Probabilistic Reinforcement Learning and Anhedonia

Brian D. Kangas, Andre Der-Avakian, and Diego A. Pizzagalli

Contents

Abstract Despite the prominence of anhedonic symptoms associated with diverse neuropsychiatric conditions, there are currently no approved therapeutics designed to attenuate the loss of responsivity to previously rewarding stimuli. However, the search for improved treatment options for anhedonia has been reinvigorated by a recent reconceptualization of the very construct of anhedonia, including within the Research Domain Criteria (RDoC) initiative. This chapter will focus on the RDoC Positive Valence Systems construct of reward learning generally and sub-construct of probabilistic reinforcement learning specifically. The general framework emphasizes objective measurement of a subject's responsivity to reward via reinforcement

B. D. Kangas (✉) and D. A. Pizzagalli
Harvard Medical School, McLean Hospital, Belmont, MA, USA
e-mail: bkangas@mclean.harvard.edu

A. Der-Avakian
University of California, San Diego, La Jolla, CA, USA

Curr Topics Behav Neurosci (2022) 58: 355–378
https://doi.org/10.1007/7854_2022_349
Published Online: 19 April 2022

learning under asymmetrical probabilistic contingencies as a means to quantify reward learning. Indeed, blunted reward responsiveness and reward learning are central features of anhedonia and have been repeatedly described in major depression. Moreover, these probabilistic reinforcement techniques can also reveal neurobiological mechanisms to aid development of innovative treatment approaches. In this chapter, we describe how investigating reward learning can improve our understanding of anhedonia via the four RDoC-recommended tasks that have been used to probe sensitivity to probabilistic reinforcement contingencies and how such task performance is disrupted in various neuropsychiatric conditions. We also illustrate how reverse translational approaches of probabilistic reinforcement assays in laboratory animals can inform understanding of pharmacological and physiological mechanisms. Next, we briefly summarize the neurobiology of probabilistic reinforcement learning, with a focus on the prefrontal cortex, anterior cingulate cortex, striatum, and amygdala. Finally, we discuss treatment implications and future directions in this burgeoning area.

Keywords Anhedonia · Animal models · Medications development · Probabilistic reinforcement schedules · Reverse translation · Reward learning

1 Introduction

1.1 Anhedonia: Definition and Statement of Problem

Anhedonia is traditionally defined as the loss of pleasure or lack of reactivity to previously rewarding stimuli. Although often associated with major depressive disorder (MDD; American Psychiatric Association 2013), its transdiagnostic relevance has emerged across neuropsychiatric conditions, including schizophrenia (Moran et al. 2022), bipolar disorder (Whitton and Pizzagalli 2022), post-traumatic stress disorder (Vinograd et al. 2022), anxiety disorder (Taylor et al. 2022), substance use disorders (Koob 2022; Gilbert and Stone 2022), eating disorders (Murray et al. 2022), neurodevelopmental disorders (Dichter and Rodriguez-Romaguera 2022), and neurodegenerative disorders (Turner and Husain 2022). Unfortunately, there are no approved treatments for anhedonia and first-line antidepressants such as selective serotonin reuptake inhibitors (SSRI) are typically ineffective at increasing hedonic tone in MDD (Calabrese et al. 2014). Therefore, a critical need for effective therapeutics to treat anhedonic conditions has inspired coordinated bi-directional research efforts between clinical investigations and animal models designed to optimize assays of relevant phenotypes (Der-Avakian et al. 2016; Silverman et al. 2020).

1.2 Using Probabilistic Contingencies to Examine Anhedonia

The search for improved treatment options for anhedonic individuals has been catalyzed by an important reconceptualization of the very construct of anhedonia in the latest revision (National Institute of Mental Health 2016) of the Research Domain Criteria (RDoC; Insel et al. 2010). This chapter will focus on the Positive Valence Systems construct of *reward learning* generally and sub-construct of *probabilistic reinforcement learning* specifically. Blunted reward responsiveness is a hallmark feature of anhedonia and examining a subject's responsivity to reward via reinforcement learning under asymmetrical probabilistic contingencies yields an objective probe to quantify reward learning. Indeed, a recent meta-analysis found that blunted reward bias was the metric most consistently associated with MDD (Halahakoon et al. 2020). In turn, these techniques can reveal neurobiological mechanisms and inform novel approaches to treat MDD and other neuropsychiatric conditions prominently characterized by anhedonic phenotypes and reductions in reward learning. In this review, we first explain how investigating reward learning can improve our understanding of anhedonia. To this end, we describe the four recommended tasks chosen for the reward learning subdomain of the Positive Valence Systems in the latest revision of the RDoC (NIMH 2016). These paradigms are used to probe sensitivity to probabilistic contingencies and their influence on choice behavior, and how such behavior is disrupted in various neuropsychiatric conditions. We then discuss promising examples of reverse translation of probabilistic assays in laboratory animals and the mechanistic understanding they have uncovered. Next, we summarize the neurobiology of probabilistic reinforcement learning, with a focus on the prefrontal cortex (PFC), anterior cingulate cortex (ACC), striatum, and amygdala. We end by discussing treatment implications and future directions in this burgeoning area.

2 How Probabilistic Contingencies Inform the Study of Anhedonia and Its Symptoms

This section highlights findings from the four RDoC-recommended behavioral tasks that have been designed to probe the reward learning subdomain across clinical populations. It should be noted that these empirical efforts are a significant departure from traditional clinical assessments and diagnostic tools that primarily rely on self-report questionnaires (Wang et al. 2022). Importantly, although the tactics vary among the tasks highlighted below, the approaches share a common strategy that emphasizes probabilistic reinforcement contingencies as an objective means to quantify responsivity to reward and participants' ability to learn from consequences, as well as investigate these processes across neuropsychiatric disorders.

2.1 Probabilistic Reward Task

The Probabilistic Reward Task (PRT) developed by Pizzagalli et al. (2005; modified after Tripp and Alsop 1999; see also Henriques et al. 1994) is a laboratory procedure designed to provide a quantitative measure of reward learning (i.e., ability to modulate behavior as a function of reinforcement history). The PRT uses probabilistic discrimination methodology to quantify responsiveness to changes in reinforcer frequency. In the prototypical computerized task, human participants are instructed to discriminate between two briefly presented mouths that vary minimally in length on a cartoon face. Unbeknownst to the participants, probabilistic contingencies are arranged so that correct responses on one alternative are rewarded 3 times more often (e.g., long line: rich alternative) than correct responses on the other alternative (e.g., short line: lean alternative). As predicted by signal detection theory (Luc et al. 2021; McCarthy and Davison 1979), healthy control participants consistently develop a response bias in favor of the rich alternative and do so without disruption in overall task discriminability (i.e., performance accuracy; Pizzagalli et al. 2005, 2008b).

During the last 17 years, the PRT has been widely used across laboratories and is one of the most common probabilistic reinforcement learning tasks used to study clinical populations (>85 empirical publications). Among others, selected studies have shown that response bias toward the more frequently rewarded stimulus: (a) is inversely related to current anhedonic symptoms in unselected adults (e.g., Pizzagalli et al. 2005), relatives of patients with MDD (Liu et al. 2016), and in a transdiagnostic sample with depression and anxiety disorders (Reilly et al. 2020); (b) predicts self-reported anhedonic symptoms 38 days later (Pizzagalli et al. 2005) and a diagnosis of MDD 8 weeks later (Vrieze et al. 2013); (c) is blunted in individuals with increased depressive symptoms (Pizzagalli et al. 2005), current MDD (e.g., Pizzagalli et al. 2008c; Vrieze et al. 2013, Liu et al. 2011; but see Reilly et al. 2020), and past MDD (e.g., Liu et al. 2011, 2016; Pechtel et al. 2013; but see Audrain-McGovern et al. 2014), particularly those with elevated anhedonic symptoms (Vrieze et al. 2013) or melancholic depression (Fletcher et al. 2015); (d) is blunted in youth reporting anhedonia across various DSM diagnoses (Morris et al. 2015) and individuals with PTSD and elevated anhedonia (Eskelund et al. 2018) but not schizophrenia (e.g., Barch et al. 2017); (e) is linked to functional, electrophysiological, and molecular markers within mesolimbic pathways (e.g., Bogdan et al. 2011; Santesso et al. 2009; Kaiser et al. 2018); (f) is potentiated by pharmacological challenges hypothesized to increase dopaminergic signaling (e.g., nicotine, amphetamine, k-opioid receptor antagonism) in both humans and rats (e.g., Barr et al. 2008; Der-Avakian et al. 2013; Kangas et al. 2020; Krystal et al. 2020; Lamontagne et al. 2018); (g) is reduced by pharmacological challenges hypothesized to decrease dopaminergic signaling in both humans and rats (e.g., Der-Avakian et al. 2013; Grob et al. 2012; Lamontagne et al. 2018; Pizzagalli et al. 2008a); and (h) is amenable to computational modeling that allows to parse reward sensitivity and learning rate (e.g., Huys et al. 2013).

