter: role in neurotoxicity and human disease. Toxicol. Appl. Pharmacol. 204 (3), 355–360.

Barone, P., Scarzella, L., Marconi, R., Antonini, A., Morgante, L., Bracco, F., Zappia, M., Musch, B., Depression/Parkinson Italian Study Group, 2006. Pramipexole versus sertraline in the treatment of depression in Parkinson's disease: a national multicenter parallel-group randomized study. J. Neurol. 253 (5), 601–607.

Balleine, B.W., O'Doherty, J.P., 2010. Human and rodent homologies in motor control: cortico-striatal determinants of goal-directed and habitual action. Neuropsychopharmacology 5 (1), 48–69.

Baskin-Sommers, A.R., Foti, D., 2015. Abnormal reward functioning across substance use disorders and major depressive disorder: considering reward as a transdiagnostic mechanism. Int. J. Psychophysiol. (February), http://dx.doi.org/10.1016/j.ijpsycho.2015.01.011, pii:S0167-8760(15)00028-8.

Beaulieu, J.M., Gainetdinov, R.R., 2011. The physiology, signaling, and pharmacology of dopamine receptors. Pharmacol. Rev. 63 (1), 182–217.

Beck, A.T., Alford, B.A., 2009. Depression: Causes and Treatment. University of Pennsylvania Press, Philadelphia.

Belujon, P., Grace, A.A., 2014. Restoring mood balance in depression: ketamine reverses deficit in dopamine-dependent synaptic plasticity. Biol. Psychiatry 76 (12), 927–936.

Belzung, C., Turiault, M., Griebel, G., 2014a. Optogenetics to study the circuits of fear- and depression-like behaviors: a critical analysis. Pharmacol. Biochem. Behav. 122, 144–157.

Belzung, C., Willner, P., Philippot, P., 2014b. Depression: from psychopathology to pathophysiology. Curr. Opin. Neurobiol. 30C, 24–30.

Berg, B.A., Schoenbaum, G., McDannald, M.A., 2014. The dorsal raphe nucleus is integral to negative prediction errors in Pavlovian fear. Eur. J. Neurosci., http://dx.doi.org/10.1111/ejn.12676

Bergfeld, I.O., Mantione, M., Hoogendoorn, M.L., Denys, D., 2013. Cognitive functioning in psychiatric disorders following deep brain stimulation. Brain Stimul. 6 (4), 532–537.

Berlim, M.T., Van den Eynde, F., Daskalakis, Z.J., 2013. A systematic review and meta-analysis on the efficacy and acceptability of bilateral repetitive transcranial magnetic stimulation (rTMS) for treating major depression. Psychol. Med. 43 (11), 2245–2254.

Berlim, M.T., van den Eynde, F., Tovar-Perdomo, S., Daskalakis, Z.J., 2014. Response, remission and drop-out rates following high-frequency repetitive transcranial magnetic stimulation (rTMS) for treating major depression: a systematic review and meta-analysis of randomized, double-blind and sham-controlled trials. Psychol. Med. 44 (2), 225–239.

Berridge, K.C., 2007. The debate over dopamine's role in reward: the case for incentive salience. Psychopharmacology (Berlin) 191, 391–431.

Berridge, K.C., 2012. From prediction error to incentive salience: mesolimbic computation of reward motivation. Eur. J. Neurosci. 35 (7), 1124–1143.

Berridge, K.C., O'Doherty, J.P., 2014. From experienced utility to decision utility. In: Glimcher, P., Fehr, E. (Eds.), Neuroeconomics: Decisions and the Brain. , 2nd ed. Academic Press, pp. 335–354.

Berridge, K.C., Robinson, T.E., 1998. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res. Rev. 28 (3), 309–369.

Berridge, K.C., Robinson, T.E., Aldridge, J.W., 2009. Dissecting components of reward: 'liking', 'wanting', and learning. Curr. Opin. Pharmacol. 9 (1), 65–73.

Bershad, A.K., Jaffe, J.H., Childs, E., de Wit, H., 2015. Opioid partial agonist buprenorphine dampens responses to psychosocial stress in humans. Psychoneuroendocrinology 52, 281–288.

Berton, O., Hahn, C.G., Thase, M.E., 2012. Are we getting closer to valid translational models for major depression? Science 338 (6103), 75–79.

Berton, O., Nestler, E.J., 2006. New approaches to antidepressant drug discovery: beyond monoamines. Nat. Rev. Neurosci. 7 (2), 137–151.

Bewernick, B.H., Hurlemann, R., Matusch, A., Kayser, S., Grubert, C., Hadrysiewicz, B., Axmacher, N., Lemke, M., Cooper-Mahkorn, D., Cohen, M.X., Brockmann, H., Lenartz, D., Sturm, V., Schlaepfer, T.E., 2010. Nucleus accumbens deep brain stimulation decreases ratings of depression and anxiety in treatment-resistant depression. Biol. Psychiatry 67 (2), 110–116.

Bewernick, B.H., Kayser, S., Sturm, V., Schlaepfer, T.E., 2012. Long-term effects of nucleus accumbens deep brain stimulation in treatment-resistant depression: evidence for sustained efficacy. Neuropsychopharmacology 37 (9), 1975–1985.

Blanco, N.J., Otto, A.R., Maddox, W.T., Beevers, C.G., Love, B.C., 2013. The influence of depression symptoms on exploratory decision-making. Cognition 129 (3), 563–568.

Blier, P., El Mansari, M., 2013. Serotonin and beyond: therapeutics for major depression. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 368 (1615), 20120536.

Bogdan, R., Nikolova, Y.S., Pizzagalli, D.A., 2013. Neurogenetics of depression: a focus on reward processing and stress sensitivity. Neurobiol. Dis. 52, 12–23.

Bornstein, A.M., Daw, N.D., 2012. Dissociating hippocampal and striatal contributions to sequential prediction learning. Eur. J. Neurosci. 35 (7), 1011–1023.

Boureau, Y.L., Dayan, P., 2011. Opponency revisited: competition and cooperation between dopamine and serotonin. Neuropsychopharmacology 36 (1), 74–97.

