



OPEN ACCESS

EDITED BY

Adam Steel,
Dartmouth College, United States

REVIEWED BY

Lilianne Rivka Mujica-Parodi,
Stony Brook University, United States
Nicholas W. Simon,
University of Memphis, United States
Roshan Cools,
Radboud University, Netherlands

*CORRESPONDENCE

Kenneth T. Kishida
kkishida@wakehealth.edu

†PRESENT ADDRESS

L. Paul Sands,
Human Neuroimaging Laboratory,
Computational Psychiatry Unit, Fralin
Biomedical Research Institute at
Virginia Tech Carilion, Roanoke, VA,
United States

SPECIALTY SECTION

This article was submitted to
Computational Psychiatry,
a section of the journal
Frontiers in Psychiatry

RECEIVED 28 February 2022

ACCEPTED 23 September 2022

PUBLISHED 20 October 2022

CITATION

Liebenow B, Jones R, DiMarco E,
Trattner JD, Humphries J, Sands LP,
Spry KP, Johnson CK, Farkas EB,
Jiang A and Kishida KT (2022)
Computational reinforcement
learning, reward (and punishment),
and dopamine in psychiatric disorders.
Front. Psychiatry 13:886297.
doi: 10.3389/fpsy.2022.886297

COPYRIGHT

© 2022 Liebenow, Jones, DiMarco,
Trattner, Humphries, Sands, Spry,
Johnson, Farkas, Jiang and Kishida.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Computational reinforcement learning, reward (and punishment), and dopamine in psychiatric disorders

Brittany Liebenow^{1,2}, Rachel Jones^{1,2}, Emily DiMarco^{1,2},
Jonathan D. Trattner^{1,2}, Joseph Humphries^{1,2}, L. Paul Sands^{1,2†},
Kasey P. Spry^{1,2}, Christina K. Johnson², Evelyn B. Farkas³,
Angela Jiang² and Kenneth T. Kishida^{1,2,4,5*}

¹Neuroscience Graduate Program, Wake Forest University School of Medicine, Winston-Salem, NC, United States, ²Department of Physiology and Pharmacology, Wake Forest University School of Medicine, Winston-Salem, NC, United States, ³Georgia State University Undergraduate Neuroscience Institute, Atlanta, GA, United States, ⁴Department of Neurosurgery, Wake Forest University School of Medicine, Winston-Salem, NC, United States, ⁵Department of Biomedical Engineering, Wake Forest University School of Medicine, Winston-Salem, NC, United States

In the DSM-5, psychiatric diagnoses are made based on self-reported symptoms and clinician-identified signs. Though helpful in choosing potential interventions based on the available regimens, this conceptualization of psychiatric diseases can limit basic science investigation into their underlying causes. The reward prediction error (RPE) hypothesis of dopamine neuron function posits that phasic dopamine signals encode the difference between the rewards a person expects and experiences. The computational framework from which this hypothesis was derived, temporal difference reinforcement learning (TDRL), is largely focused on reward processing rather than punishment learning. Many psychiatric disorders are characterized by aberrant behaviors, expectations, reward processing, and hypothesized dopaminergic signaling, but also characterized by suffering and the inability to change one's behavior despite negative consequences. In this review, we provide an overview of the RPE theory of phasic dopamine neuron activity and review the gains that have been made through the use of computational reinforcement learning theory as a framework for understanding changes in reward processing. The relative dearth of explicit accounts of punishment learning in computational reinforcement learning theory and its application in neuroscience is highlighted as a significant gap in current computational psychiatric research. Four disorders comprise the main focus of this review: two disorders of traditionally hypothesized hyperdopaminergic function, addiction and schizophrenia, followed by two disorders of traditionally hypothesized hypodopaminergic function, depression and post-traumatic stress disorder (PTSD). Insights gained from a reward processing based reinforcement learning framework about underlying dopaminergic mechanisms and the role of punishment learning (when available) are explored

in each disorder. Concluding remarks focus on the future directions required to characterize neuropsychiatric disorders with a hypothesized cause of underlying dopaminergic transmission.

KEYWORDS

punishment learning, reward learning, reward prediction error (RPE), temporal difference (TD) learning, depression, addiction

Introduction

Psychiatric conditions diagnosed within DSM-5 (1) criteria are based on clinical phenomenon (e.g., patient-reported symptoms and clinician-observation signs) rather than quantifiable physiological markers. Therapeutic regimens are often devised through trial-and-error in an attempt to treat observed symptoms. Unfortunately, underlying etiologies for most psychiatric disorders are currently unknown, so it is unclear how current treatments effect their biologic causes; as a result, unwanted side-effects are a common result (2). Computational Psychiatry research attempts to bridge the gaps in our understanding of the relationship between brain activity and human behavior, particularly in the context of psychiatric medicine, using computational tools including approaches that have long been developed in computational neuroscience basic research (3, 4). A particularly fruitful avenue of work has been the investigation of human decision-making and choice behavior where computational reinforcement learning (5, 6) has been a guiding conceptual framework.

Computational reinforcement learning (RL) theory (5, 6) describes a well-developed field of computer science research that has identified optimal algorithms by which theoretical “agents” may learn to make choices in order to optimize well defined objective functions. RL expanded into neuroscience with the discovery that RL concepts accurately model learning in mammalian brains (7–9). Specifically, temporal difference reward prediction errors (TD-RPE) have been shown to be encoded by dopamine neurons in the midbrains of non-human primates (7, 8) and rodents (9). The general idea of a “reward prediction error” calculation is “the reward received” vs. “the reward expected.” Neural and behavioral correlates to this general calculation can be found in nearly all experiments where expectations and deviations from these expectations can be controlled and robustly delivered to research subjects. However, some calculations of RPEs can explain more of the nuances in behavior as well as associated neural activity. For example, temporal difference reinforcement learning (TDRL) and the TD-RPE hypothesis of dopamine neuron activity can provide a mechanistic explanation of how an unconditioned stimulus (US) becomes associated with a conditioned stimulus (CS) in Pavlovian conditioning behavioral paradigms (Box 1); whereas

Rescorla-Wagner based calculations of RPEs fail to provide this insight (12). Further, in quantitatively controlled experiments, the value of the RPEs can be designed to be significantly different depending on the manner in which the RPE is calculated.

Recent work suggests that dopaminergic signals in the human brain also encode reward prediction errors (10, 11, 13), but whether these dopaminergic signals are specifically “temporal difference reward prediction errors” has not been demonstrated. Whether dopamine neurons encode temporal difference reward prediction errors vs. some other calculation of a “reward prediction error” signal is important because different specifications of the mathematical theory lead to different predictions about the relationship between the organism’s (i.e., human) behavior and the underlying neural activity supporting it. Human behavior is not driven solely by the pursuit of reward and reward maximization. Negative feedback (e.g., injury) and the anticipation of negative outcomes (e.g., the threat of injury or death) have significant influence on the neurobehavioral processes underlying our behavior (14). Despite this, reward-focused TDRL-based accounts of the neural underpinnings of human decision-making and behavior have not explicitly addressed limitations in the ability of dopamine neuron encoding of TD-RPEs to account for and encode punishing experiences (15–18).

Temporal difference reinforcement learning and its use in computational neuroscience and computational psychiatry treats appetitive and aversive experiences as opposite ends of a unidimensional reward spectrum (5, 17, 18). In this way, aversive experiences are modeled mathematically in TDRL as a negatively signed “reward,” represented by a single scalar value. This traditional unidimensional representation of valence in TDRL theory stipulates that appetitive and aversive experiences are inherently anti-correlated in the natural environment and, further, predicts that rewards and punishments are processed and influence behavioral control processes symmetrically, putatively *via* the dopamine system. However, empirical and theoretical studies in psychology and neuroscience indicate that this assumption might not accurately reflect the nature of reward and punishment learning and associated behavioral responses in humans and other animals (15–23).

BOX 1 Temporal difference reinforcement learning.

Temporal difference reinforcement learning provides a computational framework to investigate how an agent might learn directly from experience. The goal of TDRL algorithms are to estimate the value of a particular state or of an action paired with a state (a state-action value) in order to maximize the rewards that can be obtained. At the core of this family of algorithms is a teaching signal called a temporal difference reward prediction error (TD-RPE), Equation 1:

$$\delta_t = [outcome_t + \gamma V(S_{t+1})] - V(S_t) \quad (1)$$

Here, δ_t is the TD-RPE quantity at time t , which is determined by evaluating the *outcome* magnitude (i.e., reward) observed at time t added to the expected future value, $V(S_{t+1})$, of being in the present state. $V(S_{t+1})$ is discounted by the term γ . All of this is compared to the overall value expected at time $V(S_t)$. This amounts generally to the idea of reward received vs. reward expected. However, the time indices cause the agent to evaluate not only what is received in the present state, but also consider how being in the present state increases or decreases the overall value of future states. It can be shown that the agent only needs to estimate one step into the future (5, 6) to have an optimal approach of associating the estimated value of future states with states that predict the occurrence of those future high (or low) value states. This reward prediction error (δ_t) is then used to update the estimated value of the current state following Equation 2:

$$V'(S_t) \leftarrow V(S_t) + (\alpha \bullet \delta_t) \quad (2)$$

where α is a fractional multiplier that controls how much the current reward prediction error updates the (apparently wrong) estimate of the value of the current state $V(S_t)$ to a new estimate $V'(S_t)$.