2.2 *Probabilistic Stimulus Selection Task*

The Probabilistic Stimulus Selection Task (PSST) developed by Frank et al. (2004) is also a computerized task using visual discrimination methodology and probabilistic conditions. This laboratory protocol consists of two phases. First, in the *acquisition phase*, subjects are presented with three different stimulus pairs across trials that have varied asymmetric probabilistic contingencies arranged (A:B, 80%:20%; C:D, 70%:30%; E:F, 60%:40%). Following discrimination mastery, subjects are then exposed to a *transfer test phase* in which they are presented with the same stimuli, but in novel arrangements and without feedback, to enable examination of whether response biases that emerge are a function of *choosing* the more frequently rewarded (rich) stimulus or *avoiding* the less frequently rewarded (lean) stimulus. This task was originally designed to characterize reward learning via positive vs. negative feedback in patients with Parkinson's disease while either unmedicated or medicated with L-dopa. These initial studies verified the expected findings in reward responsiveness (i.e., the inability to learn from trial and error); however, this task was also able to reveal selectivity in the effects of positive vs. negative feedback. Specifically, impairment in learning under probabilistic contingencies was driven by insensitivity to positive feedback when unmedicated relative to their performance under medicated conditions and, also, sensitivity to negative feedback under unmedicated conditions that was greater than when medicated (Frank et al. 2004). These observations of functional segregation between responses to positive and negative outcomes, in turn, were examined further using computational models to mechanistically interrogate the so-called "Go" and "NoGo" dopaminergic signaling pathways, primarily in the basal ganglia which has well-known dopamine depletion in Parkinson's disease patients (Frank 2005).

The value of this experimental framework was extended in patients with schizophrenia (Waltz et al. 2007), a clinical population also known to have dopamine dysfunction in the basal ganglia and, often more critically, in the prefrontal cortex (Weinberger 1987; Weinberger and Berman 1988). These system deficits have been long associated with poor reinforcement learning rates, anhedonic phenotypes, and negative symptoms of schizophrenia (Kirkpatrick and Buchanan 1990). Pronounced deficits in prefrontal cortex function were indeed corroborated by an inability of most patients with schizophrenia to successfully learn to discriminate between the standard PSST stimuli (Hiragana characters) used in the studies with Parkinson's patients highlighted above. However, patients with schizophrenia were able to successfully engage with a task variant that used more familiar clip art images as stimuli (which could, however, introduce working memory requirements that could make result interpretations challenging). In addition, the modified task confirmed reduced reward learning in patients with schizophrenia. Specifically, reduced learning from positive, but not negative, outcomes were observed and have since been replicated (Dowd et al. 2016; see also Strauss et al. 2014 for a review on the role of reward learning in the motivational impairment of schizophrenic disorders).

The PSST has also been used to examine reward responsiveness in MDD participants. For example, Admon et al. (2017) conducted a randomized controlled trial in unmedicated depressed patients and healthy control participants receiving either placebo or a single, low dose of the D_2/D_3 receptor antagonist amisulpride (thought to increase dopamine signaling through presynaptic autoreceptor blockade). As hypothesized, depressed patients showed a reduced probability of selecting previously rewarding stimuli. However, despite the ability of amisulpride to potentiate corticostriatal functional connectivity (examined with fMRI) in response to monetary rewards in the same study, drug treatment did not modulate behavioral performance. Similarly reduced reward learning in PSST performance was also observed in a sample of women with remitted MDD and a history of childhood sexual abuse (Pechtel and Pizzagalli 2013), which included concurrent electrophysiological measurement; source-localized electroencephalographic (EEG) activity revealed blunted differentiation between correct and incorrect responses (feedback-related negativity and error-related negativity) and increased activation in the subgenual anterior cingulate cortex in the clinical sample. Cavanagh et al. (2019) extended our understanding of PSST performance in MDD participants by associating selective features of EEG responses to probabilistic reward and punishment by examining positive prediction errors (when the outcome is better than expected) and negative prediction errors (when the outcome is worse than expected). By teasing apart depressive and anxious dimensional aspects of MDD, the authors were able to document elevated anxiety as reliably associated with avoidance learning due to a tighter coupling of negative prediction error signaling (i.e., the mismatch between reward expectancy and actual reward omission) with punishment-specific EEG features (i.e., ERPs related to punishment stimuli and associated theta-band dynamics). Conversely, depressive symptoms were reliably associated with smaller reward-related EEG signature (i.e., smaller reward-specific ERPs and associated delta-band dynamics). These dissociations between diverse dimensions of MDD support further an RDoC view of multifaceted neuropsychiatric disorders.

More recently, Brown et al. (2021) examined in participants with MDD the ability of cognitive behavioral therapy (CBT) to improve probabilistic reward (and loss) learning during fMRI imaging of prediction error and value signaling in the striatum. Among the participants with MDD, expected reductions in reward learning rates, associations between prediction error and expected value in ventral striatum, and anhedonia were observed relative to healthy controls. Following CBT, participants with MDD exhibited expected reductions of anhedonic and negative affect symptoms and, as well, significantly higher reward learning rates and ventral striatum signaling to prediction error and expected value. Moreover, a correlation was observed between reported symptom change and task-related behavioral and neural responses, thus demonstrating that this nonpharmacological treatment strategy can have desirable effects on reinforcement learning processes. Importantly, however, inconsistent findings have been observed when examining behavioral and neural responses to probabilistic reinforcement conditions in participants with MDD. For example, Rutledge et al. (2017) found no evidence of reward learning reduction

using fMRI, computational modeling, and smartphone-based metrics between depressed participants and healthy controls in monetary earnings, choice accuracy, and reaction times, nor were differences observed in reward prediction errors in BOLD responses in the reward-relevant regions of interest in the ventral striatum.

Interestingly, healthy subjects also display blunted reward responsivity in the PSST following acute exposure to stressful conditions, which are known etiological factors in MDD (Hammen 2005). For example, Berghorst et al. (2013) examined the effects of threat-of-shock experimental protocols in healthy female subjects on self-report measures, cortisol, and PSST performance. Although not all subjects had expected elevations in self-reported stress or elevations in cortisol, those who were sensitive to the laboratory stressor were characterized by blunted learning from reward, but not punishment, as assayed by the PSST.

2.3 Probabilistic Pavlovian Conditioning Task

Examination of Pavlovian conditioning can provide insight into additional aspects of fundamental adaptive behavior that, unlike operant conditioning, allows for assessments of reward learning during passive stimulus-response exposure rather than through volitional behavioral responses determined by programmed response-reinforcement contingencies. Although there are numerous ways to arrange classical conditioning paradigms (Bouton 2016; Pavlov 1927), in keeping with the theme of this chapter, O'Doherty et al. (2004) promulgated a probabilistic variant of a Pavlovian conditioning task which has been subsequently refined for use in clinical studies of anhedonic phenotypes. The task was initially developed to serve as a control condition for a probabilistic operant task designed to examine the extent to which the ventral and dorsal striatum contributes to instrumental conditioning. In the operant task, subjects are exposed to two trial types: either reward trials or neutral trials. During reward trial types, one of two stimuli is presented that was either associated with a relatively high (60%) or a relatively low (30%) probability of obtaining a palatable juice reward. During neutral trial types, subjects are presented with two different stimuli that are also associated with either a relatively high (60%) or a relatively low (30%) probability of obtaining a neutral tasteless solution. In the probabilistic Pavlovian conditioning task, subjects are exposed to the same conditions, but in a passive manner with the computer making the selection that exposed the subject to what would become conditioned stimuli immediately preceding either palatable or neutral stimuli. Because the ventral striatum has been long associated with reward learning and motivation (Cardinal et al. 2002), whereas the dorsal striatum is implicated in learning stimulus-response associations (Packard and Knowlton 2002), the active (operant) and passive (Pavlovian) tasks were conducted under fMRI conditions to examine reward learning, during both variants of conditioning, in the striatum. And, indeed, behavioral and neuroimaging outcomes largely supported these dissociable roles.