Bress, J.N., Foti, D., Kotov, R., Klein, D.N., Hajcak, G., 2013. Blunted neural response to rewards prospectively predicts depression in adolescent girls. Psychophysiology 50 (1), 74–81.
Brito, N.H., Noble, K.G., 2014. Socioeconomic status and structural brain development. Front. Neurosci. 8, 276, http://dx.doi.org/10.3389/fnins.2014.00276
Brunoni, A.R., Vanderhasselt, M.A., 2014. Working memory improvement with non-invasive brain stimulation of the dorsolateral prefrontal cortex: a systematic review and meta-analysis. Brain Cogn. 86, 1–9.
Brunswick, D.J., Amsterdam, J.D., Mozley, P.D., Newberg, A., 2003. Greater availability of brain dopamine transporters in major depression shown by [99m Tc]TRODAT-1 SPECT imaging. Am. J. Psychiatry 160 (10), 1836–1841.
Burnett, F.E., Scott, L.V., Weaver, M.G., Medbak, S.H., Dinan, T.G., 1999. The effect of naloxone on adrenocorticotropin and cortisol release: evidence for a reduced response in depression. J. Affect. Disord. 53 (3), 263–268.
Burokas, A., Martín-García, E., Gutiérrez-Cuesta, J., Rojas, S., Herance, J.R., Gispert, J.D., Serra, M.Á., Maldonado, R., 2014. Relationships between serotonergic and cannabinoid system in depressive-like behavior: a PET study with [11C]-DASB. J. Neurochem. 130 (1), 126–135.
Buschman, T.J., Miller, E.K., 2014. Goal-direction and top-down control. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 369 (1655), 20130471.
Cabib, S., Puglisi-Allegra, S., 2012. The mesoaccumbens dopamine in coping with stress. Neurosci. Biobehav. Rev. 36 (1), 79–89.
Caldarone, B.J., Zachariou, V., King, S.L., 2015. Rodent models of treatment-resistant depression. Eur. J. Pharmacol. 753, 51–65.
Cassano, P., Lattanzi, L., Soldani, F., Navari, S., Battistini, G., Gemignani, A., Cassano, G.B., 2004. Pramipexole in treatment-resistant depression: an extended follow-up. Depress. Anxiety 20 (3), 131–138.
Castro, D.C., Berridge, K.C., 2014. Advances in the neurobiological bases for food 'liking' versus 'wanting'. Physiol. Behav., http://dx.doi.org/10.1016/j.physbeh.2014.05.022, pii:S0031-9384(14)00299-6.
Chamberlain, S.R., Robbins, T.W., 2013. Noradrenergic modulation of cognition: therapeutic implications. J. Psychopharmacol. 27 (8), 694–718.
Chase, H.W., Frank, M.J., Michael, A., Bullmore, E.T., Sahakian, B.J., Robbins, T.W., 2010. Approach and avoidance learning in patients with major depression and healthy controls: relation to anhedonia. Psychol. Med. 40 (3), 433–440.
Chase, H.W., Kumar, P., Eickhoff, S.B., Dombrovski, A.Y., 2015. Reinforcement learning models and their neural correlates: an activation likelihood estimation meta-analysis. Cogn. Affect. Behav. Neurosci., http://dx.doi.org/10.3758/s13415-015-0338-7
Chatham, C., Frank, M.J., Badre, D., 2014. Corticostriatal output gating during selection from working memory. Neuron 81, 930–942.
Chaudhry, F.A., Edwards, R.H., Fonnum, F., 2008. Vesicular neurotransmitter transporters as targets for endogenous and exogenous toxic substances. Annu. Rev. Pharmacol. Toxicol. 48, 277–301.
Chaudhury, D., Walsh, J.J., Friedman, A.K., Juarez, B., Ku, S.M., Koo, J.W., Ferguson, D., Tsai, H.C., Pomeranz, L., Christoffel, D.J., Nectow, A.R., Ekstrand, M., Domingos, A., Mazei-Robison, M.S., Mouzon, E., Lobo, M.K., Neve, R.L., Friedman, J.M., Russo, S.J., Deisseroth, K., Nestler, E.J., Han, M.H., 2013. Rapid regulation of depression-related behaviours by control of midbrain dopamine neurons. Nature 493 (7433), 532–536.
Chen, C., 2014. Intelligence moderates reinforcement learning: a mini-review of the neural evidence. J. Neurophysiol., http://dx.doi.org/10.1152/jn.00600.2014 (Published online: 3 September 2014).
Chen, C., Omiya, Y., Yang, S., 2014. Dissociating contributions of ventral and dorsal striatum to reward learning. J. Neurophysiol., http://dx.doi.org/10.1152/jn.00873.2014(Published online: 19 November 2014)
Chernoloz, O., El Mansari, M., Blier, P., 2012. Long-term administration of the dopamine D3/2 receptor agonist pramipexole increases dopamine and serotonin neurotransmission in the male rat forebrain. J. Psychiatry Neurosci. 37 (2), 113–121.
Chiuccariello, L., Houle, S., Miler, L., Cooke, R.G., Rusjan, P.M., Rajkowska, G., Levitan, R.D., Kish, S.J., Kolla, N.J., Ou, X., Wilson, A.A., Meyer, J.H., 2014. Elevated monoamine oxidase a binding during major depressive episodes is associated with greater severity and reversed neurovegetative symptoms. Neuropsychopharmacology 39 (4), 973–980.
Chowdhury, R., Guitart-Masip, M., Lambert, C., Dayan, P., Huys, Q., Duzel, E., Dolan, R., 2013. Dopamine restores reward prediction errors in old age. Nat. Neurosci. 16, 648–653.
Clark, L., Chamberlain, S.R., Sahakian, B.J., 2009. Neurocognitive mechanisms in depression: implications for treatment. Annu. Rev. Neurosci. 32, 57–74.
Cohen, M.X., 2007. Individual differences and the neural representations of reward expectation and reward prediction error. Soc. Cogn. Affect. Neurosci. 2 (1), 20–30.
Costa, V.D., Tran, V.L., Turchi, J., Averbeck, B.B., 2015. Reversal learning and dopamine: a Bayesian perspective. J. Neurosci. 35 (6), 2407–2416.
Covington III, H.E., Lobo, M.K., Maze, I., Vialou, V., Hyman, J.M., Zaman, S., et al., 2010. Antidepressant effect of optogenetic stimulation of the medial prefrontal cortex. J. Neurosci. 30, 16082–16090.
Cusin, C., Iovieno, N., Iosifescu, D.V., Nierenberg, A.A., Fava, M., Rush, A.J., Perlis, R.H., 2013. A randomized, double-blind, placebo-controlled trial of pramipexole augmentation in treatment-resistant major depressive disorder. J. Clin. Psychiatry 74 (7), e636–e641.
Daw, N.D., 2013. Advanced reinforcement learning. In: Glimcher, P., Fehr, E. (Eds.), Neuroeconomics: Decision Making and the Brain. , 2nd ed. Elsevier, San Diego, pp. 299–320.
Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., Dolan, R.J., 2011. Model-based influences on humans' choices and striatal prediction errors. Neuron 69 (6), 1204–1215.
Daw, N.D., Niv, Y., Dayan, P., 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. 8 (12), 1704–1711.
Daw, N.D., O'Doherty, J.P., 2013. Multiple systems for value learning. In: Glimcher, P., Fehr, E. (Eds.), Neuroeconomics: Decision Making and the Brain. , 2nd ed. Elsevier, San Diego, pp. 393–410.
Daw, N.D., Tobler, P., 2013. Value learning through reinforcement: the basics of dopamine and reinforcement learning. In: Glimcher, P., Fehr, E. (Eds.), Neuroeconomics: Decision Making and the Brain. , 2nd ed. Elsevier, San Diego, pp. 283–298.
Dayan, P., 2009. Goal-directed control and its antipodes. Neural Netw. 22, 213–219.
Dayan, P., Huys, Q.J., 2009. Serotonin in affective control. Annu. Rev. Neurosci. 32, 95–126.
Deakin, J., 2013. The origins of '5-HT and mechanisms of defence' by Deakin and Graeff: a personal perspective. J. Psychopharmacol. 27 (12), 1084–1089.
de Kloet, E.R., Joëls, M., Holsboer, F., 2005. Stress and the brain: from adaptation to disease. Nat. Rev. Neurosci. 6 (6), 463–475.
Delgado, M.R., Li, J., Schiller, D., Phelps, E.A., 2008. The role of the striatum in aversive learning and aversive prediction errors. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 363 (1511), 3787–3800.
Der-Avakian, A., Markou, A., 2012. The neurobiology of anhedonia and other reward-related deficits. Trends Neurosci. 35 (1), 68–77.
de Wit, S., Watson, P., Harsay, H.A., Cohen, M.X., van de Vijver, I., Ridderinkhof, K.R., 2012. Corticostriatal connectivity underlies individual differences in the balance between habitual and goal-directed action control. J. Neurosci. 32 (35), 12066–12075.
Dickinson, A., Charnock, D.J., 1985. Contingency effects with maintained instrumental reinforcement. Q. J. Exp. Psychol. B 37, 397–416.
Dillon, D.G., Dobbins, I.G., Pizzagalli, D.A., 2014. Weak reward source memory in depression reflects blunted activation of VTA/SN and parahippocampus. Soc. Cogn. Affect. Neurosci. 9 (10), 1576–1583.
Dimidjian, S., Barrera Jr., M., Martell, C., Muñoz, R.F., Lewinsohn, P.M., 2011. The origins and current status of behavioral activation treatments for depression. Annu. Rev. Clin. Psychol. 7, 1–38.
Dolan, R.J., Dayan, P., 2013. Goals and habits in the brain. Neuron 80 (2), 312–325.
Doll, B.B., Simon, D.A., Daw, N.D., 2012. The ubiquity of model-based reinforcement learning. Curr. Opin. Neurobiol. 22 (6), 1075–1081.
Dombrovski, A.Y., Clark, L., Siegle, G.J., Butters, M.A., Ichikawa, N., Sahakian, B.J., Szanto, K., 2010. Reward/Punishment reversal learning in older suicide attempters. Am. J. Psychiatry 167 (6), 699–707.
Dombrovski, A.Y., Szanto, K., Clark, L., Aizenstein, H.J., Chase, H.W., Reynolds III, C.F., Siegle, G.J., 2014. Corticostriatothalamic reward prediction error signals and executive control in late-life depression. Psychol. Med., http://dx.doi.org/10.1017/S0033291714002517 (Published online: 16 October 2014).
Dombrovski, A.Y., Szanto, K., Clark, L., Reynolds, C.F., Siegle, G.J., 2013. Reward signals, attempted suicide, and impulsivity in late-life depression. JAMA Psychiatry, http://dx.doi.org/10.1001/jamapsychiatry.2013.75
Dombrovski, A.Y., Szanto, K., Siegle, G.J., Wallace, M.L., Forman, S.D., Sahakian, B., Reynolds III, C.F., Clark, L., 2011. Lethal forethought: delayed reward discounting differentiates high- and low-lethality suicide attempts in old age. Biol. Psychiatry 70 (2), 138–144.
Domschke, K., Dannlowski, U., Ohrmann, P., Lawford, B., Bauer, J., Kugel, H., Heindel, W., Young, R., Morris, P., Arolt, V., Deckert, J., Suslow, T., Baune, B.T., 2008. Cannabinoid receptor 1 (CNR1) gene: impact on antidepressant treatment response and emotion processing in major depression. Eur. Neuropsychopharmacol. 18 (10), 751–759.
Downar, J., Geraci, J., Salomons, T.V., Dunlop, K., Wheeler, S., McAndrews, M.P., Bakker, N., Blumberger, D.M., Daskalakis, Z.J., Kennedy, S.H., Flint, A.J., Giacobbe, P., 2014. Anhedonia and reward-circuit connectivity distinguish nonresponders from responders to dorsomedial prefrontal repetitive transcranial magnetic stimulation in major depression. Biol. Psychiatry 76 (3), 176–185.
Doya, K., 2008. Modulators of decision making. Nat. Neurosci. 11 (4), 410–416.
Duncan, J., Johnson, S., Ou, X.M., 2012. Monoamine oxidases in major depressive disorder and alcoholism. Drug Discov. Ther. 6 (3), 112–122.
Dunlop, B.W., Nemeroff, C.B., 2007. The role of dopamine in the pathophysiology of depression. Arch. Gen. Psychiatry 64 (3), 327–337.
Dunlop, B.W., Nemeroff, C.B., 2010. The role of dopamine in the pathophysiology and treatment of major depressive disorder. In: Iversen, L., Iversen, S., Dunnett, S., Bjorklund, A. (Eds.), Dopamine Handbook. Oxford University Press, Inc., New York, pp. 572–589.
Dzirasa, K., Covington III, H.E., 2012. Increasing the validity of experimental models for depression. Ann. N. Y. Acad. Sci. 1265, 36–45.
Epstein, S., 1994. Integration of the cognitive and the psychodynamic unconscious. Am. Psychol. 49, 709–724.
Eriksen, J., Jørgensen, T.N., Gether, U., 2010. Regulation of dopamine transporter function by protein-protein interactions: new discoveries and methodological challenges. J. Neurochem. 113 (1), 27–41.
Eshel, N., Roiser, J.P., 2010. Reward and punishment processing in depression. Biol. Psychiatry 68 (2), 118–124.
Fan, Y., Chen, P., Li, Y., Zhu, M.Y., 2013. Effects of chronic social defeat on expression of dopamine β-hydroxylase in rat brains. Synapse 67 (6), 300–312.