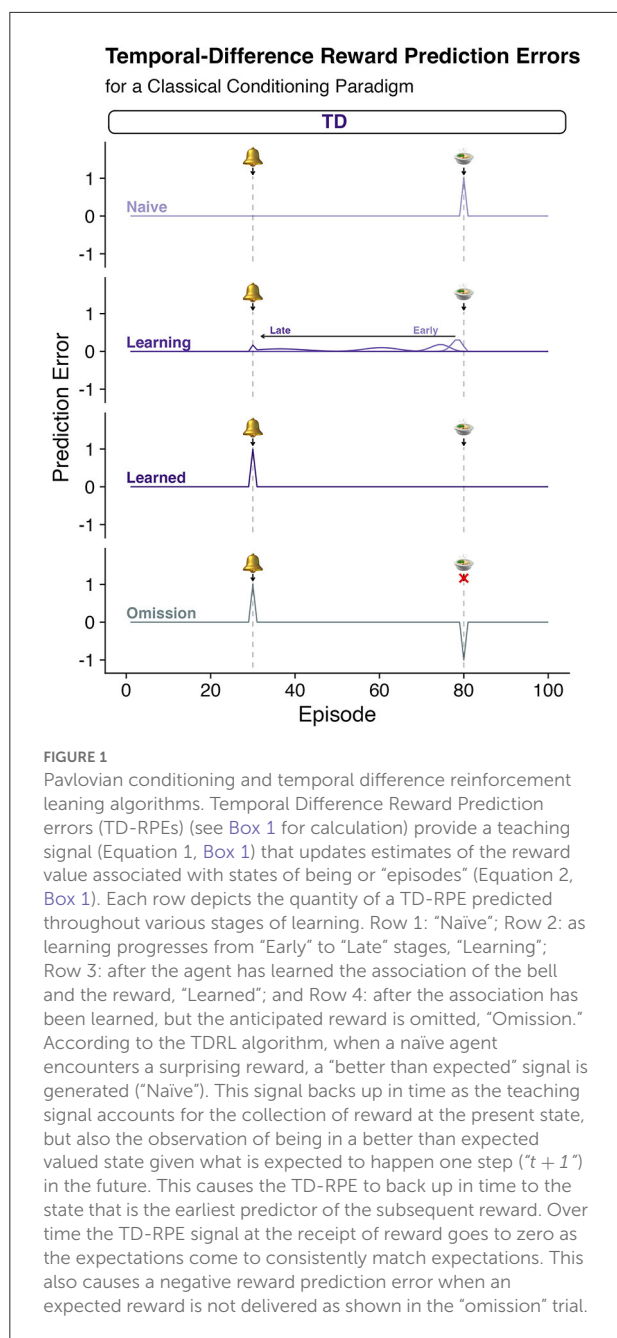
Using these computations, one can simulate simple Pavlovian conditioning paradigms (Figure 1; or more complex operant behavior in humans). Figure 1 shows a simple CS-US pairing where the conditioning stimulus (CS) precedes a reliable reward (i.e., the unconditioned stimulus, US). Figure 1 shows that, prior to any learning, the surprising reward elicits a positive TD-RPE. Over time the learning rules expressed in equations 1 and 2 lead to the TD-RPE backing up in time (to earlier episodes of the trial) to the earliest reliable predictor: here, the bell. The lack of the response at $t = 80$ is consistent with there being a TD-RPE equal to zero – the reward received is exactly as expected. If, however, the reward is omitted (fourth row, “Omission”) a negative reward prediction error is emitted signaling that things are worse than expected.

Human decision-making can and has been successfully studied using this computational framework and state-action elaborations depicted with Q-learning. By interacting with the environment, we learn the consequences of our actions and adapt our behavior accordingly. These behavioral processes and TDRL models have been linked to fluctuations in the firing rates of mid-brain dopaminergic neurons, which have been shown to encode temporal difference (TD) reward prediction errors (RPEs; Equation 1) in response to better-than-expected or worse-than-expected outcomes (positive and negative RPEs, respectively) (10, 11). However, it remains unclear how these neurons can encode variations in the magnitude of punishment since negative RPEs (from simple omissions) drive the neurons to stop responding, which means that worse than expected outcomes that vary in magnitude (e.g., worse than simply not getting a reward) cannot be differentiated.

Genetic research into learning from feedback indicates that separate genes may govern the dopaminergic mechanisms underlying learning to positive and negative outcomes (24). Further, mammalian mesencephalic dopamine neurons demonstrate low baseline firing rates (25), and it is therefore unclear how pauses in dopamine neuron activity are able to effectively communicate magnitude variations across all aversive experiences (negative RPEs) in the same manner as for rewards (positive RPEs), or how downstream brain regions might decode this information conveyed by dopaminergic neuron silence and use it for further behavioral control (15). In primates, midbrain dopaminergic neurons demonstrate excitatory firing behavior following appetitive rewards and pauses following aversive punishments (7, 8, 26). However, evidence suggests that a separate population of dopaminergic neurons may demonstrate excitatory activity in response to punishing information (26–28). For example, different populations of dopaminergic neurons in the rodent ventral tegmental area (VTA) have been reported to respond with excitatory activity to rewarding or punishing outcomes, respectively (27). Local field potential (LFP) data in rodents corroborates this separation of dopaminergic excitatory activity for rewarding vs. aversive learning; namely, theta oscillations increase with rewarding—but not punishing—feedback (28).

While these data support the idea that the “dopaminergic system” can encode both reward and punishment prediction errors, it also suggests that there are separate neural systems within the dopaminergic neuron pool that appear to partition along positively and negatively valenced feedback, which in turn may underlie dopamine’s role in learning to escape aversive situations (29). That separate dopaminergic neurons (but also other neuromodulatory systems) may encode both reward or punishment related information independently complicates an otherwise simple explanation. However, the application of computational models may be one way that clarity can be gained as these models are explicit about how separate systems combine information and contribute to complex behaviors.

These and other insights have inspired new RL-based computational theories that aim to capture diverse aspects of the influence of rewards and punishments on human neural activity, choice behaviors, and affective experiences (18, 23, 29, 30). Still, the vast majority of RL computational models employed to study human neural systems and decision-making behaviors rely on a unidimensional representation of outcome valence as the modeled driver of adaptive learning. Though contemporary applications of TDRL theory to understanding human decision-making in psychiatric conditions have seen early successes, reward-centric computational theories are



ultimately limited in their ability to distinguish the unique influence of rewards and punishments on human neural activity and associated affective behaviors that are manifest in a variety of psychiatric conditions.

Newer reinforcement learning models have incorporated novel reinforcement learning terms as a method to separately investigate behavior and neurochemical responses to appetitive and aversive stimuli (18, 24, 31). Collins and Frank conceptualize stimuli as evaluated by a “critic” encoded

by phasic dopamine in the ventral striatum in Opponent actor learning (OpAL) (31). “Go” and “NoGo” weights are calculated separately for appetitive and aversive stimuli (31). These are hypothesized to be encoded by D1 receptors in the direct pathway, and D2 receptors in the indirect pathway, respectively (31). Another proposed way to decouple algorithmic representations of rewards and punishments is a valence-partitioned reinforcement learning (VPRL) proposed by Sands and Kishida (16–18). We use VPRL (Box 2) as one putative example for how models incorporating aversive feedback may differ from traditional RL models. VPRL is based on the observation that dopaminergic neurons have a low baseline firing rate (4 Hz), and may not have sufficient bandwidth to encode variations in punishment magnitude by decreases in firing (though see Bayer et al.) (17, 18, 25). VPRL provides a generative account that may explain asymmetric representations of positive and negative outcomes that guide adaptive decision-making. This includes the relative weighting of benefits vs. costs when making decisions. Further, it retains the successful accounting of TDRL-based optimal reward learning while hypothesizing a parallel independent process for punishment learning.

In this review, we discuss the benefits that reward-centric TDRL accounts have provided for computational psychiatric investigations into the brain-behavior based accounts of psychiatric conditions and highlight where considerations of punishment learning have also provided insight. We begin our discussion with two disorders that are classically hypothesized to have predominantly hyperdopaminergic action as a potential underlying etiology: addiction and schizophrenia (32, 33). In both behavioral and substance based additions, it is hypothesized that dopaminergic reinforcement systems are manipulated by repetitive risky decisions or pharmacological agents directly or indirectly (32, 33). In schizophrenia, hallmark positive symptoms like psychosis may result from hyperdopaminergic activity of $D_{2/3}$ receptors in the striatum (34–36). However, application of a reinforcement learning framework demonstrates that hypodopaminergic action of stimuli may account for the duality of negative symptoms (37). We continue our discussion with the insights that computational reinforcement learning has had on understanding reward processing in disorders of primarily hypothesized hypodopaminergic signaling. Depressive symptoms found in major depressive disorder (MDD) may arise from altered reward and punishment learning, consistent with depression as a disorder of learning from feedback (38–40). We also highlight here that PTSD may be understood as a reinforcement learning disorder and discuss how future investigations using a computational framework may provide a roadmap for understanding the etiology of PTSD as well as the efficacy of current exposure-based therapies for PTSD.

BOX 2 Valence-partitioned reinforcement learning.

One hypothesized solution to the limitations of TDRL (Box 1) is to partition “outcomes” according to their valence. Positively valenced outcomes (e.g., those that promote survival and reproduction) are handled *via* a positive-valence system, whereas negatively valenced outcomes (e.g., those that would lead to death if unchecked) are handled by a negatively valenced system. Both systems are hypothesized, here, to learn optimally from experience and hence use the TDRL-prediction error, but for partitioned valence specific receptive fields (Equations 3 and 4):

$$\delta_t^P = \begin{cases} outcome_t + \gamma^P V^P(S_{t+1}) - V^P(S_t) & \text{if } outcome_t > 0 \\ 0 + \gamma^P V^P(S_{t+1}) - V^P(S_t) & \text{if } outcome_t \leq 0 \end{cases} \quad (3)$$

$$\delta_t^N = \begin{cases} |outcome_t| + \gamma^N V^N(S_{t+1}) - V^N(S_t) & \text{if } outcome_t < 0 \\ 0 + \gamma^N V^N(S_{t+1}) - V^N(S_t) & \text{if } outcome_t \geq 0 \end{cases} \quad (4)$$

δ_t^P and δ_t^N are each TD-prediction errors that are calculated in much the same way as TD-RPEs in Box 1, except that the “outcome” processed at time t are conditioned on the positive (greater than zero, Equation 3) or negative (less than zero, Equation 4) valence. If the valence of the outcome is not within the respective positive or negative receptive field, then the outcome is treated as a null, or zero, outcome for that respective system.

As in TDRL, the prediction errors update representations of the appetitive (Positive system) and aversive (Negative system) values $V^P(S_t)$ (Equation 5) and $V^N(S_t)$ (Equation 6), respectively.

$$V^P(S_t) \leftarrow V^P(S_t) + (\alpha^P \bullet \delta_t^P) \quad (5)$$

$$V^N(S_t) \leftarrow V^N(S_t) + (\alpha^N \bullet \delta_t^N) \quad (6)$$

The estimated appetitive (positive system) and aversive (negative system) values $V^P(S_t)$ (Equation 5) and $V^N(S_t)$ (Equation 6) can then (at any time) be integrated to provide an overall estimate of expected value (Equation 7):

$$V(S_t) \leftarrow V^P(S_t) - V^N(S_t) \quad (7)$$

This approach maintains the optimal learning algorithm from TDRL reward processing and it is expected that the results of experiments that only vary reward will result in the same predictions when comparing traditional TDRL (Box 1) and this instantiation of VPRL (Figure 2). However, in experiments where punishment learning is investigated, the two approaches will have very different predictions about the calculated reward and punishment learning signals (Figure 2).