The general approach of including assessments of Pavlovian mechanisms in the behavioral and neural study of anhedonia was subsequently advanced by Kumar et al. (2008). Dopaminergic function has been long known to encode highly specific and brief phasic reward learning signals to unconditioned reinforcers and, as well, track behavioral measures of classical conditioning until the conditioned response produces dopamine release following the conditioned stimulus alone (Montague et al. 1996; Schultz 2002; Schultz and Dickinson 2000; McClure et al. 2003; Tobler et al. 2006). These mechanisms have been repeatedly documented to be blunted in MDD populations (Gershon et al. 2007; Gradin et al. 2011). Therefore, dysfunction in phasic reward learning signals was interrogated by Kumar et al. (2008) in medicated but SSRI treatment-resistant MDD patients and in healthy control subjects following acute treatment with the antidepressant citalopram. Computer-based photographic stimuli served as conditioned stimuli (A and B), which were presented prior to small volume water deliveries in fluid-deprived subjects. Probabilistic schedules associated with the conditioned stimuli and water delivery were systematically varied across five 20-trial blocks (e.g., A:B, 80%:0%; A:B, 50%:20%; A:B, 0%:90%; A:B, 20%:20%; A:B, 80%: 0%) to allow for repeated measures of conditioning and re-conditioning of differing response strength during fMRI recording. Findings showed that patients with MDD had expected blunting in reward learning signals in the ventral striatum, rostral and dorsal anterior cingulate, retrosplenial cortex, midbrain and hippocampus, with a magnitude that correlated with anhedonic severity. In addition, they observed that acute administration of citalopram in healthy control subjects blunted reward learning and its associated neurophysiological activity, which is consistent with evidence that typical antidepressants initially suppress dopamine function before enhancing it following chronic treatment as illustrated by their well-known delayed onset of action (Taylor et al. 2006).

This probabilistic Pavlovian conditioning task was subsequently used in conjunction with fMRI to examine putative dopaminergic function associated with reward learning in the ventral striatum and ventral tegmental area. Computational modeling techniques revealed that in medication-free patients with remitted recurrent depression and a high risk of recurrence, greater anhedonia was significantly associated with lower prediction error-related activation of the ventral tegmental area, whereas greater anhedonia in healthy controls was associated with higher prediction error-related activation of the ventral tegmental area (Geugies et al. 2019). These findings are consistent with assumptions regarding the dissociation of MDD and anhedonia and the latter's resistance to frontline antidepressant treatment even when it successfully reduced depressive symptoms and led to remission (Admon and Pizzagalli 2015; Calabrese et al. 2014). In other studies that paired the probabilistic Pavlovian conditioning task with fMRI to assess reward value encoding and event-related connectivity, Rupprechter et al. (2021) observed in unmedicated participants with MDD both blunted striatal activation following presentation of reward and negative encoding of reward value in the hippocampus and rostral anterior cingulate cortex, thus, suggesting an impaired communication between these areas as a possible culprit in the subjective valuation of rewards in MDD. Finally, probabilistic Pavlovian conditioning tasks have also been modified to examine both appetitive and

aversive outcomes under fMRI conditions and using computational modeling. For example, in studies designed to investigate how the habenula encodes negative value of stimuli associated with punishment contingencies, healthy participants (Lawson et al. 2014) and unmedicated patients with current MDD (Lawson et al. 2017) were exposed to abstract computerized images that were followed by probabilistically high (75%) or low (25%) positive (e.g., win money), negative (e.g., lose money, painful electric shock), or 100% neutral outcomes. Findings showed that habenula activation increased in response to conditioned stimuli more strongly predictive of negative outcomes, especially electric shock; however, the opposite was observed in participants with MDD (i.e., habenula activation decreased in the presence of conditioned stimuli more strongly predictive of shock). Moreover, habenula volume was negatively correlated with self-reported anhedonic symptoms in participants with MDD, leading the authors to speculate that habenula dysfunction may contribute to a poorer ability to avoid aversive stimuli, thereby, exacerbating MDD symptomology.

2.4 Drifting Double Bandit Task

The Drifting Double Bandit task (also known as the Two-step task) was developed by Daw et al. (2011) and designed to examine another aspect of reward learning, namely, a subject's reliance on goal-directed behavior versus habit-based behavior (e.g., inflexible responding based on previously experienced contingencies). This task consists of two stages. In the first stage, the subject is presented with two visual stimuli (A and B). A fixed probability is programmed for the stimulus pair such that a response to stimulus A results in a second stimulus pair (C and D) 70% of the time or another stimulus pair (E and F) 30% of the time, whereas a response to stimulus B results in a second stimulus pair (C and D) 30% of the time or (E and F) 70% of the time. In this second stage, responses to C and D *or* E and F are rewarded with monetary outcomes that are programmed with variable probabilistic schedules that change slowly and independently throughout the session. This arrangement is designed to examine the extent to which subjects are relatively habit-based and make choices based on the fixed probabilities arranged during the first stage stimulus pair or relatively goal-directed and remain flexible in response allocation as the probabilities change during the second stage stimulus pairs. This task also lends itself well to computational modeling strategies that can be used to define a subject's response style to determine reward learning processes. Although the ability of this task to probe reward learning processes as they relate to anhedonic phenotypes in this subdomain is highly probable, there have yet to be any published reports using the Drifting Double Bandit expressly for this pursuit, despite it being a recommended task in the most recent RDoC revision (NIMH 2016).

3 Reverse Translation of Probabilistic Assays in Laboratory Animals

Given the correspondence between behavioral outcomes under probabilistic contingencies and anhedonic phenotypes across diverse clinical populations, there have been increasing efforts to reverse translate these tasks for use in laboratory animals. As reviewed above, although task performance in human participants has revealed critical information regarding neurophysiological mechanisms which, in turn, have allowed an ability to appraise novel behavioral and pharmacological treatment strategies, there is considerable value in the ability to conduct similar studies in animals while healthy and following conditions designed to produce anhedonic-like phenotypes. Functional similarities in task outcome are the primary objective; however, recent advances in apparatus technologies have also afforded the ability to maintain certain formalistic features of various computerized cognitive tasks. More generally, the expectation is that this coordinated bi-directional approach will help bridge the preclinical gap between therapeutic discovery and treatment (Der-Avakian and Pizzagalli 2018).

One prominent example of this approach has been the reverse translation of the PRT into rats and nonhuman primates. The first variant of this task designed for laboratory animals established a protocol using tone duration discriminations in rats which, after acquisition, were programmed with a 3:1 rich:lean probabilistic contingency modeled after the human task detailed above (Der-Avakian et al. 2013). Expected task outcomes were observed, including a reliable response bias toward the more richly rewarded stimulus alternative and a pharmacological blunting of the response bias following administration of low doses of pramipexole (thought to decrease dopaminergic signaling via presynaptic autoreceptor activation) as seen previously in humans (Pizzagalli et al. 2008a). Subsequent independent studies advanced this approach by documenting task sensitivity to chronic stress, with rats exposed to social defeat exhibiting a blunted response bias relative to non-stressed controls (Der-Avakian et al. 2017) and highlighted the role of dopamine and glucocorticoid systems in reward responsiveness (Lamontagne et al. 2018).

Subsequent efforts to reverse translate the PRT capitalized on recent advances in touchscreen technology (Kangas and Bergman 2017) to develop a task variant using visual line-length discriminations under probabilistic contingencies designed for rats (Kangas et al. 2020) and nonhuman primates (Wooldridge et al. 2021). In addition to enhanced formal similarity of the touchscreen-based animal task variant to the computerized human task, expected response biases were observed in both species that closely approximated values observed in humans. Subsequent drug studies using these reverse-translated PRT variants in laboratory animals have confirmed the ability of putative antidepressants and pro-hedonics, such as amphetamine, scopolamine, and ketamine, to dose-dependently enhance reward learning. These findings confirm and extend their therapeutic promise previously documented in clinical populations using traditional metrics (Jaffe et al. 2013; Kim et al. 2019; McIntyre et al. 2017). Most recently, studies in rats have confirmed the ability of the

touchscreen PRT to characterize enduring deficits in reward responsiveness during adulthood long after exposure to a rodent model of early-life adversity and simulated poverty (Kangas et al. 2022).