Favier, M., Duran, T., Carcenac, C., Drui, G., Savasta, M., Carnicella, S., 2014. Pramipexole reverses Parkinson's disease-related motivational deficits in rats. Mov. Disord. 29 (7), 912–920.
Finberg, J.P., 2014. Update on the pharmacology of selective inhibitors of MAO-A and MAO-B: focus on modulation of CNS monoamine neurotransmitter release. Pharmacol. Ther. 143 (2), 133–152.
Forbes, E.E., Dahl, R.E., 2012. Research Review: altered reward function in adolescent depression: what, when and how? J. Child Psychol. Psychiatry 53 (1), 3–15.
Forbes, E.E., Olino, T.M., Ryan, N.D., Birmaher, B., Axelson, D., Moyles, D.L., Dahl, R.E., 2010. Reward-related brain function as a predictor of treatment response in adolescents with major depressive disorder. Cogn. Affect. Behav. Neurosci. 10 (1), 107–118.
Ford, C.P., 2014. The role of D2-autoreceptors in regulating dopamine neuron activity and transmission. Neuroscience 282C, 13–22.
Franco-Chaves, J.A., Mateus, C.F., Luckenbaugh, D.A., Martinez, P.E., Mallinger, A.G., Zarate Jr., C.A., 2013. Combining a dopamine agonist and selective serotonin reuptake inhibitor for the treatment of depression: a double-blind, randomized pilot study. J. Affect. Disord. 149 (1–3), 319–325.
Frank, M.J., Claus, E.D., 2006. Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making and reversal. Psychol. Rev. 113, 300–326.
Friedel, E., Schlagenhauf, F., Beck, A., Dolan, R.J., Huys, Q.J.M., Rapp, M.A., Heinz, A., 2014. The effects of life stress and neural learning signals on fluid intelligence. Eur. Arch. Psychiatry Clin. Neurosci. (Published online: 21 Aug 2014).
Friedman, A., Frankel, M., Flaumenhaft, Y., Merenlender, A., Pinhasov, A., Feder, Y., Taler, M., Gil-Ad, I., Abeles, M., Yadid, G., 2009. Programmed acute electrical stimulation of ventral tegmental area alleviates depressive-like behavior. Neuropsychopharmacology 34 (4), 1057–1066.
Friedman, A., Friedman, Y., Dremencov, E., Yadid, G., 2008. VTA dopamine neuron bursting is altered in an animal model of depression and corrected by desipramine. J. Mol. Neurosci. 34 (3), 201–209.
Friedman, A., Lax, E., Abraham, L., Tischler, H., Yadid, G., 2012. Abnormality of VTA local field potential in an animal model of depression was restored by patterned DBS treatment. Eur. Neuropsychopharmacol. 22 (1), 64–71.
Friedman, A.K., Walsh, J.J., Juarez, B., Ku, S.M., Chaudhury, D., Wang, J., Li, X., Dietz, D.M., Pan, N., Vialou, V.F., Neve, R.L., Yue, Z., Han, M.H., 2014. Enhancing depression mechanisms in midbrain dopamine neurons achieves homeostatic resilience. Science 344 (6181), 313–319.
García-Gutiérrez, M.S., Pérez-Ortiz, J.M., Gutiérrez-Adán, A., Manzanares, J., 2010. Depression-resistant endophenotype in mice overexpressing cannabinoid CB(2) receptors. Br. J. Pharmacol. 160 (7), 1773–1784.
Garrison, J., Erdeniz, B., Done, J., 2013. Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. Neurosci. Biobehav. Rev. 37 (7), 1297–1310.
Gaynes, B.N., Lloyd, S.W., Lux, L., Gartlehner, G., Hansen, R.A., Brode, S., Jonas, D.E., Swinson Evans, T., Viswanathan, M., Lohr, K.N., 2014. Repetitive transcranial magnetic stimulation for treatment-resistant depression: a systematic review and meta-analysis. J. Clin. Psychiatry 75 (5), 477–489.
Gazit, T., Friedman, A., Lax, E., Samuel, M., Zahut, R., Katz, M., Abraham, L., Tischler, H., Teicher, M., Yadid, G., 2015. Programmed deep brain stimulation synchronizes VTA gamma band field potential and alleviates depressive-like behavior in rats. Neuropharmacology 91, 135–141.
Gerfen, C.R., Surmeier, D.J., 2011. Modulation of striatal projection systems by dopamine. Annu. Rev. Neurosci. 34, 441–466.
Gershon, A.A., Vishne, T., Grunhaus, L., 2007. Dopamine D2-like receptors and the antidepressant response. Biol. Psychiatry 61 (2), 145–153.
Gersner, R., Toth, E., Isserles, M., Zangen, A., 2010. Site-specific antidepressant effects of repeated subconvulsive electrical stimulation: potential role of brain-derived neurotrophic factor. Biol. Psychiatry 67 (2), 125–132.
Gillies, G.E., Virdee, K., McArthur, S., Dalley, J.W., 2014. Sex-dependent diversity in ventral tegmental dopaminergic neurons and developmental programing: a molecular, cellular and behavioral analysis. Neuroscience 282C, 69–85.
Gläscher, J., Daw, N., Dayan, P., O'Doherty, J.P., 2010. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron 66 (4), 585–595.
Glimcher, P.W., 2011. Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. Proc. Natl. Acad. Sci. U.S.A. 108 (Suppl. 3), 15647–15654.
Glimcher, P., Fehr, E., 2013. Neuroeconomics: Decision Making and the Brain, 2nd ed. Elsevier, San Diego.
Goldwater, D.S., Pavlides, C., Hunter, R.G., Bloss, E.B., Hof, P.R., McEwen, B.S., Morrison, J.H., 2009. Structural and functional alterations to rat medial prefrontal cortex following chronic restraint stress and recovery. Neuroscience 164 (2), 798–808.
Gorzalka, B.B., Hill, M.N., 2011. Putative role of endocannabinoid signaling in the etiology of depression and actions of antidepressants. Prog. Neuropsychopharmacol. Biol. Psychiatry 35 (7), 1575–1585.
Gotlib, I.H., Hammen, C.L., 2014. Handbook of Depression, 3rd ed. Guilford Press, New York.
Gotlib, I.H., Joormann, J., 2010. Cognition and depression: current status and future directions. Annu. Rev. Clin. Psychol. 6, 285–312.
Gottesman, I.I., Gould, T.D., 2003. The endophenotype concept in psychiatry: etymology and strategic intentions. Am. J. Psychiatry 160 (4), 636–645.
Gradin, V.B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Reid, I., Hall, J., Steele, J.D., 2011. Expected value and prediction error abnormalities in depression and schizophrenia. Brain 134 (6), 1751–1764.
Haber, S.N., 2014. The place of dopamine in the cortico-basal ganglia circuit. Neuroscience 282C, 248–257.
Haber, S.N., Behrens, T.E., 2014. The neural network underlying incentive-based learning: implications for interpreting circuit disruptions in psychiatric disorders. Neuron 83 (5), 1019–1039.
Haber, S.N., Knutson, B., 2010. The reward circuit: linking primate anatomy and human imaging. Neuropsychopharmacology 35 (1), 4–26.
Hajcak, G., Moser, J.S., Yeung, N., Simons, R.F., 2005. On the ERN and the significance of errors. Psychophysiology 42 (2), 151–160.
Hamani, C., Amorim, B.O., Wheeler, A.L., Diwan, M., Driesslein, K., Covolan, L., Butson, C.R., Nobrega, J.N., 2014. Deep brain stimulation in rats: different targets induce similar antidepressant-like effects but influence different circuits. Neurobiol. Dis. 71, 205–214.
Hamani, C., Diwan, M., Macedo, C.E., Brandão, M.L., Shumake, J., Gonzalez-Lima, F., Raymond, R., Lozano, A.M., Fletcher, P.J., Nobrega, J.N., 2010. Antidepressant-like effects of medial prefrontal cortex deep brain stimulation in rats. Biol. Psychiatry 67 (2), 117–124.
Hamani, C., Machado, D.C., Hipólide, D.C., Dubiela, F.P., Suchecki, D., Macedo, C.E., Tescarollo, F., Martins, U., Covolan, L., Nobrega, J.N., 2012. Deep brain stimulation reverses anhedonic-like behavior in a chronic model of depression: role of serotonin and brain derived neurotrophic factor. Biol. Psychiatry 71 (1), 30–35.
Hammen, C., 2005. Stress and depression. Annu. Rev. Clin. Psychol. 1, 293–319.
Han, R., Takahashi, T., 2012. Psychophysics of time perception and valuation in temporal discounting of gain and loss. Physica A 391 (24), 6568–6576.
Hasler, G., 2012. Can the neuroeconomics revolution revolutionize psychiatry? Neurosci. Biobehav. Rev. 36 (1), 64–78.
Henn, F.A., Vollmayr, B., 2005. Stress models of depression: forming genetically vulnerable strains. Neurosci. Biobehav. Rev. 29 (4–5), 799–804.
Hill, M.N., Carrier, E.J., McLaughlin, R.J., Morrish, A.C., Meier, S.E., Hillard, C.J., Gorzalka, B.B., 2008. Regional alterations in the endocannabinoid system in an animal model of depression: effects of concurrent antidepressant treatment. J. Neurochem. 106 (6), 2322–2336.
Hill, M.N., Hillard, C.J., Bambico, F.R., Patel, S., Gorzalka, B.B., Gobbi, G., 2009. The therapeutic potential of the endocannabinoid system for the development of a novel class of antidepressants. Trends Pharmacol. Sci. 30 (9), 484–493.
Hirvonen, J., Hietala, J., Kajander, J., Markkula, J., Rasi-Hakala, H., Salminen, J.K., Någren, K., Aalto, S., Karlsson, H., 2011. Effects of antidepressant drug treatment and psychotherapy on striatal and thalamic dopamine D2/3 receptors in major depressive disorder studied with [11C]raclopride PET. J. Psychopharmacol. 25 (10), 1329–1336.
Holroyd, C.B., Coles, M.G., 2002. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol. Rev. 109 (4), 679–709.
Holtzheimer, P.E., Kelley, M.E., Gross, R.E., Filkowski, M.M., Garlow, S.J., Barrocas, A., Wint, D., Craighead, M.C., Kozarsky, J., Chismar, R., Moreines, J.L., Mewes, K., Posse, P.R., Gutman, D.A., Mayberg, H.S., 2012. Subcallosal cingulate deep brain stimulation for treatment-resistant unipolar and bipolar depression. Arch. Gen. Psychiatry 69 (2), 150–158.
Howlett, J.R., Paulus, M.P., 2013. Decision-making dysfunctions of counterfactuals in depression: who might I have been? Front Psychiatry 4, 143, http://dx.doi.org/10.3389/fpsyt.2013.00143 (eCollection 2013).
Hsiao, M.C., Lin, K.J., Liu, C.Y., Schatz, D.B., 2013. The interaction between dopamine transporter function, gender differences, and possible laterality in depression. Psychiatry Res. 211 (1), 72–77.
Huettel, S.A., 2012. Event-related fMRI in cognition. Neuroimage 62 (2), 1152–1156.
Huys, Q.J.M., Daw, N.D., Dayan, P., 2015. Depression: a decision theoretic analysis. Annu. Rev. Neurosci. 38 (Author's proof).
Huys, Q.J., Moutoussis, M., Williams, J., 2011. Are computational models of any use to psychiatry? Neural Netw. 24 (6), 544–551.
Huys, Q.J., Pizzagalli, D.A., Bogdan, R., Dayan, P., 2013. Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. Biol. Mood Anxiety Disord. 3 (1), 12.
Inoue, T., Kitaichi, Y., Masui, T., Nakagawa, S., Boku, S., Tanaka, T., Suzuki, K., Nakato, Y., Usui, R., Koyama, T., et al., 2010. Pramipexole for stage 2 treatment-resistant major depression: an open study. Prog. Neuropsychopharmacol. Biol. Psychiatry 34 (8), 1446–1449.
Jackson, F., Nelson, B.D., Proudfit, G.H., 2015. In an uncertain world, errors are more aversive: evidence from the error-related negativity. Emotion 15 (1), 12–16.
Jocham, G., Klein, T.A., Ullsperger, M., 2011. Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. J. Neurosci. 31 (5), 1606–1613.
Juhasz, G., Chase, D., Pegg, E., Downey, D., Toth, Z.G., Stones, K., Platt, H., Mekli, K., Payton, A., Elliott, R., Anderson, I.M., Deakin, J.F., 2009. CNR1 gene is associated with high neuroticism and low agreeableness and interacts with recent negative life events to predict current depressive symptoms. Neuropsychopharmacology 34 (8), 2019–2027.
Kahneman, D., 2011. Thinking, Fast and Slow. Farrar, Straus and Giroux, New York.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47 (2), 263–291.
Kanter, J.W., Manos, R.C., Bowe, W.M., Baruch, D.E., Busch, A.M., Rusch, L.C., 2010. What is behavioral activation? A review of the empirical literature. Clin. Psychol. Rev. 30 (6), 608–620.
Kapur, S., Mann, J.J., 1992. Role of the dopaminergic system in depression. Biol. Psychiatry 32 (1), 1–17.
Karhunen, T., Tilgmann, C., Ulmanen, I., Panula, P., 1995. Catechol-O-methyltransferase (COMT) in rat brain: immunoelectron microscopic study