Behavioral addictions, substance use disorder, and impulse control disorder

Addiction has long been conceptualized as a spectrum disorder that includes substance based as well as behavioral addictions (41–43). The repetitive engagement in behaviors detrimental to an individual is required for the diagnosis of all addictions—including substance use disorders and behavioral addiction. Also, a loss of control or inhibition often defines the pathology of addictive disorders (1, 41–44). Work investigating the dopaminergic system and associated RL-models, in addiction, are largely focused on the behavioral reinforcing aspects of addictive stimuli, but what is largely lost in these depictions is the diminished influence of the punishing or aversive stimuli that are also experienced by the patient yet appear to have little influence on behavioral change. The following section will consider processes common to behavioral addictions and substance use disorders as part of an extended addiction spectrum and details relevant to specific disorders highlighted by name.

The initial stages of any addiction are suggestive of both classical and operant conditioning (45, 46). On first use, addictive substances or behaviors elicit physiological and psychological unconditioned responses, which increase the likelihood of repeating the actions preceding the drug-taking or behaviorally rewarding experience (47). TDRL prescribes an algorithmic approach for associating the unconditioned ‘reward’ response with the conditioning stimuli or operant behavior (5, 6). Physiologically, these events are strongly associated with a sudden increase in extracellular dopamine levels and is likely associated with the psychological experience “better than expected” (11). In the TDRL framework, this reward prediction error and would serve as a teaching signal for updating the “value” of the preceding states and actions (5, 6). Importantly, the actual “value” of the psychological experience need not correspond to the “value” being updated in the dopaminergic-TDRL hypothesis (7, 8). Substances of abuse can directly and indirectly cause an increase in extracellular dopamine levels (32, 47, 48), which the brain may interpret as a signal to increase the estimated value of an on-drug state (32, 33). The hypothesis is that brain receives a signal that neurochemically reinforces the drug-taking state — akin to this state being of

Temporal-Difference versus Valence-Partitioned Prediction Errors

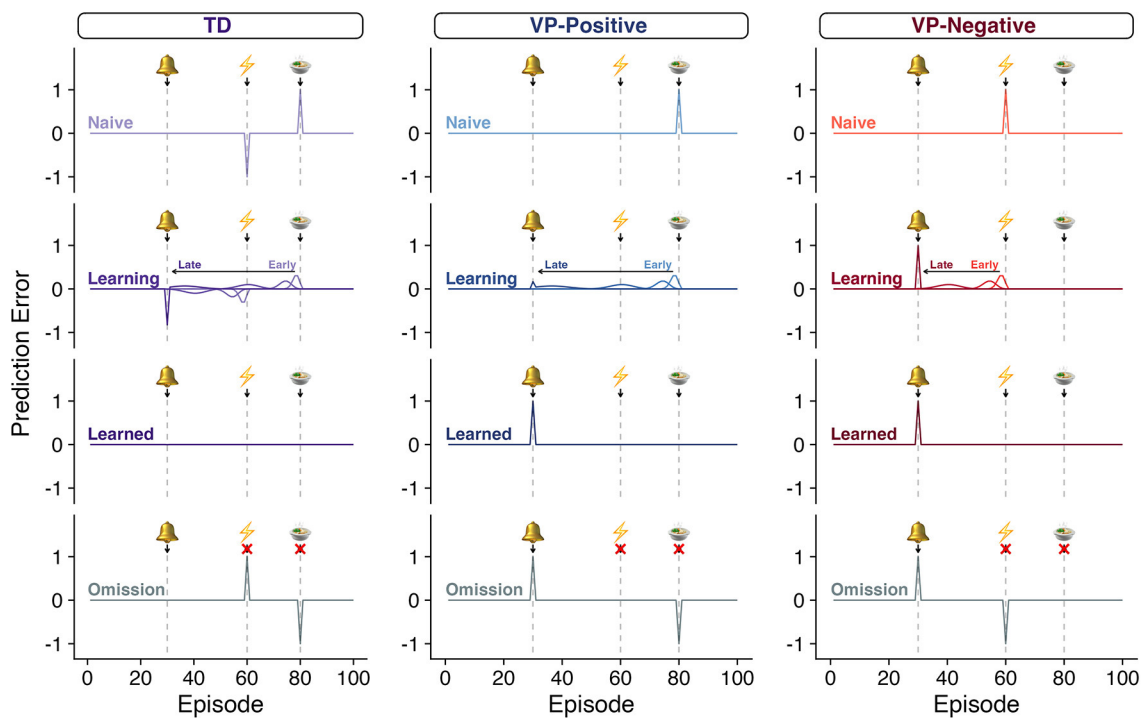


FIGURE 2

Classic TDRL algorithms do not distinguish rewards and punishments. As in Figure 1, each row depicts the quantity of a “prediction error” predicted throughout various stages of learning according to TDRL (column 1, Box 1, Equations 1–2) or valence-partitioned reinforcement learning (VPRL, columns 2 and 3, Box 2, Equations 3–7). Row 1: “Naive”; Row 2: as learning progresses from “Early” to “Late” stages, “Learning”; Row 3: after the agent has learned the association of the bell and the outcomes, “Learned”; and Row 4: after the association has been learned, but the anticipated outcome is omitted, “Omission.” In a paradigm that contains both a reward and a punishment in each trial, classic-TDRL generates a positive RPE for surprising gains and a negative RPE for surprising losses (TD - “Naive,” row 1). As these outcomes are learned, the TD-RPE backs up in time as specified in Eqs. 1, 2 and shown (TD - “Learning,” row 2; for rewards only in Figure 1). However, once these associations are learned (“Learned,” row 3) the positive and negative values of the eventual outcome cancel each other out and the associated TD-RPEs. This behavior of the TDRL algorithm is symptomatic of a system that does not independently distinguish rewards and punishments. As one hypothetical remedy, we propose VPRL (columns 2 and 3, Box 2, Equations 3–7). The prediction errors generated by the Positive System (VP-Positive) and Negative System (VP-Negative) behave in much the same way as the traditional TDRL algorithms, but by having these systems partitioned they are able to independently provide learning signals that can update both the positive and negative aspects of episodes and actions (if applied in Q-learning contexts). This would allow a system to use both error signals (predictive positive and predictive negative) to motivate behavior and evaluate the associated cost and benefits without unintentionally superimposing the valence of coincident events.

“higher value” than expected. This neurochemical reinforcement is not necessary associated with a positive subjective experience; in fact, individuals suffering from addiction may continue an addictive drug or behavior in an attempt to return to prior state (49–51). Notably, these repetitive behaviors in the face of negative feedback would appear as a relative underweighting of aversive or punishing feedback in addition to the relative overweighting of (no longer) rewarding feedback (49–51).

From this viewpoint, the pathology of addiction can be conceptualized as a disorder of learning caused by an overvaluation of behaviors associated with artificial increases in dopamine, but also an underweighting of consequent negative experiences. Dopamine increases are caused by pathological levels of a potentially addictive substance or behavior (51, 52). Such a connection between temporal difference reinforcement learning (TDRL) and addiction was put forth by Redish, who

posited that aberrant RPE signaling in individuals suffering from substance use disorder (SUD) may lead to persistent use of addictive substances (32). As Redish noted, many drugs with addictive potential increase dopamine levels directly, like cocaine, or indirectly, like nicotine and heroin, and SUDs may be understood in the context of aberrant dopaminergic TDRL signaling (32). Indeed, individuals with SUDs have been observed to overvalue immediate, drug-associated stimuli and undervalue future alternative rewards (53, 54). This discounting of future outcomes is more severe in patients with SUDs including opioid-dependence and cocaine use disorder in studies assessing delay-discounting by survey tools and monetarily incentivized decision-making tasks (55, 56). Temporal discounting behavior has been shown to predict short-term relapse for individuals completing an inpatient detoxification program for cocaine, opioid, marijuana, and

amphetamine addictions (57). Studies of methamphetamine use disorder also find these individuals to have altered learning with associated hyperactivation of striatal regions while performing reinforcement learning tasks in patients who have relapsed vs. those who have maintained abstinence (58, 59).

The conceptual link between reinforcement learning and addiction has been expanded to include the entire addiction spectrum, including behavioral addictions (BAs) (33). Individuals with BAs share patterns in cognition and decision-making with patients suffering from SUDs (60, 61). In light of these findings, gambling disorder was recategorized in the DSM-5 as the first non-substance, or behavioral, addiction (62). It is hypothesized that the same underlying dopaminergic pathways critical to reinforcement learning are altered in all forms of addictions (33). If this is true, then core characteristics in patients suffering along the spectrum of addictions (including substance use disorders) may be revealed by studying decision-making in specific behavioral addictions (1, 63). There is, however, a major difference between substance induced addictions and behavioral addictions — addictive substances can be neurotoxic at the levels typically consumed by individuals with substance use disorders, whereas behavioral addictions are not known to have similar neurotoxic or neurodegenerative profiles (1, 61, 63–65). Thus, the study of BAs may allow for a separation of the underlying neural mechanisms shared by all addictions without the confounding factor of the neurotoxicity-induced changes that are caused by specific addictive substances (i.e., cocaine vs. alcohol).

A special class of BAs, Impulse Control Disorder (ICD), has furthered the investigation of addiction and the role reinforcement learning mechanisms may play. Impulse control disorders (ICDs) are identified as rapid changes in risky behavior that arise secondary to dopaminergic therapies for the treatment of Parkinson's disease (PD), particularly dopamine receptor agonists (66). The current gold standard for assessing ICD history is the Questionnaire for Impulsive-Compulsive Disorders in Parkinson's Disease-Rating Scale (QUIP-RS), which is a self-report survey instrument (67, 68). In the QUIP-RS, ICD is defined as meeting a threshold of excessive behavior in one or a combination of the following categories: gambling, sex, buying, and eating (67, 68). These ICDs can be present as specific symptoms (for example, gambling or eating disorders) or as a combined group of ICD symptoms; the unifying factor is that the symptoms present as a rapid change in risky behavior elicited by dopaminergic action (67, 68). In a study of over 3,000 patients with PD, dopamine agonist therapy increased the odds of developing ICD by 2–3.5 times (69). Within the subgroup of patients with any ICD, over a quarter had two or more ICDs (69).