Reverse translation of the other probabilistic tasks highlighted in this chapter has either yet to be developed for use to examine anhedonic phenotypes or has yet to be subjected to extensive pharmacological and neurophysiological analysis in healthy and chronically stressed animals. Some tasks (e.g., the probabilistic Pavlovian conditioning task) should be relatively straightforward to adapt for laboratory animals with aims to study anhedonic phenotypes, whereas other tasks (e.g., the PSST) will likely require creative modifications given the well-documented difficulty in reliably establishing transfer of function in laboratory animals (Lionello-DeNolf 2009; Zentall et al. 2014). Nevertheless, as illustrated above, coordinated translational efforts studying clinical populations and animal subjects can yield complementary approaches and mutually beneficial advances from clinical observations and laboratory discoveries.

4 Neurobiological Mechanisms of Probabilistic Reinforcement Learning

Several studies in both humans and laboratory animals have implicated corticolimbic circuits, modulated primarily by dopamine, norepinephrine, and serotonin in probabilistic reinforcement learning. In this section, we will provide a brief overview of neurobiological mechanisms that underlie probabilistic reinforcement learning.

4.1 Prefrontal Cortex and Probabilistic Learning

Two areas of the prefrontal cortex (PFC) that are heavily implicated in decision-making processes associated with probabilistic reinforcement include the orbitofrontal cortex (OFC) and anterior cingulate cortex (ACC). The OFC encodes reward value and responds to reward expectancy (Gottfried et al. 2003; Schoenbaum and Roesch 2005). Thus, the OFC is sensitive to both the magnitude and probability of future rewards, and lesions or pharmacological impairment of this area generally results in an inability to select optimal outcomes in the face of uncertainty, which may occur when the probability of obtaining a reward is relatively low (Mobini et al. 2002; Rogers et al. 1999a). For example, electrophysiological activity in the OFC is correlated with reward valence and expectancy (Hikosaka and Watanabe 2000; Schoenbaum and Roesch 2005; Kennerley et al. 2011). Similarly, in nonhuman primates, the magnitude of a reward modulates activity of OFC neurons, which can be modulated by reward expectancy and history (Saez et al. 2017). Moreover, cerebral blood flow is increased in the OFC in humans making a choice between

small rewards with a relatively high outcome probability and large rewards with a relatively low outcome probability (Rogers et al. 1999b). Lesioning the OFC in rats has been shown to increase risky decision-making and preference for uncertain rewards, whereby animals become more likely to respond for rewards that are large but have a relatively low probability (Stopper et al. 2014). Evidence suggests that this change in choice preference can be partially, but not exclusively, explained by deficits in reward valuation that are accompanied by OFC lesions (Stalnaker et al. 2015). That is, the OFC is important for coding reward expectancy under probabilistic conditions.

The ACC has also been implicated in signaling reward expectancy. In particular, the ACC is thought to code reward prediction errors, whereby a mismatch occurs between expected and actual reward outcomes (Hyman et al. 2017). Evidence in humans and nonhuman primates also suggests that the ACC codes for reward valuation as well (Amiez et al. 2006; Kolling et al. 2016). As with the OFC, inactivation of the prelimibic cortex (PrL) in rats, which is thought to approximate human ACC area 32/25, also increased risky decision-making, but only when the probability of reward decreased over time (St Onge and Floresco 2010). Interestingly, inactivation of this region decreased risky decision-making when reward probability increased over time, suggesting that the ACC plays an important role in updating reward probabilities based on outcome to help guide future decision-making. Thus, like the OFC, reward expectancy signals in the ACC may contribute to the coding of rewards of a particular magnitude and probability of outcome.

It is unlikely, however, that two distinct PFC regions play functionally identical roles with regard to reward expectancy. Differences between these two PFC areas may emerge in the rate at which they track reward probability, and thus expectancy, over time. Soltani and Izquierdo (2019) recently suggested that while the ACC may be responsible for rapid updating of reward probabilities based on immediate computation of unexpected events, the OFC may provide slower, longer-term updates on changes in reward valuation and expectancy. Alternatively, Winstanley and Floresco (2016) have suggested that the OFC plays a role in risky decision-making when one of the options includes an aversive stimulus, thereby promoting the value of the appetitive option. On the other hand, the ACC may help guide choices of two or more uncertain rewards to ensure maximal possible outcomes. Given that these two areas maintain reciprocal connections, it is important to also consider that discrete functional processes specific to one area are likely communicated to the other area to help guide decision-making during probabilistic reinforcement learning.

Both norepinephrine and serotonin appear to play neuromodulatory roles in the PFC with regard to reward expectancy signaling. Norepinephrine is thought to regulate the balance between exploitation and exploratory behavior as animals navigate different actions with varying probabilities of reward outcomes (Aston-Jones and Cohen 2005). The firing rates of noradrenergic cells originating from the locus coeruleus change with alternating reward contingencies. When reward outcomes are uncertain, tonic firing of noradrenergic cells facilitates alternating behavior from current actions that may be suboptimal to new actions that may produce

more certain outcomes (i.e., exploration; Aston-Jones et al. 1999). On the other hand, when reward outcomes become more certain, phasic firing of noradrenergic cells promotes optimized task performance (i.e., exploitation; Aston-Jones and Cohen 2005). Additionally, noradrenergic signaling in the orbitofrontal cortex may facilitate the learning of current or prior associative states (Sadacca et al. 2017). That is, the ability to recognize and adapt to changes in reward probabilities and expectations may require an understanding of different task states whereby different actions yield different outcomes depending on the task state. Maximizing reward outcomes requires actions to be implemented that are appropriate for a given state. The OFC is believed to mediate learning of these task states and may promote rapid learning under conditions of changing and unexpected reward contingencies.

Serotonin originating from the midbrain dorsal raphe nucleus (DRN) is also involved in reward expectancy and probabilistic learning and may regulate the timescale of reward predictions (Miyazaki et al. 2020). Whereas midbrain dopamine activity encodes prediction error signals, serotonin is believed to modulate the degree to which these prediction error signals for uncertain outcomes are integrated into action. Given the dense reciprocal connections between the dorsal raphe nucleus and OFC, it is possible that this serotonergic modulation of prediction error signaling is at least partially mediated by the OFC.

4.2 Striatum and Probabilistic Learning

Both ventral and dorsal striatum, which form corticostriatal loops with the PFC areas described above, appear to be involved in probabilistic reinforcement learning. In humans, parts of the midbrain that send dopaminergic projections to the NAc respond to stimulus uncertainty, and activity of these dopaminergic cells correlates with reward probability (Dreher et al. 2006). Moreover, increasing reward probability is associated with increased blood flow in the striatum in humans (Abler et al. 2006). Consistent with these findings in humans, lesions of the NAc in rats promote risk-aversive behavior by biasing choices away from large rewards with a low probability of occurrence and toward small rewards with a high probability of occurrence, while discrimination of the reward value of different choices remains largely intact (Cardinal and Howes 2005). Interestingly, despite the role of the shell subregion of the NAc in processing the hedonic value of rewards, the suppression of risky behavior described above was specific to the core subregion of the NAc, as NAc shell lesions had no effect on choice behavior based on reward probability. In both nonhuman primates (Costa et al. 2016) and rodents (St Onge et al. 2012), lesions of the dorsal striatum also impaired learning during probabilistic, but not deterministic, reward schedules. Much of the role of the striatum in signaling reward expectancy has focused on the neurotransmitter dopamine. Midbrain dopamine neuronal activity encodes the mismatch between expected and actual reward error signals (Schultz et al. 1997). That is, during positive prediction errors, the firing of

midbrain dopamine neurons is increased, whereas during negative prediction errors, firing of midbrain dopamine neurons is reduced. Thus, either too much or too little dopamine signaling may disrupt prediction error processing, thereby impairing learning during activities with unexpected or probabilistic reward schedules.