with an antiserum against rat recombinant COMT protein. Neurosci. Lett. 187 (1), 57–60.
Kato, M., Chang, C.M., 2013. Augmentation treatments with second-generation antipsychotics to antidepressants in treatment-resistant depression. CNS Drugs 27 (Suppl. 1), S11–S19.
Kennedy, S.E., Koeppe, R.A., Young, E.A., Zubieta, J.K., 2006. Dysregulation of endogenous opioid emotion regulation circuitry in major depression in women. Arch. Gen. Psychiatry 63 (11), 1199–1208.
Kennedy, S.H., Giacobbe, P., Rizvi, S.J., Placenza, F.M., Nishikawa, Y., Mayberg, H.S., Lozano, A.M., 2011. Deep brain stimulation for treatment-resistant depression: follow-up after 3 to 6 years. Am. J. Psychiatry 168 (5), 502–510.
Kerestes, R., Davey, C.G., Stephanou, K., Whittle, S., Harrison, B.J., 2014. Functional brain imaging studies of youth depression: a systematic review. Neuroimage Clin. 4, 209–231.
Knutson, B., Bhanji, J.P., Cooney, R.E., Atlas, L.Y., Gotlib, I.H., 2008. Neural responses to monetary incentives in major depression. Biol. Psychiatry 63 (7), 686–692.
Knutson, B., Heinz, A., 2015. Probing psychiatric symptoms with the monetary incentive delay task. Biol. Psychiatry 77 (5), 418–420.
Kocabas, N.A., 2012. Catechol-O-methyltransferase (COMT) pharmacogenetics in the treatment response phenotypes of major depressive disorder (MDD). CNS Neurol. Disord. Drug Targets 11 (3), 264–272.
Koob, G., Kreek, M.J., 2007. Stress, dysregulation of drug reward pathways, and the transition to drug dependence. Am. J. Psychiatry 164 (8), 1149–1159.
Kruk-Slomka, M., Michalak, A., Biala, G., 2015. Antidepressant-like effects of the cannabinoid receptor ligands in the forced swimming test in mice: mechanism of action and possible interactions with cholinergic system. Behav. Brain Res. 284C, 24–36.
Kumar, P., Waiter, G., Ahearn, T., Milders, M., Reid, I., Steele, J.D., 2008. Abnormal temporal difference reward-learning signals in major depression. Brain 131 (Pt 8), 2084–2093.
Laasonen-Balk, T., Kuikka, J., Viinamäki, H., Husso-Saastamoinen, M., Lehtonen, J., Tiihonen, J., 1999. Striatal dopamine transporter density in major depression. Psychopharmacology (Berlin) 144 (3), 282–285.
Lam, R.W., Stewart, J.N., 1996. The validity of atypical depression in DSM-IV. Compr. Psychiatry 37 (6), 375–383.
Lammers, C.H., Diaz, J., Schwartz, J.C., Sokoloff, P., 2000. Selective increase of dopamine D3 receptor gene expression as a common effect of chronic antidepressant treatments. Mol. Psychiatry 5 (4), 378–388.
Lammel, S., Lim, B.K., Ran, C., Huang, K.W., Betley, M.J., Tye, K.M., Deisseroth, K., Malenka, R.C., 2012. Input-specific control of reward and aversion in the ventral tegmental area. Nature 491 (7423), 212–217.
Landauer, T.K., 1969. Reinforcement as consolidation. Psychol. Rev. 76 (1), 82–96.
Larsen, T., O'Doherty, J.P., 2014. Uncovering the spatio-temporal dynamics of value-based decision-making in the human brain: a combined fMRI-EEG study. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 369 (1655), 20130473.
Lecca, S., Meye, F.J., Mameli, M., 2014. The lateral habenula in addiction and depression: an anatomical, synaptic and behavioral overview. Eur. J. Neurosci. 39 (7), 1170–1178.
Lee, D., 2013. Decision making: from neuroscience to psychiatry. Neuron 78, 233–248.
Lee, D., Seo, H., Jung, M.W., 2012a. Neural basis of reinforcement learning and decision making. Annu. Rev. Neurosci. 35, 287–308.
Lee, R.S., Hermens, D.F., Porter, M.A., Redoblado-Hodge, M.A., 2012b. A meta-analysis of cognitive deficits in first-episode Major Depressive Disorder. J. Affect. Disord. 140 (2), 113–124.
Leentjens, A.F., 2011. The role of dopamine agonists in the treatment of depression in patients with Parkinson's disease: a systematic review. Drugs 71 (3), 273–286.
Li, Y., Zhu, Z.R., Ou, B.C., Wang, Y.Q., Tan, Z.B., Deng, C.M., Gao, Y.Y., Tang, M., So, J.H., Mu, Y.L., Zhang, L.Q., 2015. Dopamine D2/D3 but not dopamine D1 receptors are involved in the rapid antidepressant-like effects of ketamine in the forced swim test. Behav. Brain Res. 279, 100–105.
Liljeholm, M., O'Doherty, J.P., 2012. Contributions of the striatum to learning, motivation, 205 and performance: an associative account. Trends Cogn. Sci. 16 (9), 467–475.
Liu, Y., Edwards, R.H., 1997. The role of vesicular transport proteins in synaptic transmission and neural degeneration. Annu. Rev. Neurosci. 20, 125–156.
Lowry, C.A., Hale, M.W., Evans, A.K., Heerkens, J., Staub, D.R., Gasser, P.J., Shekhar, A., 2008. Serotonergic systems, anxiety, and affective disorder: focus on the dorsomedial part of the dorsal raphe nucleus. Ann. N. Y. Acad. Sci. 1148, 86–94.
Lozano, A.M., Giacobbe, P., Hamani, C., Rizvi, S.J., Kennedy, S.H., Kolivakis, T.T., Debonnel, G., Sadikot, A.F., Lam, R.W., Howard, A.K., Ilcewicz-Klimek, M., Honey, C.R., Mayberg, H.S., 2012. A multicenter pilot study of subcallosal cingulate area deep brain stimulation for treatment-resistant depression. J. Neurosurg. 116 (2), 315–322.
Lozano, A.M., Mayberg, H.S., Giacobbe, P., Hamani, C., Craddock, R.C., Kennedy, S.H., 2008. Subcallosal cingulate gyrus deep brain stimulation for treatment-resistant depression. Biol. Psychiatry 64 (6), 461–467.
Lupien, S.J., McEwen, B.S., Gunnar, M.R., Heim, C., 2009. Effects of stress throughout the lifespan on the brain, behaviour and cognition. Nat. Rev. Neurosci. 10 (6), 434–445.
Lutz, P.E., Kieffer, B.L., 2013. Opioid receptors: distinct roles in mood disorders. Trends Neurosci. 36 (3), 195–206.
Maddox, W.T., Chandrasekaran, B., Smayda, K., Yi, H.G., Koslov, S., Beevers, C.G., 2014. Elevated depressive symptoms enhance reflexive but not reflective auditory category learning. Cortex 58, 186–198.
Maia, T.V., Frank, M.J., 2011. From reinforcement learning models to psychiatric and neurological disorders. Nat. Neurosci. 14, 154–162.