As ICDs occur naturally through the course of dopaminergic therapy for patients with Parkinson's disease, patients with ICD have been studied as a way to investigate the relationship between dopamine and behavioral addiction (66, 70, 71). Elevated impulsivity has been shown to be a hallmark both

in patients with pathological gambling, an ICD, and patients with SUDs (72). [¹¹C] raclopride positron emission tomography (PET) measurements during a gambling task have revealed increases in ventral striatal dopamine release in patients with Parkinson's disease with ICD compared to patients with Parkinson's disease without ICD (73).

A gap in the human addiction literature is a rigorous characterization of the individual neurochemical differences present in those who suffer from addiction and how those differences map to behavior. Future work using ICD as a model for behavioral addiction may be a promising avenue to uncover dopaminergic and behavioral signals with the potential to elucidate differences in reward and punishment processing and provide insight for therapeutic development. Dopamine increase during addictive drug use has been suggested to heighten learning following positive reinforcement, thus driving the transition between drug use and addiction and SUD (51, 71, 74). Similarly, it is possible that dopaminergic therapies used to treat PD may amplify dopaminergic signals involved in positive reinforcement, pushing those with an underlying predisposition for ICD into a symptomatic state. This may be particularly true for dopamine receptor agonists, where research suggests that these agents alter dopaminergic action specifically in the ventral and dorsal striatum (75). Directly investigating real-time dopamine signaling in response to reward and punishment in humans with and without a history of ICD is possible (11, 13, 76–78), and should provide critical insight into the neurochemical mechanisms underlying ICD and other addictive disorders.

Recent work suggests that punishment learning ought to also be more closely investigated in behavioral addiction and substance use disorder (79–82). In one study, pramipexole—a dopamine receptor agonist known to cause ICD—was given to individuals with SUD and found to increase learning from punishment (80). When patients with PD, with and without ICD, were presented with a probabilistic learning task including rewards and punishments, ICD patients on dopamine receptor agonists demonstrated a relatively lower learning rate to negative reward prediction errors (83, 84). Individuals with opioid use disorder completed the same task, but the researchers concluded that individual variability in learning to both rewards and punishments did not allow for detection of differences across individuals with and without SUD (84, 85). A separate study investigating heroin dependency found that individuals with a heroin dependency learned to successfully acquire rewards with equal accuracy as a control group, but the control group outperformed the group with heroin dependency when challenged to avoid punishments (86).

Further research on punishment learning in addiction may expand treatment options for addictive disorders, even employing medications already on the market for alternative uses (87). For example, the antipsychotic sulpiride, a D₂ receptor antagonist, has already demonstrated differential action on

reversal learning to rewards and punishments in gambling disorder vs. a control group (87). These results suggest that further research targeting investigation of punishment may reveal translational potential within differences in decision-making to punishing outcomes in ICD and across the addiction spectrum (81, 82).

Schizophrenia

Schizophrenia is a psychiatric disorder diagnosed based on a spectrum of positive and negative symptoms (1, 88, 89) (Supplementary Table 1). Positive symptoms are often conceptualized as the outwards presentation of the disorder, notably hallucinations, delusions, and disordered speech (1). Negative symptoms include the inwards presentation of the disorder, which may consist of diminished emotional expression, asociality, anhedonia, avolition, and alogia (1). The dopamine hypothesis of schizophrenia suggests that dopamine dysregulation explains the characteristic signs and symptoms of the disorder (37, 89, 90). Recent animal and neurochemical imaging research — including positron emission topography (PET) and single-photon emission computed topography (SPECT) — refined the dopamine hypothesis to specifically include elevated presynaptic dopamine in the striatum as a key factor underlying schizophrenia (91). This hypothesis predicts that the divergent symptomology experienced in patients with schizophrenia (Sz) may result from altered dopaminergic receptor activity affecting distinct brain regions. Evidence from additional PET and SPECT imaging experiments shows that positive symptoms result from hyperdopaminergic activity of D_{2/3} receptors in the striatum (34–36). Negative symptoms and cognitive impairments may be a result of hypofrontality, or decreased D₁ receptor density in the prefrontal cortex (89, 91, 92). Alternatively, schizophrenia may be caused by increased spontaneous dopamine release punctuated by relatively decreased phasic responses to reinforcing stimuli (37). Based on this evidence, current research aims to determine how cognitive and behavioral processes regulated by striatal dopamine and dopaminergic projections to cortical regions are altered in Sz.

Cognitive impairments may be the sentinel sign of Sz—often appearing prior to positive symptoms classically associated with the disorder (93). Set-shifting in particular is a potential cognitive impairment in Sz with applicability to the reinforcement learning field (94–96). Set-shifting tasks require participants to ‘shift’ their attention across different goals based on feedback; for example, the Wisconsin Card Sorting Test (WCST) requires participants to sort cards based on rules that change throughout the test (96–98). It is acknowledged that set-shifting may measure a variety of cognitive skills, including working memory (96). Working memory is an important intermediate for RL because it is essential for

integrating relevant information during goal-directed behavior (99, 100). Anatomically, working memory is modulated by dopamine in the prefrontal cortex and striatum. PET imaging and behavioral studies show that dopamine levels and working memory are related in a U-shaped fashion, with both excess and insufficient levels of dopamine in the prefrontal cortex and striatum impairing working memory (101–105). In Sz, these alterations to dopamine function impair RL processes leading to inaccurate value calculation and non-optimal weighing of potential outcomes (106). Therefore, patients with Sz who possess more intact working memory make more optimal decisions, whereas those with alterations to their working memory make less optimal decisions (106).

In addition to weighing potential outcomes, working memory is associated with alterations in delay discounting of rewards in Sz (101). Delay discounting is the tendency to discount future rewards as a function of time (i.e. the subjective value of a reward declines as the time to the reward increases) (101, 107). Patients with Sz have been shown to have greater delay discounting of rewards, meaning they prefer smaller immediate rewards rather than larger later rewards (101). While other cognitive processes have been hypothesized, it is believed that impaired working memory contributes to greater delay discounting (101, 102, 108). These findings suggest that greater delay discounting may add noise to value estimation functions resulting in altered reward value calculations in Sz (108). Thus, patients with Sz who exhibit impaired working memory, demonstrate greater delay discounting, which results in abnormal reward learning (108).

A hypothesis of shared striatal circuitry that modulates both reinforcement learning (RL) and the dopaminergic dysfunction in Sz has led research into the relationship between schizophrenia and altered adaptive decision-making behavior. Findings based on PET and SPECT research, human neuroimaging studies, and computational models suggest that RPE signaling in the striatum is impaired in Sz compared to controls (91, 106, 109–111). Differences in RL are specifically evident in tasks where patients with Sz are required to utilize positive feedback to optimize rewarding outcomes, resulting in abnormal RPE signaling and altered activity in the ventral striatum (106, 109–111). This abnormal RPE signaling and altered striatal activation is thought to present as behavioral differences in reward and punishment learning in Sz. Reward learning in Sz differs based on the patient’s current experience of positive or negative symptoms. When positive symptoms are experienced, patients with Sz exhibited alterations in reward learning specifically as a rigidity to maintain certain cue-reward associations, even when it was evident that this association was no longer rewarding (106). This association between reward learning and positive symptoms is exemplified during a probabilistic reward and punishment task, in which patients with Sz performed similar to patients with bipolar disorder with psychotic features, both disorders

sharing psychosis as a potential symptom (109). Recent work has attempted to model choice proportions and reaction-time simultaneously using a reinforcement learning drift diffusion model (RLDDM) (112). A RLDDM was able to mathematically explain both choice proportions and reaction time performance during reward learning for those with Sz, and a trend for significant group differences in learning rate was found when comparing those diagnosed with Sz to healthy controls (112).

Patients with Sz who display negative symptoms show alterations in reward learning that lead to a tendency to fail to appropriately estimate and use future rewards to guide decision-making behavior (110, 111). This means patients experiencing negative symptoms tend to struggle to rapidly learn relative reward value and they tend to favor the immediate preceding reward choice, even if that option is less likely to be rewarding in the future (110, 111). Therefore, patients with Sz experiencing negative symptoms tend to select the option with the greatest immediate reward even if longer term rewards will be greater (110, 111). The alteration in reward learning to select and stay with immediate rewarding outcomes is significantly correlated with avolition, defined as the lack of motivation or ability to initiate or persist in goal-directed behaviors (110). Using a RL framework to model these differential associations and alterations in reward learning may provide additional understanding of the mechanisms underlying both positive and negative symptoms in Sz.

In addition to reward learning, patients with Sz show differences in punishment learning, though results in the literature are mixed. For example, patients with Sz have been shown to maintain sensitivity to rewarding stimuli but undervalue potential losses (108). Another study reported no differences in punishment learning on a probabilistic punishment task between patients with Sz and control individuals (109). However, other work found a reduction in the use of punishments to guide future choice behavior in patients with Sz (98, 113, 114). These conflicting results may be clarified by an approach using RL that explicitly accounts for the valence differential of both punishment learning and reward learning. A potential approach that explicitly accounts for a punishment learning system that is distinct from the reward learning system would allow one to experimentally and computationally control different aspects of reward and punishment to determine whether one or the other or both are altered in patients with complex presentations of positive and negative systems (18).

Schizophrenia is a severe psychiatric disorder characterized by alterations in dopaminergic activity eventuating in positive, negative, and cognitive symptoms (115). Alterations in dopamine receptor activity are known to be critical for differences in reinforcement learning behavior in Sz. Changes in RL in Sz appear to be contingent on one's current experience with positive or negative symptoms, which result in differences in working memory, delay discounting, and value calculation

(101, 102, 106, 108). These neural and behavioral changes are thought to underlie aberrant decision-making in Sz (101, 102, 106, 108, 115). Using novel computational RL models that explicitly express the role of various factors could provide clarification of conflicting results and further the understanding of the divergent symptomatology experienced by patients with Sz.

Depression

Major depressive disorder is a leading cause of global disability, affecting over 300 million people worldwide (116). Despite this prevalence ~50% of individuals with MDD do not fully recover using traditional interventions, signifying the need for a better understanding of the neural mechanisms underlying depressive behaviors (117). Though depressed mood is standard for MDD diagnosis, symptoms that meet the DSM-5 criteria vary and may include differences in cognitive function and decision-making (118–120). Emerging evidence suggests these differences may arise from altered reward and punishment learning, characterizing depression as a reinforcement learning disorder (38–40). These behavioral differences are shown to correlate with altered brain activity in a myriad of regions responsive to rewarding and punishing feedback (for example, the prefrontal cortex, nucleus accumbens, amygdala, habenula, anterior insula, or orbitofrontal cortex) (121–125). Reinforcement learning models provide objective characterizations of dynamic brain activity and behavior that are lacking in clinical assessments and may lead to a better understanding of how various brain regions and their interactions are functionally and structurally altered in depression (126). However, the ability of these models to accurately represent the neural processes underlying reward and punishment learning will determine their utility in advancing our understanding and treatment of depression.