4.3 Basolateral Amygdala (BLA) and Probabilistic Learning

Evidence suggests that the BLA represents expected reward valuation and learning from changes in the expected value of rewards (Stolyarova and Izquierdo 2017). As described above, the OFC is also involved in coding the valuation of rewards, and it may do so via reciprocal connections with the BLA. Inactivation of the BLA in rats results in a shift toward risky decision-making, although this effect may not just rely on the value of positive outcomes. For example, if a particular choice leads to negative or aversive events, the BLA is thought to bias choice away from the aversive event. Further evidence supports the role of the BLA in rapidly detecting unexpected changes (positive or negative) to reward outcomes (Wassum and Izquierdo 2015). This rapid signaling of changes to expected reward outcomes could be mediated via reciprocal connections with the ACC, OFC, and insula. For example, amygdala connectivity with these cortical regions shifts preference toward smaller, certain rewards compared to larger, uncertain rewards (Ghods-Sharifi et al. 2009).

4.4 Overlapping Neural Circuits Underlying Probabilistic Learning and Anhedonia

The brain regions and neurotransmitters described above that support probabilistic reinforcement learning are strongly implicated in the symptom of anhedonia and several psychiatric disorders characterized by anhedonia. Activation of the OFC, and in particular the medial OFC that signals the value of rewards, is suppressed in MDD, impairing reward-related processes that likely contribute to the symptom of anhedonia. In contrast, the lateral OFC, which is responsible for signaling non-reward or aversive outcomes, is overactive in depression (e.g., Rolls 2019). Thus, suppression and potentiation of OFC subregions responsible for computing the value of rewarding and aversive outcomes, respectively, are thought to bias an individual with depression away from pleasant experiences and toward negative states. Indeed, the acute administration of the rapid-acting antidepressant ketamine in patients with treatment-resistant depression suppressed lateral OFC activity, and this suppression correlated with the alleviation of anhedonia (Lally et al. 2015). Activity of the ACC, and functional connectivity with surrounding cortical and

limbic areas, is also suppressed in patients with depression (Pizzagalli and Roberts 2022). Given the role of the ACC in encoding differences between expected and actual reward outcomes (i.e., prediction errors), the value of chosen rewards, and the integration of prior behavioral actions and subsequent reward outcomes, suppression of this region would be expected to impair reward-guided behavior. The striatum is involved in many aspects of reward-related behavior that are also impaired in patients with anhedonia. Suppression of activity in this region can negatively impact reward valuation, anticipation/expectancy, and motivation, each of which would hinder reward-guided behavior and manifest as anhedonia.

In summary, the neural computations of probabilistic learning when engaged in choices about different rewards (or aversive events) involves a diverse set of cortical and subcortical structures that are tightly interconnected, each of which computes different variables related to reward probability. The connections from cortical areas involved in reward valuation and uncertainty to subcortical areas are also widely involved in the pathophysiology of depression and other psychiatric disorders characterized by anhedonia. Thus, the different reward-related deficits observed in patients with anhedonia are likely reflected by disruptions in one or several corticolimbic structures that normally process reward valuation and expectancy signals.

5 Conclusions and Future Directions

The ability to learn from reward and adjust behavior accordingly is fundamental to survival across the animal kingdom. Here, we reviewed and integrated convergent preclinical and clinical findings highlighting the centrality of abnormalities in probabilistic reinforcement learning across neuropsychiatric disorders. Several important conclusions can be extracted from this burgeoning area. First, psychiatric conditions reporting similar levels of anhedonia, such as MDD and schizophrenia, are characterized by divergent patterns of reward learning abnormalities. For example, whereas MDD has been linked to blunted reward learning in implicit reinforcement tasks (such as the PRT), schizophrenia has been linked to (surprisingly) preserved implicit reward learning but blunted explicit reward learning (Barch et al. 2017; for an extended discussion, see Moran et al. 2022). This dissociation points to partially non-overlapping neurobiological abnormalities in the manifestation of anhedonia (i.e., MDD: more striatal-based vs. schizophrenia: more PFC-based), which implies that different therapeutic strategies might be needed to address anhedonia in these conditions. Second, by focusing on objective behavioral metrics that can be precisely quantified across species (e.g., rodents, nonhuman primates, humans) using functionally identical tasks, the field has an unprecedented opportunity to accelerate translational discoveries toward the development of novel treatments for anhedonia. In this context, it is noteworthy that, in both rats (Kangas et al. 2022) and humans (Pechtel and Pizzagalli 2013), early adversity has been

linked to blunted reward learning abilities in adulthood. Owing to such parallel findings, promising (and safe) compounds with efficacy to restore reward learning abilities in preclinical models could be useful for anhedonic individuals with a history of early-life adversity. Along similar lines, recent neuroimaging and behavioral findings – which were inspired by robust preclinical data highlighting kappa opioid blockade as a promising target for anhedonia – indicate that a kappa opioid receptor antagonist increased reward-related activation in the nucleus accumbens, boosted reward learning, and reduced self-reported anhedonia in a transdiagnostic sample (Krystal et al. 2020; Pizzagalli et al. 2020). In light of this evidence of "target engagement," clinical studies evaluating kappa opioid receptor antagonists to reverse anhedonic phenotypes are warranted. Third, as recently demonstrated by Ang and colleagues, parsing the heterogeneity of MDD using objective measures of reward learning abilities might provide a means to guide treatment selection, and thus speed up recovery (Ang et al. 2020). Specifically, in that study, reward learning rates that more closely approximated those observed in healthy control participants predicted response to the atypical antidepressant bupropion after failing 8 weeks of treatment with an SSRI. Finally, although we highlighted several possible pharmacological targets, it is important to emphasize that other treatment modalities are currently under intense investigation to tackle anhedonic phenotypes, including psychological treatments inspired by the RDoC (Sandman and Craske 2022) and neurostimulation (Siddiqi et al. 2022). With respect to the latter strategy, it is interesting to note that, among healthy controls, reward learning (as assessed by the PRT) could be potentiated by high-frequency rapid TMS (Ahn et al. 2013) or intermittent theta burst stimulation (Duprat et al. 2016) over the left dorsolateral PFC. Such findings raise the possibility that psychiatric conditions characterized by reward learning dysfunction might benefit from similar types of neurostimulation.

In spite of significant progress in this area, there are important outstanding questions for future studies. First and foremost, although reward learning abnormalities have emerged across tasks in specific psychiatric disorders (e.g., MDD), the causal status of blunted reward learning in anhedonia needs to be directly evaluated. Specifically, do improvements in anhedonia drive reward learning or does the resolution of anhedonia require normalization of reward learning? Dense sampling (e.g., within the context of a randomized clinical trial) of both constructs would be needed to clarify the temporal (and putatively, causal) relationship between them (e.g., early improvements in reward learning in week 1 predicts reduction in self-reported anhedonia in week 2). Second, reward learning abnormalities have often emerged using tasks (e.g., PRT) that include only adjusting behavior as a function of rewards. Thus, in such studies, it is unclear whether the documented abnormalities are specific to reward or might reflect more global (non-specific) learning deficits.

Ultimately, and as reviewed in detail in other chapters within this volume, we believe that parsing anhedonia into subdomains that are biologically more homogenous, can be probed in similar ways across species, and are subserved by distinct neurobiological pathways will give us the best chance at developing more efficacious and much needed treatments for anhedonia and reward learning deficits.

Acknowledgments BDK was partially supported by grant R01 DA047575 from the National Institute on Drug Abuse. AD was partially supported by grants R01 MH121352 and UH3 MH109334 from the National Institute of Mental Health and R01 AA026560 from the National Institute on Alcohol Abuse and Alcoholism. DAP was partially supported by grants R01 MH095809, P50 MH119467, R01 MH108602, R37 MH068376, and UH3 MH109334 from the National Institute of Mental Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Health.