Maier, S., Seligman, M., 1976. Learned helplessness: theory and evidence. J. Exp. Psychol.: Gen. 105 (1), 3–46.
Maier, S.F., Watkins, L.R., 2005. Stressor controllability and learned helplessness: the roles of the dorsal raphe nucleus, serotonin, and corticotropin-releasing factor. Neurosci. Biobehav. Rev. 29 (4–5), 829–841.
Maj, J., Rogóz, Z., 1999. Synergistic effect of pramipexole and sertraline in the forced swimming test. Pol. J. Pharmacol. 51 (6), 471–475.
Malone Jr., D.A., Dougherty, D.D., Rezai, A.R., Carpenter, L.L., Friehs, G.M., Eskandar, E.N., Rauch, S.L., Rasmussen, S.A., Machado, A.G., Kubu, C.S., Tyrka, A.R., Price, L.H., Stypulkowski, P.H., Giftakis, J.E., Rise, M.T., Malloy, P.F., Salloway, S.P., Greenberg, B.D., 2009. Deep brain stimulation of the ventral capsule/ventral striatum for treatment-resistant depression. Biol. Psychiatry 65 (4), 267–275.
Männistö, P.T., Kaakkola, S., 1999. Catechol-O-methyltransferase (COMT): biochemistry, molecular biology, pharmacology, and clinical efficacy of the new selective COMT inhibitors. Pharmacol. Rev. 51 (4), 593–628.
Markman, K.D., Miller, A.K., 2006. Depression, control, and counterfactual thinking: functional for whom? J. Soc. Clin. Psychol. 25 (2), 210–227.
Markman, K.D., Weary, G., 1996. The influence of chronic control concerns on counterfactual thought. Soc. Cogn. 14 (4), 292–316.
Markou, A., Salamone, J.D., Bussey, T.J., Mar, A.C., Brunner, D., Gilmour, G., Balsam, P., 2013. Measuring reinforcement learning and motivation constructs in experimental animals: relevance to the negative symptoms of schizophrenia. Neurosci. Biobehav. Rev. 37 (9 Pt B), 2149–2165.
Martinez-Aran, A., Vieta, E., 2015. Cognition as a target in schizophrenia, bipolar disorder and depression. Eur. Neuropsychopharmacol., http://dx.doi.org/10.1016/j.euroneuro.2015.01.007, pii:S0924-977X(15)00008-5.
Mayberg, H.S., Lozano, A.M., Voon, V., McNeely, H.E., Seminowicz, D., Hamani, C., Schwalb, J.M., Kennedy, S.H., 2005. Deep brain stimulation for treatment-resistant depression. Neuron 45 (5), 651–660.
Mazzucchelli, T., Kane, R., Rees, C., 2009. Behavioral Activation treatments for depression in adults: a meta-analysis and review. Clin. Psychol. Sci. Pract. 16 (4), 383–411.
McDannald, M.A., Jones, J.L., Takahashi, Y.K., Schoenbaum, G., 2014. Learning theory: a driving force in understanding orbitofrontal function. Neurobiol. Learn. Mem. 108, 22–27.
McDermott, L.M., Ebmeier, K.P., 2009. A meta-analysis of depression severity and cognitive function. J. Affect. Disord. 119 (1–3), 1–8.
McEwen, B.S., Gianaros, P.J., 2010. Central role of the brain in stress and adaptation: links to socioeconomic status, health, and disease. Ann. N. Y. Acad. Sci. 1186, 190–222.
McEwen, B.S., Morrison, J.H., 2013. The brain on stress: vulnerability and plasticity of the prefrontal cortex over the life course. Neuron 79 (1), 16–29.
McIntyre, R.S., Filteau, M.J., Martin, L., Patry, S., Carvalho, A., Cha, D.S., Barakat, M., Miguelez, M., 2014. Treatment-resistant depression: definitions, review of the evidence, and algorithmic approach. J. Affect. Disord. 156, 1–7.
McKenzie, S., Frank, A.J., Kinsky, N.R., Porter, B., Rivière, P.D., Eichenbaum, H., 2014. Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. Neuron 83 (1), 202–215.
McNeely, H.E., Mayberg, H.S., Lozano, A.M., Kennedy, S.H., 2008. Neuropsychological impact of Cg25 deep brain stimulation for treatment-resistant depression: preliminary results over 12 months. J. Nerv. Ment. Dis. 196 (5), 405–410.
Meiser, J., Weindl, D., Hiller, K., 2013. Complexity of dopamine metabolism. Cell Commun. Signal. 11 (1), 34.
Meyer, J.H., Ginovart, N., Boovariwala, A., Sagrati, S., Hussey, D., Garcia, A., Young, T., Praschak-Rieder, N., Wilson, A.A., Houle, S., 2006. Elevated monoamine oxidase a levels in the brain: an explanation for the monoamine imbalance of major depression. Arch. Gen. Psychiatry 63 (11), 1209–1216.
Michelsen, K.A., Schmitz, C., Steinbusch, H.W., 2007. The dorsal raphe nucleus – from silver stainings to a role in depression. Brain Res. Rev. 55 (2), 329–342.
Mizoguchi, K., Yuzurihara, M., Ishige, A., Sasaki, H., Chui, D.H., Tabira, T., 2000. Chronic stress induces impairment of spatial working memory because of prefrontal dopaminergic dysfunction. J. Neurosci. 20 (4), 1568–1574.
Montague, P.R., Dolan, R.J., Friston, K.J., Dayan, P., 2012. Computational psychiatry. Trends Cogn. Sci. 16 (1), 72–80.
Moraga-Amaro, R., Gonzalez, H., Pacheco, R., Stehberg, J., 2014. Dopamine receptor D3 deficiency results in chronic depression and anxiety. Behav. Brain Res. 274, 186–193.
Morgan, J.K., Olino, T.M., McMakin, D.L., Ryan, N.D., Forbes, E.E., 2013. Neural response to reward as a predictor of increases in depressive symptoms in adolescence. Neurobiol. Dis. 52, 66–74.
Mueller, E.M., Pechtel, P., Cohen, A.L., Douglas, S.R., Pizzagalli, D.A., 2015. Potentiated processing of negative feedback in depression is attenuated by anhedonia. Depress. Anxiety, http://dx.doi.org/10.1002/da.22338
Must, A., Horvath, S., Nemeth, V.L., Janka, Z., 2013. The Iowa gambling task in depression – what have we learned about sub-optimal decision-making strategies? Front. Psychol. 4, 732.
Nakahara, H., 2014. Multiplexing signals in reinforcement learning with internal models and dopamine. Curr. Opin. Neurobiol. 25, 123–129.
Naranjo, C.A., Tremblay, L.K., Busto, U.E., 2001. The role of the brain reward system in depression. Prog. Neuropsychopharmacol. Biol. Psychiatry 25 (4), 781–823.
Nassar, M.R., Rumsey, K.M., Wilson, R.C., Parikh, K., Heasly, B., Gold, J.I., 2012. Rational regulation of learning dynamics by pupil-linked arousal systems. Nat. Neurosci. 15 (7), 1040–1046.