Temporal difference reinforcement learning models (5, 6) of human choice behavior are often used to study psychiatric populations including patients with depression. These standard algorithms are grounded in the RPE hypothesis that states that dopaminergic neurons encode the difference between an expected and experienced outcome. The role of reward has been well studied in the context of TDRL models of dopaminergic activity (7, 8, 17). TDRL models are frequently paired with decision-making tasks and functional magnetic resonance imaging (fMRI) to examine choice behavior and associated changes in neural response as humans adaptively learn from rewards and punishments (22, 127–129). This methodology is foundational to our present understanding of reward learning differences specific to depression, whereby many studies support diminished RPE signaling in response to rewarding outcomes, such as monetary rewards, in unmedicated patients with MDD (121). This diminished neural activity is shown within cortico-striatal circuits including the ventral tegmental area (VTA), nucleus accumbens, ventral striatum,

amygdala, hippocampus, and prefrontal cortex (121, 122, 130, 131). Diminished functional connectivity during reward learning is further observed in unmedicated patients with MDD between regions such as the VTA and striatum and between the nucleus accumbens and midcingulate cortex (132). Collectively, studies assessing these reward-learning circuits demonstrate that strength of connectivity correlates negatively with both number of depressive episodes and depression severity (130, 132–135). It is hypothesized that this weakened RPE signaling in depression stems from dopamine signaling deficits downstream of the VTA, which is posited to generate RPE signals (136, 137). For example, one study found intact striatal RPE signaling in patients with MDD during a non-learning task (138). Other studies support diminished RPE signaling in this same brain region in patients with depression during reward learning (99, 130, 131). Together, these findings suggest that reward learning impairments in depression may stem from deficits in RPE signaling rather than a fundamental failure of RPE computation.

Multiple treatment modalities are suggested to improve reward learning impairments experienced by patients with MDD, including antidepressant medications, psychotherapies, and brain stimulation techniques (100, 113, 139). The mechanisms of action for these treatments are diverse, but evidence points (overall) to a role that allows an increase in the representation of RPEs that are generally hypothesized to be encoded by phasic dopamine. Antidepressants are pharmacological agents that modulate serotonin, norepinephrine, or dopamine concentrations in the brain (140). One study found that patients with MDD receiving long-term citalopram, a selective serotonin reuptake inhibitor (SSRI), show increased RPE signaling in the VTA (141). In contrast, diminished VTA RPE signaling was observed among unmedicated patients with MDD and in controls after receiving a 20 mg dosage of citalopram for only 3 days (141). Electrophysiological evidence suggests that SSRIs initially inhibit spontaneously active VTA dopamine neurons through 5-HT 2B/2C receptor pathways (142, 143). However, after long-term administration, these receptors become hyposensitive while DA D2/D3 receptors have been shown to increase (142, 143). The resultant rise in spontaneously active VTA dopamine neurons is hypothesized to drive the enhanced RPE signaling shown in antidepressant-responsive patients (141). The increased responsiveness of DA D2 and D3 receptors are similarly seen in the other types of antidepressants to increase mesolimbic dopaminergic function over time (144). Numerous other studies indicate normalized neural response to RPEs in regions such as the ventral striatum, ventrolateral prefrontal cortex, and orbitofrontal cortex after antidepressant treatment, suggesting the long-term action of these medications may be to modulate RPE signaling and reward learning (145, 146). Further, these findings are consistent with a tight interaction between opponent serotonin and dopamine neural systems generally, but also in the underlying pathology of MDD (15, 147).

Depression is also treated with cognitive behavioral therapy (CBT), an efficacious psychotherapy theorized to reduce symptoms in part through changing learning (114). A recent study found that after 12 weeks of CBT, normalization of reward learning rate occurred in patients with depression alongside improved anhedonia and negative affect (148). Another fMRI study demonstrates that increased neural activity in the ventral striatum, which is known to contain dense regions of dopaminergic terminals, correlates with increased RPE signaling and response to CBT (149). In addition, transcranial magnetic stimulation (TMS) is a common noninvasive brain stimulation treatment that delivers magnetic pulses to regulate the activity of cortical and subcortical structures (150). One study reported increased functional connectivity between the VTA, striatum, and prefrontal cortex among patients with depression responsive to TMS compared with those unresponsive to the procedure (151). Overall, these studies support improved reward learning among patients with depression responsive to various treatment modalities. These behavioral results suggest that prolonged treatments induce a series of adaptive signaling changes downstream of respective therapeutic targets which, in effect, improves reward learning and neural plasticity. It is not clear what mechanism allows the observed increases in reward learning rates following these brain stimulation treatments; however, a positive change in the ‘reward learning rate’ parameter in TDRL models is consistent with the hypothesis that RPEs (hypothesized to be encoded by phasic changes in dopamine) gain influence in effecting behavioral change after these treatments.

Certain brain regions are posited to integrate reward and punishment information encoded by separate, opponent, neural systems relevant to depression. This opponent systems theory hypothesizes that neural systems that track negative reward prediction errors may include serotonergic or a subset of dopaminergic neurons distinct from those that encode RPEs (16, 17, 23). Prediction-error-driven learning is also associated with subjective feelings, a critical aspect of depression as a disease of suffering. ‘Better or worse than expected’ events drive learning and changes in behavior but also are hypothesized to drive dopaminergic signals and thereby affect momentary mood in response to rewarding outcomes (152, 153). One study implemented a “happiness” model using trial-level choice parameters including the value of certain rewards, the expected value of a chosen gamble, and associated RPEs to predict subjective ratings of happiness (138). Results showed that baseline mood parameters correlate with depressive symptoms, though there was unexpectedly no difference between patients with MDD and controls when assessing the expression of dopaminergic RPEs and their impact on self-reported happiness (138).

There has also been investigation into the structural abnormalities associated with reward and punishment learning

systems in depression. These structural foundations for which functional (reward and punishment) prediction error signals are carried are an important aspect to consider as decreases in the integrity of physical connections within reward and punishment valuation networks may underlie loss of reward and punishment efficacy in depression (154). At this time, results appear inconsistent across studies, with disagreement in precisely where the changes exist with respect to these pathways. The major reward system pathways include the cingulum bundle (CB; ventromedial frontal cortex to posterior parietal and temporal cortices), uncinate fasciculus (UF; ventromedial frontal cortex to amygdala), and superolateral medial forebrain bundle (slMFB; anterior limb of internal capsule to frontal brain regions) (155–157). One of the main indices derived using diffusion tensor imaging (DTI) is fractional anisotropy (FA), which is a measure of water movement. FA can provide information about the structural integrity and axonal properties within white matter tracts, as healthier, more myelinated or organized tracts will restrict movement to one direction along the axons. Several DTI studies have shown decreased FA, indicating diminished white matter integrity, in various areas within the CB of patients with MDD (158–161). However, other DTI studies have found no changes in FA of the CB (162–164). These inconsistencies may be due to several factors, such as the effects of medication and disease history that differ between studies (155). The microstructure of the CB may also be indicative of vulnerability to depression, as suggested by family history studies, making it a target for study in depression identification and diagnosis (165, 166). Within the UF, most studies agree that FA is decreased in adult patients with MDD (164, 167, 168). However, results are less consistent in adolescents, with some studies reporting increases in FA while others report decreases, indicating that age and development may be a determining factor (162, 163, 169). Microstructural alterations have also been shown within the slMFB in patients with acute and treatment-resistant depression, making it a target for therapies such as deep brain stimulation to alleviate depression symptoms (170–175). However, unlike CB abnormalities, changes in the UF and slMFB have yet to provide evidence of a biomarker for depression in those with familial history. Interestingly, several studies show that in these pathways of the reward circuit, decreased FA is associated with higher overall melancholic symptoms and depression severity as determined by depression rating scales (155, 166, 170).

Current TDRL research investigating depression also suggests altered punishment learning within certain brain regions (for example, the habenula, insula, or medial prefrontal cortex) (123, 124). The lack of uniform incorporation of punishment learning in dopamine-TDRL theory may have produced conflicting results in the literature regarding differences in punishment learning reported for patients with and without depression (125, 176, 177). For example, fMRI studies show increased “negative-reward-prediction-error”-associated neural responses to monetary loss in the habenula,

anterior insula, and lateral orbitofrontal cortex in unmedicated patients with MDD (123, 134). Further, negative prediction error signals in the habenula are shown to correlate positively with number of depressive episodes, suggesting punishment-related habenula activation increases with disease burden (130, 178). Alternatively, unmedicated patients with MDD have also shown decreased negative prediction error response to monetary loss in the insula, habenula, and prefrontal cortex (177). Other studies show no group differences in habenula activation or connectivity strength between the VTA and habenula for unmedicated patients with MDD and healthy controls during gain or loss conditions (130). Importantly, these studies use similar monetarily-incentivized reinforcement learning tasks yet report different directionality of negative-reward-prediction-error-responsive neural activity. There are also reports of overlapping positive reward prediction error and negative reward prediction error signaling (179); for example, the anterior cingulate cortex shows increased reward prediction error and negative reward prediction error activity among individuals genetically at-risk for depression, and the insula and VTA show increased reward prediction error and negative reward prediction error activity among unmedicated patients with MDD (180).

There has been much less focus linking punishment learning and white matter differences associated with depression. One study did report a significant decrease in mean FA within the VTA-IOFC (ventral tegmental area — lateral orbitofrontal cortex) connection tract of the MFB associated with punishment processing in depressed individuals compared to non-depressed controls (170). However, these results were significant only between controls and patients suffering from melancholic MDD and did not remain significant between controls and non-melancholic MDD. As such, white matter structural abnormalities may be a key feature in the reward and punishment dysfunction associated with depression symptoms and severity.

Overall, reward learning has been widely investigated in depression, but future research may use computational frameworks to explicitly parse reward and punishment learning processes to identify and delineate hypothesized reward and punishment learning neural circuits. In turn, this may help pinpoint reward and punishment system pathways in the human brain and better capture how humans—and especially patients with MDD—may react differently from complicated interactions of appetitive and aversive outcomes in natural experience. While computational formulations that better account for these systems will help, it is important to note that factors such as clinical heterogeneity, different subtypes of depression, and variability in study methodology likely also contribute to discrepancies in the current literature. Still, more detailed computational characterizations of patient behavior, characteristics, and patient populations may provide an approach to better stratify patient subtypes for more effective targeting

of individualized therapeutic strategies. The precision gained through a computational lens may provide hypotheses about how distinct learning processes and associated circuits are altered in depression. These data, paired with white matter abnormalities in individual patients with depression, including MDD, may provide important behavioral, functional, and structural measures to advance our current understanding and treatment of depression.