Disclosures Over the past 3 years, Dr. Pizzagalli has received consulting fees from Albright Stonebridge Group, Boehringer Ingelheim, Compass Pathways, Concert Pharmaceuticals, Engrail Therapeutics, Neumora Therapeutics (former BlackThorn Therapeutics), Neurocrine Biosciences, Neuroscience Software, Otsuka Pharmaceuticals, and Takeda Pharmaceuticals; honoraria from the Psychonomic Society (for editorial work), one honorarium from Alkermes, and research funding from National Institute of Mental Health, Dana Foundation, Brain and Behavior Research Foundation, Millennium Pharmaceuticals. In addition, he has received stock options from BlackThorn Therapeutics, Compass Pathways and Neuroscience Software. Dr. Pizzagalli has a financial interest in Neumora Therapeutics (former BlackThorn Therapeutics), which has licensed the copyright to the Probabilistic Reward Task through Harvard University. Dr. Pizzagalli's interests were reviewed and are managed by McLean Hospital and Massachusetts General Brigham in accordance with their conflict-of-interest policies. Dr. Der-Avakian holds equity ownership in PAASP US. No funding from these entities was used to support the current work, and all views expressed are solely those of the authors. Dr. Kangas reports no financial relationships with commercial interest.

References

Abler B, Walter H, Erk S, Kammerer H, Spitzer M (2006) Prediction error as a linear function of reward probability is coded in human nucleus accumbens. NeuroImage 31:790–795

Admon R, Pizzagalli DA (2015) Dysfunctional reward processing in depression. Curr Opin Psychol 4:114–118

Admon R, Kaiser RH, Dillon DG, Beltzer M, Goer F, Olson DP, Vitaliano G, Pizzagalli DA (2017) Dopaminergic enhancement of striatal response to reward in major depression. Am J Psychiatry 174:378–386

Ahn HM, Kim SE, Kim SH (2013) The effects of high-frequency rTMS over the left dorsolateral prefrontal cortex on reward responsiveness. Brain Stimul 6:310–314

American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders: DSM-V. American Psychiatric Association, Washington

Amiez C, Joseph JP, Procyk E (2006) Reward encoding in the monkey anterior cingulate cortex. Cereb Cortex 16:1040–1055

Ang YS, Kaiser R, Deckersbach T, Almeida J, Phillips ML, Chase HW, Webb CA, Parsey R, Fava M, McGrath P, Weissman M, Adams P, Deldin P, Oquendo MA, McInnis MG, Carmody T, Bruder G, Cooper CM, Chin Fatt CR, Trivedi MH, Pizzagalli DA (2020) Pretreatment reward sensitivity and frontostriatal resting-state functional connectivity are associated with response to bupropion after sertraline nonresponse. Biol Psychiatry 88:657–667

Aston-Jones G, Cohen JD (2005) An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu Rev Neurosci 28:403–450

Aston-Jones G, Rajkowski J, Cohen J (1999) Role of locus coeruleus in attention and behavioral flexibility. Biol Psychiatry 46:1309–1320

Audrain-McGovern J, Wileyto EP, Ashare R, Cuevas J, Strasser AA (2014) Reward and affective regulation in depression-prone smokers. Biol Psychiatry 76:689–697

Barch DM, Carter CS, Gold JM, Johnson SL, Kring AM, MacDonald AW, Pizzagalli DA, Ragland JD, Silverstein SM, Strauss ME (2017) Explicit and implicit reinforcement learning across the psychosis spectrum. J Abnorm Psychol 126:694–711

Barr RS, Pizzagalli DA, Culhane MA, Goff DC, Evins AE (2008) A single dose of nicotine enhances reward responsiveness in nonsmokers: implications for development of dependence. Biol Psychiatry 63:1061–1065

Berghorst LH, Bogdan R, Frank MJ, Pizzagalli DA (2013) Acute stress selectively reduces reward sensitivity. Front Hum Neurosci 7:133

Bogdan R, Santesso DL, Fagerness J, Perlis RH, Pizzagalli DA (2011) Corticotropin-releasing hormone receptor type 1 (CRHR1) genetic variation and stress interact to influence reward learning. J Neurosci 31:13246–13254

Bouton ME (2016) Learning and behavior: a contemporary synthesis, 2nd edn. Sinauer Associates, Sunderland

Brown VM, Zhu L, Solway A, Wang JM, McCurry KL, King-Casas B, Chiu PH (2021) Reinforcement learning disruptions in individuals with depression and sensitivity to symptom change following cognitive behavioral therapy. JAMA Psychiat 78:1113–1122

Calabrese JR, Fava M, Garibaldi G, Grunze H, Krystal AD, Laughren T, Macfadden W, Marin R, Nierenberg AA, Tohen M (2014) Methodological approaches and magnitude of the clinical unmet need associated with amotivation in mood disorders. J Affect Disord 168:439–451

Cardinal RN, Howes NJ (2005) Effects of lesions of the nucleus accumbens core on choice between small certain rewards and large uncertain rewards in rats. BMC Neurosci 6:37

Cardinal RN, Parkinson JA, Hall J, Everitt BJ (2002) Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. Neurosci Biobehav Rev 26:321–352

Cavanagh JF, Bismark AW, Frank MJ, Allen JJB (2019) Multiple dissociations between comorbid depression and anxiety on reward and punishment processing: evidence from computationally informed EEG. Comput Psychiatr 3:1–17

Costa VD, Dal Monte O, Lucas DR, Murray EA, Averbeck BB (2016) Amygdala and ventral striatum make distinct contributions to reinforcement learning. Neuron 92:505–517

Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. Neuron 69:1204–1215

Der-Avakian A, Pizzagalli DA (2018) Translational assessments of reward and anhedonia: a tribute to Athina Markou. Biol Psychiatry 83:932–939

Der-Avakian A, D'Souza MS, Pizzagalli DA, Markou A (2013) Assessment of reward responsiveness in the response bias probabilistic reward task in rats: implications for cross-species translational research. Transl Psychiatry 3(8):e297

Der-Avakian A, Barnes SA, Markou A, Pizzagalli DA (2016) Translational assessment of reward and motivational deficits in psychiatric disorders. Curr Top Behav Neurosci 28:231–262

Der-Avakian A, D'Souza MS, Potter DN, Chartoff EH, Carlezon WA, Pizzagalli DA, Markou A (2017) Social defeat disrupts reward learning and potentiates striatal nociceptin/orphanin FQ mRNA in rats. Psychopharmacology 234:1603–1614

Dichter GS, Rodriguez-Romaguera J (2022) Anhedonia and hyperhedonia in autism and related neurodevelopmental disorders. Curr Top Behav Neurosci. https://doi.org/10.1007/7854_2022_312

Dowd EC, Frank MJ, Collins A, Gold JM, Barch DM (2016) Probabilistic reinforcement learning in patients with schizophrenia: relationships to anhedonia and avolition. Biol Psychiatry Cogn Neurosci Neuroimaging 1:460–473

Dreher JC, Kohn P, Berman KF (2006) Neural coding of distinct statistical properties of reward information in humans. Cereb Cortex 16:561–573

Duprat R, De Raedt R, Wu G-R, Baeken C (2016) Intermittent theta burst stimulation increases reward responsiveness in individuals with higher hedonic capacity. Front Hum Neurosci 10:294

Eskelund K, Karstoft KI, Andersen SB (2018) Anhedonia and emotional numbing in treatment-seeking veterans: behavioural and electrophysiological responses to reward. Eur J Psychotraumatol 9:1–12