Neimat, J.S., Hamani, C., Giacobbe, P., Merskey, H., Kennedy, S.H., Mayberg, H.S., Lozano, A.M., 2008. Neural stimulation successfully treats depression in patients with prior ablative cingulotomy. Am. J. Psychiatry 165 (6), 687–693.
Nestler, E.J., Carlezon Jr., W.A., 2006. The mesolimbic dopamine reward circuit in depression. Biol. Psychiatry 59 (12), 1151–1159.
Nestler, E.J., Hyman, S.E., 2010. Animal models of neuropsychiatric disorders. Nat. Neurosci. 13 (10), 1161–1169.
Nieuwenhuis, S., Holroyd, C.B., Mol, N., Coles, M.G., 2004. Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. Neurosci. Biobehav. Rev. 28 (4), 441–448.
Niv, Y., 2009. Reinforcement learning in the brain. J. Math. Psychol. 53 (3), 139–154.
Niv, Y., Schoenbaum, G., 2008. Dialogues on prediction errors. Trends Cogn. Sci. 12 (7), 265–272.
Novick, A.M., Forster, G.L., Tejani-Butt, S.M., Watt, M.J., 2011. Adolescent social defeat alters markers of adult dopaminergic function. Brain Res. Bull. 86 (1–2), 123–128.
O'Doherty, J.P., 2011. Contributions of the ventromedial prefrontal cortex to goal-directed action selection. Ann. N. Y. Acad. Sci. 1239, 118–129.
O'Doherty, J.P., Hampton, A., Kim, H., 2007. Model-based fMRI and its application to reward learning and decision making. Ann. N. Y. Acad. Sci. 1104, 35–53.
Olvet, D.M., Hajcak, G., 2008. The error-related negativity (ERN) and psychopathology: toward an endophenotype. Clin. Psychol. Rev. 28 (8), 1343–1354.
Onaivi, E.S., Ishiguro, H., Gong, J.P., Patel, S., Meozzi, P.A., Myers, L., Perchuk, A., Mora, Z., Tagliaferro, P.A., Gardner, E., Brusco, A., Akinshola, B.E., Liu, Q.R., Chirwa, S.S., Hope, B., Lujilde, J., Inada, T., Iwasaki, S., Macharia, D., Teasenfitz, L., Arinami, T., Uhl, G.R., 2008. Functional expression of brain neuronal CB2 cannabinoid receptors are involved in the effects of drugs of abuse and in depression. Ann. N. Y. Acad. Sci. 1139, 434–449.
Opmeer, E.M., Kortekaas, R., Aleman, A., 2010. Depression and the role of genes involved in dopamine metabolism and signalling. Prog. Neurobiol. 92 (2), 112–133.
Otto, A.R., Raio, C.M., Chiang, A., Phelps, E.A., Daw, N.D., 2013. Working-memory capacity protects model-based learning from stress. Proc. Natl. Acad. Sci. U.S.A. 110 (52), 20941–20946.
Papassotiropoulos, A., de Quervaine, D.J.F., 2015. Failed drug discovery in psychiatry: time for human genome-guided solutions. Trends Cogn. Sci., http://dx.doi.org/10.1016/j.tics.2015.02.002 (Published online: February 26, 2015).
Papp, M., Klimek, V., Willner, P., 1994. Effects of imipramine on serotonergic and beta-adrenergic receptor binding in a realistic animal model of depression. Psychopharmacology (Berlin) 114, 309–314.
Parker, G., Roy, K., Mitchell, P., Wilhelm, K., Malhi, G., Hadzi-Pavlovic, D., 2002. Atypical depression: a reappraisal. Am. J. Psychiatry 159 (9), 1470–1479.
Paulus, M.P., Yu, A.J., 2012. Emotion and decision-making: affect-driven belief systems in anxiety and depression. Trends Cogn. Sci. 16 (9), 476–483.
Payzan-LeNestour, E., Dunne, S., Bossaerts, P., O'Doherty, J.P., 2013. The neural representation of unexpected uncertainty during value-based decision making. Neuron 79 (1), 191–201.
Pearson-Fuhrhop, K.M., Dunn, E.C., Mortero, S., Devan, W.J., Falcone, G.J., et al., 2014. Dopamine genetic risk score predicts depressive symptoms in healthy adults and adults with depression. PLOS ONE 9 (5), e93772.
Perreault, M.L., Hasbi, A., O'Dowd, B.F., George, S.R., 2014. Heteromeric dopamine receptor signaling complexes: emerging neurobiology and disease relevance. Neuropsychopharmacology 39 (1), 156–168.
Pizzagalli, D.A., 2014. Depression, stress, and anhedonia: toward a synthesis and integrated model. Annu. Rev. Clin. Psychol. 10, 393–423.
Pizzagalli, D.A., Holmes, A.J., Dillon, D.G., Goetz, E.L., Birk, J.L., Bogdan, R., Dougherty, D.D., Iosifescu, D.V., Rauch, S.L., Fava, M., 2009. Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder. Am. J. Psychiatry 166 (6), 702–710.
Polter, A.M., Kauer, J.A., 2014. Stress and VTA synapses: implications for addiction and depression. Eur. J. Neurosci. 39 (7), 1179–1188.
Preuschoff, K., 't Hart, B.M., Einhauser, W., 2011. Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. Front. Neurosci. 5, 115.
Price, J.L., Drevets, W.C., 2010. Neurocircuitry of mood disorders. Neuropsychopharmacology 35 (1), 192–216.
Proudfit, G.H., 2014. The reward positivity from basic research on reward to a biomarker for depression. Psychophysiology, http://dx.doi.org/10.1111/psyp.12370
Proulx, C.D., Hikosaka, O., Malinow, R., 2014. Reward processing by the lateral habenula in normal and depressive behaviors. Nat. Neurosci. 17 (9), 1146–1152.
Puigdemont, D., Pérez-Egea, R., Portella, M.J., Molet, J., de Diego-Adeliño, J., Gironell, A., Radua, J., Gómez-Anson, B., Rodríguez, R., Serra, M., de Quintana, C., Artigas, F., Álvarez, E., Pérez, V., 2012. Deep brain stimulation of the subcallosal cingulate gyrus: further evidence in treatment-resistant major depression. Int. J. Neuropsychopharmacol. 15 (1), 121–133.
Quelhas, A.C., Power, M.J., Juhos, C., Senos, J., 2008. Counterfactual thinking and functional differences in depression. Clin. Psychol. Psychother. 15 (5), 352–365.
Quitkin, F.M., McGarth, P.J., Stewart, J.W., Klein, D.F., 2003. A reappraisal of atypical depression. Am. J. Psychiatry 160, 798–800.
Rachlin, H., Raineri, A., Cross, D., 1991. Subjective probability and delay. J. Exp. Anal. Behav. 55 (2), 233–244.
Radenbach, C., Reiter, A.M., Engert, V., Sjoerds, Z., Villringer, A., Heinze, H.J., Deserno, L., Schlagenhauf, F., 2015. The interaction of acute and chronic stress impairs model-based behavioral control. Psychoneuroendocrinology 53, 268–280, http://dx.doi.org/10.1016/j.psyneuen.2014.12.017 (Epub 2015 Jan 7).
Ragland, J.D., Cools, R., Frank, M., Pizzagalli, D.A., Preston, A., Ranganath, C., Wagner, A.D., 2009. CNTRICS final task selection: long-term memory. Schizophr. Bull. 35 (1), 197–212.
Rangel, A., Clithero, J.A., 2013. The computation of stimulus values in simple choice. In: Glimcher, P., Fehr, E. (Eds.), Neuroeconomics: Decision Making and the Brain. , 2nd ed. Elsevier, San Diego, pp. 125–147.
Rasetti, R., Weinberger, D.R., 2011. Intermediate phenotypes in psychiatric disorders. Curr. Opin. Genet. Dev. 21 (3), 340–348.
Rea, E., Rummel, J., Schmidt, T.T., Hadar, R., Heinz, A., Mathé, A.A., Winter, C., 2014. Anti-anhedonic effect of deep brain stimulation of the prefrontal cortex and the dopaminergic reward system in a genetic rat model of depression: an intracranial self-stimulation paradigm study. Brain Stimul. 7 (1), 21–28.
Rice, M.E., Patel, J.C., Cragg, S.J., 2011. Dopamine release in the basal ganglia. Neuroscience 198, 112–137.
Richard, J.M., Castro, D.C., Difeliceantonio, A.G., Robinson, M.J., Berridge, K.C., 2013. Mapping brain circuits of reward and motivation: in the footsteps of Ann Kelley. Neurosci. Biobehav. Rev. 37 (9 Pt A), 1919–1931.
Richard-Devantoy, S., Berlim, M.T., Jollant, F., 2014. A meta-analysis of neuropsychological markers of vulnerability to suicidal behavior in mood disorders. Psychol. Med. 44 (8), 1663–1673.
Riesel, A., Weinberg, A., Endrass, T., Kathmann, N., Hajcak, G., 2012. Punishment has a lasting impact on error-related brain activity. Psychophysiology 49 (2), 239–247.
Riva-Posse, P., Choi, K.S., Holtzheimer, P.E., McIntyre, C.C., Gross, R.E., Chaturvedi, A., Crowell, A.L., Garlow, S.J., Rajendra, J.K., Mayberg, H.S., 2014. Defining critical white matter pathways mediating successful subcallosal cingulate deep brain stimulation for treatment-resistant depression. Biol. Psychiatry 76 (12), 963–969.
Rive, M.M., van Rooijen, G., Veltman, D.J., Phillips, M.L., Schene, A.H., Ruhé, H.G., 2013. Neural correlates of dysfunctional emotion regulation in major depressive disorder. A systematic review of neuroimaging studies. Neurosci. Biobehav. Rev. 37 (10 Pt 2), 2529–2553.
Robbins, T.W., Arnsten, A.F., 2009. The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. Annu. Rev. Neurosci. 32, 267–287.
Robinson, O.J., Overstreet, C., Charney, D.R., Vytal, K., Grillon, C., 2013. Stress increases aversive prediction error signal in the ventral striatum. Proc. Natl. Acad. Sci. U.S.A. 110 (10), 4129–4133.
Roeper, J., 2013. Dissecting the diversity of midbrain dopamine neurons. Trends Neurosci. 36 (6), 336–342.
Rogóż, Z., 2013. Combined treatment with atypical antipsychotics and antidepressants in treatment-resistant depression: preclinical and clinical efficacy. Pharmacol. Rep. 65 (6), 1535–1544.
Roiser, J.P., Elliott, R., Sahakian, B.J., 2012. Cognitive mechanisms of treatment in depression. Neuropsychopharmacology 37 (1), 117–136.
Roy, M., Shohamy, D., Daw, N., Jepma, M., Wimmer, G.E., Wager, T.D., 2014. Representation of aversive prediction errors in the human periaqueductal gray. Nat. Neurosci. 17 (11), 1607–1612.
Russo, S.J., Nestler, E.J., 2013. The brain reward circuitry in mood disorders. Nat. Rev. Neurosci. 14 (9), 609–625.
Sacher, J., Rekkas, P.V., Wilson, A.A., Houle, S., Romano, L., Hamidi, J., Rusjan, P., Fan, I., Stewart, D.E., Meyer, J.H., 2015. Relationship of monoamine oxidase-a distribution volume to postpartum depression and postpartum crying. Neuropsychopharmacology 40 (2), 429–435.
Santiago, R.M., Barbiero, J., Gradowski, R.W., Bochen, S., Lima, M.M., Da Cunha, C., Andreatini, R., Vital, M.A., 2014. Induction of depressive-like behavior by intranigral 6-OHDA is directly correlated with deficits in striatal dopamine and hippocampal serotonin. Behav. Brain Res. 259, 70–77.
Savica, R., Benarroch, E.E., 2014. Dopamine receptor signaling in the forebrain: recent insights and clinical implications. Neurology 83 (8), 758–767.
Schiller, D., Levy, I., Niv, Y., LeDoux, J.E., Phelps, E.A., 2008. From fear to safety and back: reversal of fear in the human brain. J. Neurosci. 28 (45), 11517–11525.
Schilström, B., Konradsson-Geuken, A., Ivanov, V., Gertow, J., Feltmann, K., Marcus, M.M., Jardemark, K., Svensson, T.H., 2011. Effects of S-citalopram, citalopram, and R-citalopram on the firing patterns of dopamine neurons in the ventral tegmental area, N-methyl-D-aspartate receptor-mediated transmission in the medial prefrontal cortex and cognitive function in the rat. Synapse 65 (5), 357–367.
Schlaepfer, T.E., Cohen, M.X., Frick, C., Kosel, M., Brodesser, D., Axmacher, N., Joe, A.Y., Kreft, M., Lenartz, D., Sturm, V., 2008. Deep brain stimulation to reward circuitry alleviates anhedonia in refractory major depression. Neuropsychopharmacology 33 (2), 368–377.
Schmuckermair, C., Gaburro, S., Sah, A., Landgraf, R., Sartori, S.B., Singewald, N., 2013. Behavioral and neurobiological effects of deep brain stimulation in a mouse model of high anxiety- and depression-like behavior. Neuropsychopharmacology 38 (7), 1234–1244.
Schott, B.H., Frischknecht, R., Debska-Vielhaber, G., John, N., Behnisch, G., Düzel, E., Gundelfinger, E.D., Seidenbecher, C.I., 2010. Membrane-bound catechol-O-methyl transferase in cortical neurons and glial cells is intracellularly oriented. Front. Psychiatry 1, 142, http://dx.doi.org/10.3389/fpsyt.2010.00142
Schuldiner, S., Shirvan, A., Linial, M., 1995. Vesicular neurotransmitter transporters: from bacteria to humans. Physiol. Rev. 75 (2), 369–392.
Schultz, W., 2006. Behavioral theories and the neurophysiology of reward. Annu. Rev. Psychol. 57, 87–115.
Schultz, W., 2010. Dopamine signals for reward value and risk: basic and recent data. Behav. Brain Funct. 6, 24.