Post-traumatic stress disorder

A diagnosis of post-traumatic stress disorder (PTSD) by DSM-5 criteria must involve exposure to actual or threatened death, serious injury, or sexual violence; after such an event, the following must be present: (1) the presence of intrusion symptoms associated with the traumatic event; (2) persistent avoidance of stimuli associated with the traumatic event; (3) negative alterations in cognition or mood beginning or worsening after event; (4) marked alteration in arousal and reactivity associated with the traumatic event; and (5) symptoms lasting longer than a month (1, 181). Though the exact mechanism is not yet understood, PTSD has been described as a reinforcement learning deficit (182, 183). Patients who suffer from PTSD have come to associate punishment with environments which are normally neutral or even typically rewarding (1, 181). Psychologically, patients act as though they predict and relive traumatic experiences in relatively innocuous settings (1, 181). Researchers have taken first steps toward understanding reinforcement learning mechanisms underlying PTSD through work developing our understanding of safety cues (182, 183).

It is known that PTSD patients are unable to inhibit fear response in the presence of environmental cues that indicate that an environment associated with punishment will not result in punishment (184). For example, a combat veteran may experience uncontrollable fear in response to a previously learned aversive stimuli (e.g., helicopter sound) when surrounded by many other cues that signal safety (e.g., company of a loved one in a non-threatening environment) (184). Inhibition of fear potentiated startle phenomena in the presence of a safety cue was first observed in a control human cohort in 2005 (185). In a later study, inhibition of fear-potentiated startle to a safety cue was associated with PTSD symptom severity (186).

In humans the safety signal phenotype of PTSD is seen in the startle response, where startle is attenuated by the presence of the safety signal. Startle responses of PTSD patients are not attenuated by the presence of a safety signal compared to controls (187). Administration of L-Dopa during safety learning results in safety memories which are context independent (188). Enhanced dopaminergic activity following omission of aversive stimuli as seen in dopamine rebound is one endogenous mechanism which could mimic this experimental

observation. Further research is necessary to determine if physiologic dopamine rebound is related to safety learning in PTSD. Modeling dopamine rebound and safety learning in a reinforcement learning framework offers a computational approach to address the relationship between PTSD and impaired safety learning. Future work may utilize computational reinforcement learning algorithms to express explicit hypotheses about the way reward and punishment systems interact such that the mechanisms underlying PTSD may be disentangled and new therapeutics developed.

Conclusions and future directions

Temporal difference reinforcement learning theory and the calculation of reward prediction errors therein has been instrumental in providing insight into the information that dopamine neurons encode. The idea that dopamine neurons encode “reward prediction errors” that are important for updating expectations and guiding choice behavior has been a critical one for understanding motivated mammalian behavior and has led computational psychiatric investigations into a number of psychiatric conditions where subjective suffering and aberrant decision-making are hallmark features.

The gains that have been made using TDRL, and computational reinforcement learning theory more generally, to guide computational psychiatric investigations are undeniable. Future computational psychiatric investigations may utilize VPRL as it is described here or some other explicit approach to hypothesize about how the brain may learn to adapt in the face of punishing outcomes. The best approach is to be determined, but we hope to have made the case that some way of accounting for this fundamental aspect of human behavior and psychiatric illness ought be accounted for and investigated further.

We hope to provide those interested in computational psychiatry a clear picture of the importance of extending the widely successfully computational reinforcement learning framework to include punishment learning. The mathematical explication of reward learning mechanisms has allowed rigorous testing of hypothesized processes including the role dopamine neuron activity and dopamine release in the brain may play in human cognition generally, but specifically how these processes may be altered in psychiatric illness (81). We hope to encourage investigators in the area to consider the potential impact equally explicit mathematical representations of punishment learning may have, especially since this is a major gap in the current state of mental health research (189).

Author contributions

BL and KK conceived of and planned the layout and content of this review. BL, RJ, ED, JT, JH, LS, KS, CJ, EF, AJ, and KK contributed to the literature review and drafting of individual sections of this work. All authors reviewed and edited this

manuscript in its entirety for intellectual content, provided final approval of the version to be submitted for review and publication, and agree to be accountable for all aspects of the work including ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding

The authors would like to acknowledge the following funding sources that supported this work: National Institutes of Mental Health (KK: R01MH121099 and R01MH124115), National Institutes of Drug Abuse (KK: R01DA048096 and P50DA006634, BL: F30DA053176, and LS: F31DA053174 and T32DA041349), the National Center for Advancing Translational Sciences (KK: UL1TR001420 and KL2TR001421), and KS is funded by T32NS115704 from National Institute of Neurological Disorders and Stroke and National Institutes of Mental Health (The program, JSPTPN, is jointly sponsored by the NINDS and NIMH - <https://www.ninds.nih.gov/funding/training-career-development/institutional-grants/jointly-sponsored-institutional-predoctoral-training-program-jsptpn>).

References

- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5)*, Vol 51. Washington, DC: American Psychiatric Association (2013), p. 4189.
- Wiecki TV, Poland J, Frank MJ. Model-based cognitive neuroscience approaches to computational psychiatry: clustering and classification. *Clin Psychol Sci.* (2015) 3:378–99. doi: 10.1177/2167702614565359
- Maia TV. Introduction to the series on computational psychiatry. *Clin Psychol Sci.* (2015) 3:374–7. doi: 10.1177/2167702614567350
- Huys QJM, Maia TV, Frank MJ. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat Neurosci.* (2016) 19:404–13. doi: 10.1038/nn.4238
- Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press (1998). doi: 10.1109/TNN.1998.712192
- Sutton RS, Barto AG. *Reinforcement Learning, Second Edition: An Introduction*. Cambridge, MA: MIT Press (2018).
- Montague P, Dayan P, Sejnowski T. A Framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci.* (1996) 16:1936–47. doi: 10.1523/JNEUROSCI.16-05-01936.1996
- Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science.* (1997) 275:1593–9. doi: 10.1126/science.275.5306.1593
- Hart AS, Rutledge RB, Glimcher PW, Phillips PE. Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. *J Neurosci.* (2014) 34:698–704. doi: 10.1523/jneurosci.2489-13.2014
- Zaghoul KA, Blanco JA, Weidemann CT, McGill K, Jaggi JL, Baltuch GH, et al. Human substantia nigra neurons encode unexpected financial rewards. *Science.* (2009) 323:1496–9. doi: 10.1126/science.1167342
- Kishida KT, Saez I, Lohrenz T, Witcher MR, Laxton AW, Tatter SB, et al. Subsecond dopamine fluctuations in human striatum encode superposed error signals about actual and counterfactual reward. *Proc Natl Acad Sci.* (2016) 113:200–5. doi: 10.1073/pnas.1513619112
- McSweeney FK, Murphy ES. *The Wiley Blackwell Handbook of Operant and Classical Conditioning*. Hoboken, NJ: Wiley (2014). doi: 10.1002/9781118468135
- Moran RJ, Kishida KT, Lohrenz T, Saez I, Laxton AW, Witcher MR, et al. The protective action encoding of serotonin transients in the human brain. *Neuropsychopharmacology.* (2018) 43:1425–35. doi: 10.1038/npp.2017.304
- Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk: economics research center. *Econometrica.* (1979) 47:263–91. doi: 10.2307/1914185
- Daw ND, Kakade S, Dayan P. Opponent interactions between serotonin and dopamine. *Neural Netw.* (2002) 15:603–16. doi: 10.1016/S0893-6080(02)00052-7
- Montague PR, Kishida KT, Moran RJ, Lohrenz TM. An efficiency framework for valence processing systems inspired by soft cross-wiring. *Curr Opin Behav Sci.* (2016) 11:121–9. doi: 10.1016/j.cobeha.2016.08.002
- Palminteri S, Pessiglione M. Chapter 23 - Opponent brain systems for reward and punishment learning: causal evidence from drug and lesion studies in humans. In: Dreher J-C, Tremblay L, editors. *Decision Neuroscience*. San Diego, CA: Academic Press (2017), p. 291–303. doi: 10.1016/B978-0-12-805308-9.00023-3
- Kishida KT, Sands LP. A dynamic affective core to bind the contents, context, and value of conscious experience. In: Waugh CE, Kuppens P, editors. *Affect Dynamics*. Cham: Springer International Publishing (2021), p. 293–328. doi: 10.1007/978-3-030-82965-0_12
- Dickinson A, Dearing MF. Appetitive-aversive interactions and inhibitory processes. In *Mechanisms of Learning and Motivation: A Memorial Volume to Jerzy Konorski*. (1979). p. 203–31.
- Seymour B, O'Doherty JP, Koltzenburg M, Wiech K, Frackowiak R, Friston K, et al. Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat Neurosci.* (2005) 8:1234–40. doi: 10.1038/nn1527
- Seymour B, Daw N, Dayan P, Singer T, Dolan R. Differential encoding of losses and gains in the human striatum. *J Neurosci.* (2007) 27:4826–31. doi: 10.1523/JNEUROSCI.0400-07.2007
- Dayan P, Niv Y. Reinforcement learning: the good, the bad and the ugly. *Curr Opin Neurobiol.* (2008) 18:185–96. doi: 10.1016/j.conb.2008.08.003
- Boureau YL, Dayan P. Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology.* (2011) 36:74–97. doi: 10.1038/npp.2010.151
- Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc Natl Acad Sci USA.* (2007) 104:16311–6. doi: 10.1073/pnas.0706111104

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsy.2022.886297/full#supplementary-material>