Fletcher K, Parker G, Paterso A, Fava M, Iosifesc D, Pizzagalli DA (2015) Anhedonia in melancholic and non-melancholic depressive disorders. J Affect Disord 184:81–88
Frank MJ (2005) Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated parkinsonism. J Cogn Neurosci 17:51–72
Frank MJ, Seeberger LC, O'Reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 306:1940–1943
Gershon AA, Vishne T, Grunhaus L (2007) Dopamine D2-like receptors and the antidepressant response. Biol Psychiatry 61:145–153
Geugies H, Mocking RJT, Figueroa CA, Groot PFC, Marsman JC, Servaas MN, Steele JD, Schene AH, Ruhé HG (2019) Impaired reward-related learning signals in remitted unmedicated patients with recurrent depression. Brain 142:2510–2522
Ghods-Sharifi S, Onge JRS, Floresco SB (2009) Fundamental contribution by the basolateral amygdala to different forms of decision making. J Neurosci 29:5251–5259
Gilbert DG, Stone BM (2022) Anhedonia in nicotine dependence. Curr Top Behav Neurosci. https://doi.org/10.1007/7854_2022_320
Gottfried JA, O'Doherty J, Dolan RJ (2003) Encoding predictive reward value in human amygdala and orbitofrontal cortex. Science 301:1104–1107
Gradin VB, Kumar P, Waiter G, Ahearn T, Stickle C, Milders M, Reid I, Hall J, Steele JD (2011) Expected value and prediction error abnormalities in depression and schizophrenia. Brain 134: 1751–1764
Grob S, Pizzagalli DA, Dutra SJ, Stern J, Mörgeli H, Milos G, Schnyder U, Hasler G (2012) Dopamine-related deficit in reward learning after catecholamine depletion in unmedicated, remitted subjects with bulimia nervosa. Neuropsychopharmacology 37:1945–1952
Halahakoon DC, Kieslich K, O'Driscoll C, Nair A, Lewis G, Roiser JP (2020) Reward-processing behavior in depressed participants relative to healthy volunteers: a systematic review and meta-analysis. JAMA Psychiat 77:1286–1295
Hammen C (2005) Stress and depression. Annu Rev Clin Psychol 1:293–319
Henriques JB, Glowacki JM, Davidson RJ (1994) Reward fails to alter response bias in depression. J Abnorm Psychol 103:460–466
Hikosaka K, Watanabe M (2000) Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. Cereb Cortex 10:263–271
Huys QJ, Pizzagalli DA, Bogdan R, Dayan P (2013) Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. Biol Mood Anxiety Disord 3:12
Hyman JM, Holroyd CB, Seamans JK (2017) A novel neural prediction error found in anterior cingulate cortex ensembles. Neuron 95:447–456
Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, Sanislow C, Wang P (2010) Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry 167:748–751
Jaffe RJ, Novakovic V, Peselow ED (2013) Scopolamine as an antidepressant: a systematic review. Clin Neuropharmacol 36:24–26
Kaiser RH, Treadway MT, Wooten DW, Kumar P, Goer F, Murray L, Beltzer M, Pechtel P, Whitton A, Cohen AL, Alpert NM, El Fakhri G, Normandin MD, Pizzagalli DA (2018) Frontostriatal and dopamine markers of individual differences in reinforcement learning: a multi-modal investigation. Cereb Cortex 28:4281–4290
Kangas BD, Bergman J (2017) Touchscreen technology in the study of cognition-related behavior. Behav Pharmacol 28:623–629
Kangas BD, Wooldridge LM, Luc OT, Bergman J, Pizzagalli DA (2020) Empirical validation of a touchscreen probabilistic reward task in rats. Transl Psychiatry 10:285
Kangas BD, Short AK, Luc OT, Stern HS, Baram TZ, Pizzagalli DA (2022) A cross-species assay demonstrates that reward responsiveness is enduringly impacted by adverse, unpredictable early-life experiences. Neuropsychopharmacology 47:767–775

Kennerley SW, Behrens TE, Wallis JD (2011) Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. Nat Neurosci 14:1581–1589
Kim J, Farchione T, Potter A, Chen Q, Temple R (2019) Esketamine for treatment-resistant depression-first FDA-approved antidepressant in a new class. N Engl J Med 381:1–4
Kirkpatrick B, Buchanan RW (1990) The neural basis of the deficit syndrome of schizophrenia. J Nerv Ment Dis 178:545–555
Kolling N, Wittmann MK, Behrens TE, Boorman ED, Mars RB, Rushworth MF (2016) Value, search, persistence and model updating in anterior cingulate cortex. Nat Neurosci 19:1280–1285
Koob GF (2022) Anhedonia, hyperkatifeia, and negative reinforcement in substance use disorders. Curr Top Behav Neurosci. https://doi.org/10.1007/7854_2022_288
Krystal AD, Pizzagalli DA, Smoski M, Mathew SJ, Nurnberger J Jr, Lisanby SH, Iosifescu D, Murrough JW, Yang H, Weiner RD, Calabrese JR, Sanacora G, Hermes G, Keefe RSE, Song A, Goodman W, Szabo ST, Whitton AE, Gao K, Potter WZ (2020) A randomized proof-of-mechanism trial applying the 'fast-fail' approach to evaluating κ-opioid antagonism as a treatment for anhedonia. Nat Med 26:760–768
Kumar P, Waiter G, Ahearn T, Milders M, Reid I, Steele JD (2008) Abnormal temporal difference reward-learning signals in major depression. Brain 131:2084–2093
Lally N, Nugent AC, Luckenbaugh DA, Niciu MJ, Roiser JP, Zarate CA Jr (2015) Neural correlates of change in major depressive disorder anhedonia following open-label ketamine. J Psychopharmacol 29:596–607
Lamontagne SJ, Melendez SI, Olmstead MC (2018) Investigating dopamine and glucocorticoid systems as underlying mechanisms of anhedonia. Psychopharmacology 235:3103–3113
Lawson RP, Seymour B, Loh E, Lutti A, Dolan RJ, Dayan P, Weiskopf N, Roiser JP (2014) The habenula encodes negative motivational value associated with primary punishment in humans. Proc Natl Acad Sci U S A 111:11858–11863
Lawson RP, Nord CL, Seymour B, Thomas DL, Dayan P, Pilling S, Roiser JP (2017) Disrupted habenula function in major depression. Mol Psychiatry 22:202–208
Lionello-DeNolf KM (2009) The search for symmetry: 25 years in review. Learn Behav 37:188–203
Liu WH, Chan RC, Wang LZ, Huang J, Cheung EF, Gong QY, Gollan JK (2011) Deficits in sustaining reward responses in subsyndromal and syndromal major depression. Prog Neuro-Psychopharmacol Biol Psychiatry 35:1045–1052
Liu WH, Roiser JP, Wang LZ, Zhu YH, Huang J, Neumann DL, Shum DHK, Cheung EFC, Chan RCK (2016) Anhedonia is associated with blunted reward sensitivity in first-degree relatives of patients with major depression. J Affect Disord 190:640–648
Luc OT, Pizzagalli DA, Kangas BD (2021) Toward a quantification of anhedonia: unified matching law and signal detection for clinical assessment and drug development. Perspect Behav Sci 44:517–540
McCarthy D, Davison M (1979) Signal probability, reinforcement, and signal detection. J Exp Anal Behav 32:373–382
McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. Neuron 38:339–346
McIntyre RS, Lee Y, Zhou AJ, Rosenblat JD, Peters EM, Lam RW, Kennedy SH, Rong C, Jerrell JM (2017) The efficacy of psychostimulants in major depressive episodes: a systematic review and meta-analysis. J Clin Psychopharmacol 37:412–418
Miyazaki K, Miyazaki KW, Sivori G, Yamanaka A, Tanaka KF, Doya K (2020) Serotonergic projections to the orbitofrontal and medial prefrontal cortices differentially modulate waiting for future rewards. Science. Advances 6:eabc7246
Mobini S, Body S, Ho MY, Bradshaw CM, Szabadi E, Deakin JF, Anderson IM (2002) Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. Psychopharmacology 160:290–298
Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci 16:1936–1947