Schwabe, L., Wolf, O.T., 2011. Stress-induced modulation of instrumental behavior: from goal-directed to habitual control of action. Behav. Brain Res. 219 (2), 321–328.
Schwabe, L., 2013. Stress and the engagement of multiple memory systems: integration of animal and human studies. Hippocampus 23 (11), 1035–1043.
Schwartz, K., Yadid, G., Weizman, A., Rehavi, M., 2003. Decreased limbic vesicular monoamine transporter 2 in a genetic rat model of depression. Brain Res. 965 (1–2), 174–179.
Seo, H., Cai, X., Donahue, C.H., Lee, D., 2014. Neural correlates of strategic reasoning during competitive games. Science 346 (6207), 340–343.
Sesack, S.R., Grace, A.A., 2010. Cortico-basal ganglia reward network: microcircuitry. Neuropsychopharmacology 35 (1), 27–47.
Sesack, S.R., Hawrylak, V.A., Guido, M.A., Levey, A.I., 1998. Cellular and subcellular localization of the dopamine transporter in rat cortex. Adv. Pharmacol. 42, 171–174.
Sharot, T., Guitart-Masip, M., Korn, C.W., Chowdhury, R., Dolan, R.J., 2012. How dopamine enhances an optimism bias in humans. Curr. Biol. 22 (16), 1477–1481.
Sharot, T., Shiner, T., Brown, A.C., Fan, J., Dolan, R.J., 2009. Dopamine enhances expectation of pleasure in humans. Curr. Biol. 19 (24), 2077–2080.
Sharp, C., Monterosso, J., Montague, P.R., 2012. Neuroeconomics: a bridge for translational research. Biol. Psychiatry 72 (2), 87–92.
Sherdell, L., Waugh, C.E., Gotlib, I.H., 2012. Anticipatory pleasure predicts motivation for reward in major depression. J. Abnorm. Psychol. 121 (1), 51–60.
Shiffrin, R.M., Schneider, W., 1977. Controlled and automatic human information processing. II: Perceptual learning, automatic attending, and a general theory. Psychol. Rev. 84, 127–190.
Shimp, C.P., 1976. Short-term memory in the pigeon: the previously reinforced response. J. Exp. Anal. Behav. 26 (3), 487–493.
Sinclair, D., Purves-Tyson, T.D., Allen, K.M., Weickert, C.S., 2014. Impacts of stress and sex hormones on dopamine neurotransmission in the adolescent brain. Psychopharmacology (Berlin) 231 (8), 1581–1599.
Smittenaar, P., Fitz Gerald, T.H., Romei, V., Wright, N.D., Dolan, R.J., 2013. Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. Neuron 80 (4), 914–919.
Snyder, H.R., 2013. Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. Psychol. Bull. 139 (1), 81–132.
Solé, B., Jiménez, E., Martinez-Aran, A., Vieta, E., 2015. Cognition as a target in major depression: new developments. Eur. Neuropsychopharmacol., http://dx.doi.org/10.1016/j.euroneuro.2014.12.004, pii:S0924-977X(14)00351-4.
Spasojevic, N., Stanisavljevic, D., Gavrilovic, L., Jovanovic, P., Cucakovic, A., Dronjak, S., 2012. Hippocampal asymmetry in expression of catecholamine synthesizing enzyme and transporters in socially isolated rats. Neuroendocrinol. Lett. 33 (6), 631–635.
Squire, L.R., Zola, S.M., 1996. Structure and function of declarative and nondeclarative memory systems. Proc. Natl. Acad. Sci. U.S.A. 93, 13515–13522.
Starcke, K., Brand, M., 2012. Decision making under stress: a selective review. Neurosci. Biobehav. Rev. 36 (4), 1228–1248.
Steele, J.D., Kumar, P., Ebmeier, K.P., 2007. Blunted response to feedback information in depressive illness. Brain 130 (Pt 9), 2367–2374.
Stephan, K.E., Mathys, C., 2014. Computational approaches to psychiatry. Curr. Opin. Neurobiol. 25, 85–92.
Stewart, J.W., 2007. Treating depression with atypical features. J. Clin. Psychiatry 68 (Suppl. 3), 25–29.
Stewart, J.W., McGrath, P.J., Quitkin, F.M., Klein, D.F., 2009. DSM-IV depression with atypical features: is it valid? Neuropsychopharmacology 34 (13), 2625–2632.
Sutton, R.S., Barto, A.G., 1998. Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA.
Takahashi, T., 2004. Cortisol levels and time-discounting of monetary gain in humans. Neuroreport 15 (13), 2145–2147.
Takahashi, T., 2009. Theoretical frameworks for neuroeconomics of intertemporal choice. J. Neurosci. Psychol. Econ. 2, 75–90.
Takahashi, T., 2011a. A neuroeconomic theory of rational addiction and nonlinear time-perception. Neuroendocrinol. Lett. 32 (3), 221–225.
Takahashi, T., 2011b. Neuroeconomics of suicide. Neuroendocrinol. Lett. 32 (4), 400–404.
Takahashi, T., Han, R., Nishinaka, H., Makino, T., Fukui, H., 2013. The q-exponential probability discounting of gain and loss. Appl. Math. 4, 876–881, http://dx.doi.org/10.4236/am.2013.46120
Takahashi, T., Ikeda, K., Fukushima, H., Hasegawa, T., 2007. Salivary alpha-amylase levels and hyperbolic discounting in male humans. Neuroendocrinol. Lett. 28 (1), 17–20.
Takahashi, T., Ikeda, K., Hasegawa, T., 2008a. Salivary alpha-amylase levels and temporal discounting for primary reward under a simulated life-threatening condition. Neuroendocrinol. Lett. 29 (4), 451–453.
Takahashi, T., Oono, H., Inoue, T., Boku, S., Kako, Y., Kitaichi, Y., Kusumi, I., Masui, T., Nakagawa, S., Suzuki, K., Tanaka, T., Koyama, T., Radford, M.H., 2008b. Depressive patients are more impulsive and inconsistent in intertemporal choice behavior for monetary gain and loss than healthy subjects – an analysis based on Tsallis' statistics. Neuroendocrinol. Lett. 29 (3), 351–358.
Takahashi, T., Shinada, M., Inukai, K., Tanida, S., Takahashi, C., Mifune, N., Takagishi, H., Horita, Y., Hashimoto, H., Yokota, K., Kameda, T., Yamagishi, T., 2010. Stress hormones predict hyperbolic time-discount rates six months later in adults. Neuroendocrinol. Lett. 31 (5), 616–621.
Thase, M.E., 2009. Atypical depression: useful concept, but it's time to revise the DSM-IV criteria. Neuropsychopharmacology 34 (13), 2633–2641.
Treadway, M.T., Bossaller, N.A., Shelton, R.C., Zald, D.H., 2012a. Effort-based decision-making in major depressive disorder: a translational model of motivational anhedonia. J. Abnorm. Psychol. 121 (3), 553–558.
Treadway, M.T., Buckholtz, J.W., Cowan, R.L., Woodward, N.D., Li, R., Ansari, M.S., Baldwin, R.M., Schwartzman, A.N., Kessler, R.M., Zald, D.H., 2012b. Dopaminergic mechanisms of individual differences in human effort-based decision-making. J. Neurosci. 32 (18), 6170–6176.
Treadway, M.T., Buckholtz, J.W., Schwartzman, A.N., Lambert, W.E., Zald, D.H., 2009. Worth the 'EEfRT'? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. PLoS ONE 4 (8), e6598be.
Treadway, M.T., Zald, D.H., 2011. Reconsidering anhedonia in depression: lessons from translational neuroscience. Neurosci. Biobehav. Rev. 35 (3), 537–555.
Treadway, M.T., Zald, D.H., 2013. Parsing anhedonia: translational models of reward-processing deficits in psychopathology. Curr. Dir. Psychol. Sci. 22 (3), 244–249.
Trew, J.L., 2011. Exploring the roles of approach and avoidance in depression: an integrative model. Clin. Psychol. Rev. 31 (7), 1156–1168.
Tritsch, N.X., Sabatini, B.L., 2012. Dopaminergic modulation of synaptic transmission in cortex and striatum. Neuron 76 (1), 33–50.
Turner, P., Kantaria, R., Young, A.H., 2014. A systematic review and meta-analysis of the evidence base for add-on treatment for patients with major depressive disorder who have not responded to antidepressant treatment: a European perspective. J. Psychopharmacol. 28 (2), 85–98.
Tye, K.M., 2014. Neural circuit reprogramming: a new paradigm for treating neuropsychiatric disease? Neuron 83 (6), 1259–1261.
Tye, K.M., Mirzabekov, J.J., Warden, M.R., Ferenczi, E.A., Tsai, H.C., Finkelstein, J., Kim, S.Y., Adhikari, A., Thompson, K.R., Andalman, A.S., Gunaydin, L.A., Witten, I.B., Deisseroth, K., 2013. Dopamine neurons modulate neural encoding and expression of depression-related behaviour. Nature 493 (7433), 537–541.
Ullsperger, M., Danielmeier, C., Jocham, G., 2014. Neurophysiology of performance monitoring and adaptive behavior. Physiol. Rev. 94 (1), 35–79.
Vaidyanathan, U., Nelson, L.D., Patrick, C.J., 2012. Clarifying domains of internalizing psychopathology using neurophysiology. Psychol. Med. 42 (3), 447–459.
Valenti, O., Gill, K.M., Grace, A.A., 2012. Different stressors produce excitation or inhibition of mesolimbic dopamine neuron activity: response alteration by stress pre-exposure. Eur. J. Neurosci. 35 (8), 1312–1321.
Valverde, O., Torrens, M., 2012. CB1 receptor-deficient mice as a model for depression. Neuroscience 204, 193–206.
van Heeringen, K., Bijttebier, S., Desmyter, S., Vervaet, M., Baeken, C., 2014. Is there a neuroanatomical basis of the vulnerability to suicidal behavior? A coordinate-based meta-analysis of structural and functional MRI studies. Front. Hum. Neurosci. 8, 824, http://dx.doi.org/10.3389/fnhum.2014.00824
Van't Veer, A., Carlezon Jr., W.A., 2013. Role of kappa-opioid receptors in stress and anxiety-related behavior. Psychopharmacology (Berlin) 229 (3), 435–452.
Veerakumar, A., Challis, C., Gupta, P., Da, J., Upadhyay, A., Beck, S.G., Berton, O., 2014. Antidepressant-like effects of cortical deep brain stimulation coincide with pro-neuroplastic adaptations of serotonin systems. Biol. Psychiatry 76 (3), 203–212.
Vinod, K.Y., Hungund, B.L., 2006. Role of the endocannabinoid system in depression and suicide. Trends Pharmacol. Sci. 27 (10), 539–545.
Vrieze, E., Pizzagalli, D.A., Demyttenaere, K., Hompes, T., Sienaert, P., de Boer, P., Schmidt, M., Claes, S., 2013. Reduced reward learning predicts outcome in major depressive disorder. Biol. Psychiatry 73 (7), 639–645.
Wahlstrom, D., White, T., Luciana, M., 2010. Neurobehavioral evidence for changes in dopamine system activity during adolescence. Neurosci. Biobehav. Rev. 34 (5), 631–648.
Walsh, M.M., Anderson, J.R., 2012. Learning from experience: event-related potential correlates of reward processing, neural adaptation and behavioral choice. Neurosci. Biobehav. Rev. 36, 1870–1884.
Walsh, J.J., Han, M.H., 2014. The heterogeneity of ventral tegmental area neurons: projection functions in a mood-related context. Neuroscience 282C, 101–108.
Wardle, M.C., Treadway, M.T., Mayo, L.M., Zald, D.H., de Wit, H., 2011. Amping up effort: effects of D-amphetamine on human effort-based decision-making. J. Neurosci. 31 (46), 16597–16602.
Watkins, E.R., Nolen-Hoeksema, S., 2014. A habit-goal framework of depressive rumination. J. Abnorm. Psychol. 123 (1), 24–34.
Wayment, H.K., Schenk, J.O., Sorg, B.A., 2001. Characterization of extracellular dopamine clearance in the medial prefrontal cortex: role of monoamine uptake and monoamine oxidase inhibition. J. Neurosci. 21 (1), 35–44.
Wiecki, T.V., Poland, J.S., Frank, M.J., 2015. Model-based cognitive neuroscience approaches to computational psychiatry: clustering and classification. Clin. Psychol. Sci., http://dx.doi.org/10.1177/2167702614565359 (Published online: March 10, 2015).
Williams, B.A., 1991. Choice as a function of local versus molar reinforcement contingencies. J. Exp. Anal. Behav. 56 (3), 455–473.
Willner, P., Klimek, V., Golembiowska, K., Muscat, R., 1991. Changes in mesolimbic dopamine may explain stress-induced anhedonia. Psychobiology 19, 79–84.
Wilson, R.C., Takahashi, Y.K., Schoenbaum, G., Niv, Y., 2014. Orbitofrontal cortex as a cognitive map of task space. Neuron 81 (2), 267–279.
Winter, C., von Rumohr, A., Mundt, A., Petrus, D., Klein, J., Lee, T., Morgenstern, R., Kupsch, A., Juckel, G., 2007. Lesions of dopaminergic neurons in the substantia nigra pars compacta and in the ventral tegmental area enhance depressive-like behavior in rats. Behav. Brain Res. 184 (2), 133–141.
Wright, B.M., Eiland III, E.H., Lorenz, R., 2013. Augmentation with atypical antipsychotics for depression: a review of evidence-based support from the medical literature. Pharmacotherapy 33 (3), 344–359.
Wunderlich, K., Smittenaar, P., Dolan, R.J., 2012. Dopamine enhances model-based over model-free choice behavior. Neuron 75 (3), 418–424.