25. Bayer HM, Lau B, Glimcher PW. Statistics of midbrain dopamine neuron spike trains in the awake primate. *J Neurophysiol.* (2007) 98:1428–39. doi: 10.1152/jn.01140.2006
26. Mirenowicz J, Schultz W. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature.* (1996) 379:449–51. doi: 10.1038/379449a0
27. Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature.* (2012) 482:85–8. doi: 10.1038/nature10754
28. Kim Y, Wood J, Moghaddam B. Coordinated activity of ventral tegmental neurons adapts to appetitive and aversive learning. *PLoS ONE.* (2012) 7:e29766. doi: 10.1371/journal.pone.0029766
29. Lloyd K, Dayan P. Safety out of control: dopamine and defence. *Behav Brain Funct.* (2016) 12:15. doi: 10.1186/s12993-016-0099-7
30. Mikhael JG, Bogacz R. Learning reward uncertainty in the basal Ganglia. *PLoS Comput Biol.* (2016) 12:e1005062. doi: 10.1371/journal.pcbi.1005062
31. Collins AG, Frank MJ. Opponent actor learning (Opal): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychol Rev.* (2014) 121:337–66. doi: 10.1037/a0037015
32. Redish AD. Addiction as a computational process gone awry. *Science.* (2004) 306:1944–7. doi: 10.1126/science.1102384
33. Redish AD, Jensen S, Johnson A. A unified framework for addiction: vulnerabilities in the decision process. *Behav Brain Sci.* (2008) 31:415–37. discussion 37–87. doi: 10.1017/S0140525X0800472X
34. Abi-Dargham A, Gil R, Krystal J, Baldwin RM, Seibyl JP, Bowers M, et al. Increased striatal dopamine transmission in schizophrenia: confirmation in a second cohort. *Am J Psychiatry.* (1998) 155:761–7.
35. Laruelle M, Abi-Dargham A, Dyck C, Krystal J, Charney D, Innis R. Increased striatal dopamine release in drug free schizophrenic subjects. *Biol Psychiat.* (1996) 39:560. doi: 10.1016/0006-3223(96)84153-9
36. Heinz A, Murray GK, Schlagenhauf F, Sterzer P, Grace AA, Walter JA. Towards a unifying cognitive, neurophysiological, and computational neuroscience account of schizophrenia. *Schizophr Bull.* (2019) 45:1092–100. doi: 10.1093/schbul/sby154
37. Maia TV, Frank MJ. An integrative perspective on the role of dopamine in schizophrenia. *Biol Psychiatry.* (2017) 81:52–66. doi: 10.1016/j.biopsych.2016.05.021
38. Eshel N, Roiser JP. Reward and punishment processing in depression. *Biol Psychiatry.* (2010) 68:118–24. doi: 10.1016/j.biopsych.2010.01.027
39. Pizzagalli DA, Dillon DG, Bogdan R, Holmes AJ. *Reward and Punishment Processing in the Human Brain: Clues from Affective Neuroscience and Implications for Depression Research. Neuroscience of Decision Making.* New York, NY: Psychology Press (2011), p. 199–220.
40. Pizzagalli DA. Depression, stress, and anhedonia: toward a synthesis and integrated model. *Annu Rev Clin Psychol.* (2014) 10:393–423. doi: 10.1146/annurev-clinpsy-050212-185606
41. Martin PR, Petry NM. Are non-substance-related addictions really addictions? *Am J Addict.* (2005) 14:1–7. doi: 10.1080/10550490590899808
42. Piccinni A, Costanzo D, Vanelli F, Franceschini C, Cremonese I, Conversano C, et al. A controversial side of addiction: new insight in eating behavior. *Heroin Addict Relat Clin Probl.* (2013) 15:45–7.
43. Dunne FJ, Feeney S, Schipperheijn J. Eating disorders and alcohol misuse: features of an addiction spectrum. *Postgrad Med J.* (1991) 67:112–3. doi: 10.1136/pgmj.67.78.4.112
44. Augustine A, Winstanley CA, Krishnan V. Impulse control disorders in Parkinson's disease: from bench to bedside. *Front Neurosci.* (2021) 15:654238. doi: 10.3389/fnins.2021.654238
45. Pavlov PI. Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. *Ann Neurosci.* (2010) 17:136–41. doi: 10.5214/ans.0972-7531.1017309
46. Skinner BF. *The Behavior of Organisms: An Experimental Analysis.* Oxford: Appleton-Century (1938), p. 457.
47. Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci.* (2005) 8:1481–9. doi: 10.1038/nn1579
48. Adinoff B. Neurobiologic processes in drug reward and addiction. *Harv Rev Psychiatry.* (2004) 12:305–20. doi: 10.1080/10673220490910844
49. Koob GF, Le Moal M. Drug addiction, dysregulation of reward, and allostasis. *Neuropsychopharmacology.* (2001) 24:97–129. doi: 10.1016/S0893-133X(00)00195-0
50. Koob GF, Le Moal M. Addiction and the brain antireward system. *Annu Rev Psychol.* (2008) 59:29–53. doi: 10.1146/annurev.psych.59.103006.093548
51. Koob GF. Negative reinforcement in drug addiction: the darkness within. *Curr Opin Neurobiol.* (2013) 23:559–63. doi: 10.1016/j.conb.2013.03.011
52. Djamshidian A, Jha A, O'Sullivan SS, Silveira-Moriyama L, Jacobson C, Brown P, et al. Risk and learning in impulsive and nonimpulsive patients with Parkinson's disease. *Mov Disord.* (2010) 25:2203–10. doi: 10.1002/mds.23247
53. Bickel WK, Miller ML, Yi R, Kowal BP, Lindquist DM, Pitcock JA. Behavioral and neuroeconomics of drug addiction: competing neural systems and temporal discounting processes. *Drug Alcohol Depend.* (2007) 90:S85–91. doi: 10.1016/j.drugalcdep.2006.09.016
54. Baker TE, Zeigami Y, Dagher A, Holroyd CB. Smoking decisions: altered reinforcement learning signals induced by nicotine state. *Nicotine Tob Res.* (2020) 22:164–71.
55. Robles E, Huang BE, Simpson PM, McMillan DE. Delay discounting, impulsiveness, and addiction severity in opioid-dependent patients. *J Subst Abuse Treat.* (2011) 41:354–62. doi: 10.1016/j.jsat.2011.05.003
56. Cox DJ, Dolan SB, Johnson P, Johnson MW. Delay and probability discounting in cocaine use disorder: comprehensive examination of money, cocaine, and health outcomes using gains and losses at multiple magnitudes. *Exp Clin Psychopharmacol.* (2020) 28:724–38. doi: 10.1037/pha0003341
57. Stevens L, Goudriaan AE, Verdejo-García A, Dom G, Roeyers H, Vanderplasschen W. Impulsive choice predicts short-term relapse in substance-dependent individuals attending an in-patient detoxification programme. *Psychol Med.* (2015) 45:2083–93. doi: 10.1017/S003329171500001X
58. Stewart JL, Connolly CG, May AC, Tapert SF, Wittmann M, Paulus MP. Striatum and insula dysfunction during reinforcement learning differentiates abstinent and relapsed methamphetamine-dependent individuals. *Addiction.* (2014) 109:460–71. doi: 10.1111/add.12403
59. Robinson AH, Perales JC, Volpe I, Chong TT, Verdejo-García A. Are methamphetamine users compulsive? Faulty reinforcement learning, not inflexibility, underlies decision making in people with methamphetamine use disorder. *Addict Biol.* (2021) 26:e12999. doi: 10.1111/adb.12999
60. Lawrence AJ, Luty J, Bogdan NA, Sahakian BJ, Clark L. Impulsivity and response inhibition in alcohol dependence and problem gambling. *Psychopharmacology.* (2009) 207:163–72. doi: 10.1007/s00213-009-1645-x
61. Albein-Urios N, Martínez-González JM, Lozano O, Clark L, Verdejo-García A. Comparison of impulsivity and working memory in cocaine addiction and pathological gambling: implications for cocaine-induced neurotoxicity. *Drug Alcohol Depend.* (2012) 126:1–6. doi: 10.1016/j.drugalcdep.2012.03.008
62. Clark L. Disordered gambling: the evolving concept of behavioral addiction. *Ann N Y Acad Sci.* (2014) 1327:46–61. doi: 10.1111/nyas.12558
63. Römer Thomsen K, Fjorback LO, Møller A, Lou HC. Applying incentive sensitization models to behavioral addiction. *Neurosci Biobehav Rev.* (2014) 45:343–9. doi: 10.1016/j.neubiorev.2014.07.009
64. Lawrence AJ, Luty J, Bogdan NA, Sahakian BJ, Clark L. Problem gamblers share deficits in impulsive decision-making with alcohol-dependent individuals. *Addiction.* (2009) 104:1006–15. doi: 10.1111/j.1360-0443.2009.02533.x
65. Römer Thomsen K, Joensson M, Lou HC, Møller A, Gross J, Kringelbach ML, et al. Altered paralimbic interaction in behavioral addiction. *Proc Natl Acad Sci.* (2013) 110:4744–9. doi: 10.1073/pnas.1302374110
66. Schrag A, Sauerbier A, Chaudhuri KR. New clinical trials for nonmotor manifestations of Parkinson's disease. *Mov Disord.* (2015) 30:1490–504. doi: 10.1002/mds.26415
67. Evans AH, Okai D, Weintraub D, Lim SY, O'Sullivan SS, Voon V, et al. Scales to assess impulsive and compulsive behaviors in Parkinson's disease: critique and recommendations. *Mov Disord.* (2019) 34:791–8. doi: 10.1002/mds.27689
68. Weintraub D, Mamikonyan E, Papay K, Shea JA, Xie SX, Siderowf A. Questionnaire for impulsive-compulsive disorders in Parkinson's disease-rating scale. *Mov Disord.* (2012) 27:242–7. doi: 10.1002/mds.24023
69. Weintraub D, Koester J, Potenza MN, Siderowf AD, Stacy M, Voon V, et al. Impulse control disorders in Parkinson disease: a cross-sectional study of 3090 patients. *Arch Neurol.* (2010) 67:589–95. doi: 10.1001/archneurol.2010.65
70. Samuel M, Rodriguez-Oroz M, Antonini A, Brothie JM, Ray Chaudhuri K, Brown RG, et al. Management of impulse control disorders in Parkinson's disease: controversies and future approaches. *Mov Disord.* (2015) 30:150–9. doi: 10.1002/mds.26099
71. Clark CA, Dagher A. The role of dopamine in risk taking: a specific look at Parkinson's disease and gambling. *Front Behav Neurosci.* (2014) 8:196. doi: 10.3389/fnbeh.2014.00196