Moran EK, Culbreth AJ, Barch DM (2022) Anhedonia in schizophrenia. Curr Top Behav Neurosci. https://doi.org/10.1007/7854_2022_321
Morris BH, Bylsma LM, Yaroslavsky I, Kovacs M, Rottenberg J (2015) Reward learning in pediatric depression and anxiety: preliminary findings in a high-risk sample. Depress Anxiety 32:373–381
Murray SM, Brown CS, Kaye WH, Wierenga CE (2022) Anhedonia in eating disorders. Curr Top Behav Neurosci. https://doi.org/10.1007/7854_2021_287
National Institute of Mental Health (2016) Behavioral assessment methods for RDoC constructs (revised August 2016). https://www.nimh.nih.gov/about/advisory-boards-and-groups/namhc/reports/behavioral-assessment-methods-for-rdoc-constructs.shtml
O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304:452–454
Packard MG, Knowlton BJ (2002) Learning and memory functions of the Basal Ganglia. Annu Rev Neurosci 25:563–593
Pavlov IP (1927) Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. Oxford University Press
Pechtel P, Pizzagalli DA (2013) Disrupted reinforcement learning and maladaptive behavior in women with a history of childhood sexual abuse: a high-density event-related potential study. JAMA Psychiat 70:499–507
Pechtel P, Dutra SJ, Goetz EL, Pizzagalli DA (2013) Blunted reward responsiveness in remitted depression. J Psychiatr Res 47:1864–1869
Pizzagalli DA, Roberts AC (2022) Prefrontal cortex and depression. Neuropsychopharmacology 47:225–246
Pizzagalli DA, Jahn AL, O'Shea JP (2005) Toward an objective characterization of an anhedonic phenotype: a signal-detection approach. Biol Psychiatry 57:319–327
Pizzagalli DA, Evins AE, Schetter EC, Frank MJ, Pajtas PE, Santesso DL, Culhane M (2008a) Single dose of a dopamine agonist impairs reinforcement learning in humans: Behavioral evidence from a laboratory-based measure of reward responsiveness. Psychopharmacology 196:221–232
Pizzagalli DA, Goetz E, Ostacher M, Iosifescu DV, Perlis RH (2008b) Euthymic patients with bipolar disorder show decreased reward learning in a probabilistic reward task. Biol Psychiatry 64:162–168
Pizzagalli DA, Iosifescu D, Hallett LA, Ratner KG, Fava M (2008c) Reduced hedonic capacity in major depressive disorder: evidence from a probabilistic reward task. J Psychiatr Res 43:76–87
Pizzagalli DA, Smoski M, Ang YS, Whitton AE, Sanacora G, Mathew SJ, Nurnberger J Jr, Lisanby SH, Iosifescu DV, Murrough JW, Yang H, Weiner RD, Calabrese JR, Goodman W, Potter WZ, Krystal AD (2020) Selective kappa-opioid antagonism ameliorates anhedonic behavior: evidence from the fast-fail trial in mood and anxiety spectrum disorders (FAST-MAS). Neuropsychopharmacology 45:1656–1663
Reilly EE, Whitton AE, Pizzagalli DA, Rutherford AV, Stein MB, Paulus MP, Taylor CT (2020) Diagnostic and dimensional evaluation of implicit reward learning in social anxiety disorder and major depression. Depress Anxiety 37:1221–1230
Rogers RD, Everitt BJ, Baldacchino A, Blackshaw AJ, Swainson R, Wynne K, Baker NB, Hunter J, Carthy T, Booker E, London M, Deakin JF, Sahakian BJ, Robbins TW (1999a) Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms. Neuropsychopharmacology 20:322–339
Rogers RD, Owen AM, Middleton HC, Williams EJ, Pickard JD, Sahakian BJ, Robbins TW (1999b) Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. J Neurosci 19:9029–9038
Rolls ET (2019) The orbitofrontal cortex and emotion in health and disease, including depression. Neuropsychologia 128:14–43

Rupprechter S, Stankevicius A, Huys QJM, Series P, Steele JD (2021) Abnormal reward valuation and event-related connectivity in unmedicated major depressive disorder. Psychol Med 51:795–803

Rutledge RB, Moutoussis M, Smittenaar P, Zeidman P, Taylor T, Hrynkiewicz L, Lam J, Skandali N, Siegel JZ, Ousdal OT, Prabhu G, Dayan P, Fonagy P, Dolan RJ (2017) Association of neural and emotional impacts of reward prediction errors with major depression. JAMA Psychiat 74:790–797

Sadacca BF, Wikenheiser AM, Schoenbaum G (2017) Toward a theoretical role for tonic norepinephrine in the orbitofrontal cortex in facilitating flexible learning. Neuroscience 345:124–129

Saez RA, Saez A, Paton JJ, Lau B, Salzman CD (2017) Distinct roles for the amygdala and orbitofrontal cortex in representing the relative amount of expected reward. Neuron 95:70–77

Sandman CF, Craske MG (2022) Psychological treatments for anhedonia. Curr Top Behav Neurosci. https://doi.org/10.1007/7854_2021_291

Santesso DL, Evins AE, Frank MJ, Schetter EC, Bogdan R, Pizzagalli DA (2009) Single dose of a dopamine agonist impairs reinforcement learning in humans: evidence from event-related potentials and computational modeling of striatal-cortical function. Hum Brain Mapp 30:1963–1976

Schoenbaum G, Roesch M (2005) Orbitofrontal cortex, associative learning, and expectancies. Neuron 47:633–636

Schultz W (2002) Getting formal with dopamine and reward. Neuron 36:241–263

Schultz W, Dickinson A (2000) Neuronal coding of prediction errors. Annu Rev Neurosci 23:473–500

Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275:1593–1599

Siddiqi SH, Haddad N, Fox MD (2022) Circuit-targeted neuromodulation for anhedonia. Curr Top Behav Neurosci

Silverman JL, Nithianantharajah J, Der-Avakian A, Young JW, Sukoff Rizzo SJ (2020) Lost in translation: at the crossroads of face validity and translational utility of behavioral assays in animal models for the development of therapeutics. Neurosci Biobehav Rev 116:452–453

Soltani A, Izquierdo A (2019) Adaptive learning under expected and unexpected uncertainty. Nat Rev Neurosci 20:635–644

St Onge JR, Floresco SB (2010) Prefrontal cortical contribution to risk-based decision making. Cereb Cortex 20:1816–1828

St Onge JR, Stopper CM, Zahm DS, Floresco SB (2012) Separate prefrontal-subcortical circuits mediate different components of risk-based decision making. J Neurosci 32:2886–2899

Stalnaker T, Cooch N, Schoenbaum G (2015) What the orbitofrontal cortex does not do. Nat Neurosci 18:620–627

Stolyarova A, Izquierdo A (2017) Complementary contributions of basolateral amygdala and orbitofrontal cortex to value learning under uncertainty. elife 6:e27483

Stopper CM, Green EB, Floresco SB (2014) Selective involvement by the medial orbitofrontal cortex in biasing risky, but not impulsive, choice. Cereb Cortex 24:154–162

Strauss GP, Waltz JA, Gold JM (2014) A review of reward processing and motivational impairment in schizophrenia. Schizophr Bull 40:S107–S116

Taylor MJ, Freemantle N, Geddes JR, Bhagwagar Z (2006) Early onset of selective serotonin reuptake inhibitor antidepressant action: systematic review and meta-analysis. Arch Gen Psychiatry 63:1217–1223

Taylor CT, Hoffman SN, Khan AJ (2022) Anhedonia in anxiety disorders. Curr Top Behav Neurosci. https://doi.org/10.1007/7854_2022_319

Tobler PN, O'Doherty JP, Dolan RJ, Schultz W (2006) Human neural learning depends on reward prediction errors in the blocking paradigm. J Neurophysiol 95:301–310

Tripp G, Alsop B (1999) Sensitivity to reward frequency in boys with attention deficit hyperactivity disorder. J Clin Child Psychol 28:366–375

Turner V, Husain M (2022) Anhedonia in neurodegenerative diseases. Curr Top Behav Neurosci

Vinograd M, Stout D, Risbrough VB (2022) Anhedonia in posttraumatic stress disorder: prevalence, phenotypes, and neural circuitry. Curr Top Behav Neurosci. https://doi.org/10.1007/7854_2021_292

Vrieze E, Pizzagalli DA, 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 73:639–645

Waltz JA, Frank MJ, Robinson BM, Gold JM (2007) Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. Biol Psychiatry 62:756–764

Wang S, Leri F, Rizvi SJ (2022) Clinical and preclinical assessments of anhedonia in psychiatric disorders. Curr Top Behav Neurosci. https://doi.org/10.1007/7854_2022_318

Wassum KM, Izquierdo A (2015) The basolateral amygdala in reward learning and addiction. Neurosci Biobehav Rev 57:271–283

Weinberger DR (1987) Implications of normal brain development for the pathogenesis of schizophrenia. Arch Gen Psychiatry 44:660–669

Weinberger DR, Berman KF (1988) Speculation on the meaning of cerebral metabolic hypofrontality in schizophrenia. Schizophr Bull 14:157–168

Whitton AE, Pizzagalli DA (2022) Anhedonia in depression and bipolar disorder. Curr Top Behav Neurosci. https://doi.org/10.1007/7854_2022_323

Winstanley CA, Floresco SB (2016) Deciphering decision making: variation in animal models of effort-and uncertainty-based choice reveals distinct neural circuitries underlying core cognitive processes. J Neurosci 36:12069–12079

Wooldridge LM, Bergman J, Pizzagalli DA, Kangas BD (2021) Translational assessments of reward responsiveness in the marmoset. Int J Neuropsychopharmacol 24:409–418

Zentall TR, Wasserman EA, Urcuioli PJ (2014) Associative concept learning in animals. J Exp Anal Behav 101:130–151