Worbe, Y., Savulich, G., de Wit, S., Fernandez-Egea, E., Robbins, T.W., 2015. Tryptophan depletion promotes habitual over goal-directed control of appetitive responding in humans. Int. J. Neuropsychopharmacol., http://dx.doi.org/10.1093/ijnp/pyv013, pii:pyv013.

Yang, Y.K., Yeh, T.L., Yao, W.J., Lee, I.H., Chen, P.S., Chiu, N.T., Lu, R.B., 2008. Greater availability of dopamine transporters in patients with major depression – a dual-isotope SPECT study. Psychiatry Res. 162 (3), 230–235.

Yetnikoff, L., Lavezzi, H.N., Reichard, R.A., Zahm, D.S., 2014. An update on the connections of the ventral mesencephalic dopaminergic complex. Neuroscience 282C, 23–48.

Yu, A.J., Dayan, P., 2005. Uncertainty, neuromodulation, and attention. Neuron 46, 681–692.

Zajonc, R.B., 1982. Feeling and thinking: preferences need no inferences. Am. Psychol. 35, 151–175.

Zhang, J., Berridge, K.C., Tindell, A.J., Smith, K.S., Aldridge, J.W., 2009. A neural computational model of incentive salience. PLoS Comput. Biol. 5, e1000437.

Zhang, W.N., Chang, S.H., Guo, L.Y., Zhang, K.L., Wang, J., 2013. The neural correlates of reward-related processing in major depressive disorder: a meta-analysis of functional magnetic resonance imaging studies. J. Affect. Disord. 151 (2), 531–539.

Zubieta, J.K., Ketter, T.A., Bueller, J.A., Xu, Y., Kilbourn, M.R., Young, E.A., Koeppe, R.A., 2003. Regulation of human affective responses by anterior cingulate and limbic mu-opioid neurotransmission. Arch. Gen. Psychiatry 60 (11), 1145–1153.