72. Potenza MN. Should addictive disorders include non-substance-related conditions? *Addiction*. (2006) 101(Suppl 1):142–51. doi: 10.1111/j.1360-0443.2006.01591.x
73. Steeves TD, Miyasaki J, Zurowski M, Lang AE, Pellicchia G, Van Eimeren T, et al. Increased striatal dopamine release in parkinsonian patients with pathological gambling: a [11c] raclopride pet study. *Brain*. (2009) 132(Pt 5):1376–85. doi: 10.1093/brain/awp054
74. Cohen MX, Frank MJ. Neurocomputational models of basal ganglia function in learning, memory and choice. *Behav Brain Res*. (2009) 199:141–56. doi: 10.1016/j.bbr.2008.09.029
75. Wu K, Politis M, Piccini P. Parkinson disease and impulse control disorders: a review of clinical features, pathophysiology and management. *Postgrad Med J*. (2009) 85:590–6. doi: 10.1136/pgmj.2008.075820
76. Kishida KT, Sandberg SG, Lohrenz T, Comair YG, Sáez I, Phillips PE, et al. Sub-second dopamine detection in human striatum. *PLoS ONE*. (2011) 6:e23291. doi: 10.1371/journal.pone.0023291
77. Bang D, Kishida KT, Lohrenz T, White JP, Laxton AW, Tatter SB, et al. Sub-second dopamine and serotonin signaling in human striatum during perceptual decision-making. *Neuron*. (2020) 108:999–1010.e6. doi: 10.1016/j.neuron.2020.09.015
78. Liebenow B, Williams M, Wilson T, Haq IU, Siddiqui MS, Laxton AW, et al. Intracranial approach for sub-second monitoring of neurotransmitters during Dbs electrode implantation does not increase infection rate. *PLoS ONE*. (2022) 17:e0271348. doi: 10.1371/journal.pone.0271348
79. Gueguen MCM, Schweitzer EM, Konova AB. Computational theory-driven studies of reinforcement learning and decision-making in addiction: what have we learned? *Curr Opin Behav Sci*. (2021) 38:40–8. doi: 10.1016/j.cobeha.2020.08.007
80. Kanen JW, Ersche KD, Fineberg NA, Robbins TW, Cardinal RN. Computational modelling reveals contrasting effects on reinforcement learning and cognitive flexibility in stimulant use disorder and obsessive-compulsive disorder: remediating effects of dopaminergic D2/3 receptor agents. *Psychopharmacology*. (2019) 236:2337–58. doi: 10.1007/s00213-019-05325-w
81. Jean-Richard-Dit-Bressel P, Killcross S, McNally GP. Behavioral and neurobiological mechanisms of punishment: implications for psychiatric disorders. *Neuropsychopharmacology*. (2018) 43:1639–50. doi: 10.1038/s41386-018-0047-3
82. Ekhtiari H, Victor TA, Paulus MP. Aberrant decision-making and drug addiction—how strong is the evidence? *Curr Opin Behav Sci*. (2017) 13:25–33. doi: 10.1016/j.cobeha.2016.09.002
83. Piray P, Zeighami Y, Bahrami F, Eissa AM, Hewedi DH, Moustafa AA. Impulse control disorders in Parkinson's disease are associated with dysfunction in stimulus valuation but not action valuation. *J Neurosci*. (2014) 34:7814–24. doi: 10.1523/JNEUROSCI.4063-13.2014
84. Bódi N, Kéri S, Nagy H, Moustafa A, Myers CE, Daw N, et al. Reward-learning and the novelty-seeking personality: a between- and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. *Brain*. (2009) 132(Pt 9):2385–95. doi: 10.1093/brain/awp094
85. Myers CE, Sheynin J, Balsdon T, Luzzardo A, Beck KD, Hogarth L, et al. Probabilistic reward- and punishment-based learning in opioid addiction: experimental and computational data. *Behav Brain Res*. (2016) 296:240–8. doi: 10.1016/j.bbr.2015.09.018
86. Myers CE, Rego J, Haber P, Morley K, Beck KD, Hogarth L, et al. Learning and generalization from reward and punishment in opioid addiction. *Behav Brain Res*. (2017) 317:122–31. doi: 10.1016/j.bbr.2016.09.033
87. Janssen LK, Sescousse G, Hashemi MM, Timmer MH, ter Huurne NP, Geurts DE, et al. Abnormal modulation of reward versus punishment learning by a dopamine D2-receptor antagonist in pathological gamblers. *Psychopharmacology*. (2015) 232:3345–53. doi: 10.1007/s00213-015-3986-y
88. Dugré JR, Dumais A, Tikasz A, Mendrek A, Potvin S. Functional connectivity abnormalities of the long-axis hippocampal subregions in schizophrenia during episodic memory. *NPJ Schizophrenia*. (2021) 7:19. doi: 10.1038/s41537-021-00147-2
89. Nakazawa K, Sapkota K. The origin of Nmda receptor hypofunction in schizophrenia. *Pharmacol Ther*. (2020) 205:107426. doi: 10.1016/j.pharmthera.2019.107426
90. Howes OD, Kapur S. The dopamine hypothesis of schizophrenia: version iii—the final common pathway. *Schizophr Bull*. (2009) 35:549–62. doi: 10.1093/schbul/sbp006
91. Deserno L, Schlagenhaut F, Heinz A. Striatal dopamine, reward, and decision making in schizophrenia. *Dialogues Clin Neurosci*. (2016) 18:77–89. doi: 10.31887/DCNS.2016.18.1/ldeserno
92. Correll CU, Schooler NR. Negative symptoms in schizophrenia: a review and clinical guide for recognition, assessment, and treatment. *Neuropsychiatr Dis Treat*. (2020) 16:519–34. doi: 10.2147/NDT.S225643
93. Davidson M, Reichenberg A, Rabinowitz J, Weiser M, Kaplan Z, Mark M. Behavioral and intellectual markers for schizophrenia in apparently healthy male adolescents. *Am J Psychiatry*. (1999) 156:1328–35. doi: 10.1176/ajp.156.9.1328
94. Pantelis C, Barber FZ, Barnes TR, Nelson HE, Owen AM, Robbins TW. Comparison of set-shifting ability in patients with chronic schizophrenia and frontal lobe damage. *Schizophr Res*. (1999) 37:251–70. doi: 10.1016/S0920-9964(98)00156-X
95. Fleming K, Goldberg TE, Gold JM, Weinberger DR. Verbal working memory dysfunction in schizophrenia: use of a brown-Peterson paradigm. *Psychiatry Res*. (1995) 56:155–61. doi: 10.1016/0165-1781(95)02589-3
96. Pantelis C, Wood SJ, Proffitt TM, Testa R, Mahony K, Brewer WJ, et al. Attentional set-shifting ability in first-episode and established schizophrenia: relationship to working memory. *Schizophr Res*. (2009) 112:104–13. doi: 10.1016/j.schres.2009.03.039
97. Grant DA, Berg EA. Wisconsin card sorting test. *J Exp Psychol*. (1993).
98. Prentice KJ, Gold JM, Buchanan RW. The Wisconsin card sorting impairment in schizophrenia is evident in the first four trials. *Schizophr Res*. (2008) 106:81–7. doi: 10.1016/j.schres.2007.07.015
99. Epstein J, Pan H, Kocsis JH, Yang Y, Butler T, Chusid J, et al. Lack of ventral striatal response to positive stimuli in depressed versus normal subjects. *Am J Psychiatry*. (2006) 163:1784–90. doi: 10.1176/ajp.2006.163.10.1784
100. Roiser JP, Elliott R, Sahakian BJ. Cognitive mechanisms of treatment in depression. *Neuropsychopharmacology*. (2012) 37:117–36. doi: 10.1038/npp.2011.183
101. Ahn W-Y, Rasmussen O, Fridberg DJ, Bishara AJ, Forsyth JK, Breier A, et al. Temporal discounting of rewards in patients with bipolar disorder and schizophrenia. *J Abnorm Psychol*. (2011) 120:911. doi: 10.1037/a0023333
102. Wang H, Lesh TA, Maddock RJ, Fassbender C, Carter CS. Delay discounting abnormalities are seen in first-episode schizophrenia but not in bipolar disorder. *Schizophr Res*. (2020) 216:200–6. doi: 10.1016/j.schres.2019.11.063
103. Cools R, Gibbs SE, Miyakawa A, Jagust W, D'Esposito M. Working memory capacity predicts dopamine synthesis capacity in the human striatum. *J Neurosci*. (2008) 28:1208–12. doi: 10.1523/JNEUROSCI.4475-07.2008
104. Brozoski TJ, Brown RM, Rosvold H, Goldman PS. Cognitive deficit caused by regional depletion of dopamine in prefrontal cortex of rhesus monkey. *Science*. (1979) 205:929–32. doi: 10.1126/science.112679
105. Cools R, D'Esposito M. Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biol Psychiatry*. (2011) 69:e113–25. doi: 10.1016/j.biopsych.2011.03.028
106. Gold JM, Waltz JA, Prentice KJ, Morris SE, Heerey EA. Reward processing in schizophrenia: a deficit in the representation of value. *Schizophr Bull*. (2008) 34:835–47. doi: 10.1093/schbul/sbn068
107. Barry AB, Koeppl JA, Ho B-C. Impulsive decision making, brain cortical thickness and familial schizophrenia risk. *Schizophr Res*. (2020) 220:54–60. doi: 10.1016/j.schres.2020.03.072
108. Heerey EA, Bell-Warren KR, Gold JM. Decision-making impairments in the context of intact reward sensitivity in schizophrenia. *Biol Psychiatry*. (2008) 64:62–9. doi: 10.1016/j.biopsych.2008.02.015
109. Abohamza E, Weickert T, Ali M, Moustafa AA. Reward and punishment learning in schizophrenia and bipolar disorder. *Behav Brain Res*. (2020) 381:112298. doi: 10.1016/j.bbr.2019.112298
110. Saperia S, Da Silva S, Siddiqui I, Agid O, Daskalakis ZJ, Ravindran A, et al. Reward-driven decision-making impairments in schizophrenia. *Schizophr Res*. (2019) 206:277–83. doi: 10.1016/j.schres.2018.11.004
111. Waltz JA, Frank MJ, Robinson BM, Gold JM. Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biol Psychiatry*. (2007) 62:756–64. doi: 10.1016/j.biopsych.2006.09.042
112. Pine JG, Pedersen ML, Frank MJ, Barch DM. P528. Computational modeling of reward learning in schizophrenia using the reinforcement learning drift diffusion model (Rlddm). *Biol Psychiatry*. (2022) 91:S302–S3. doi: 10.1016/j.biopsych.2022.02.765
113. Schlaepfer TE, Cohen MX, Frick C, Kosel M, Brodessa D, Axmacher N, et al. Deep brain stimulation to reward circuitry alleviates anhedonia in refractory major depression. *Neuropsychopharmacology*. (2008) 33:368–77. doi: 10.1038/sj.npp.1301408
114. Burkhouse KL, Kujawa A, Kennedy AE, Shankman SA, Langenecker SA, Phan KL, et al. Neural reactivity to reward as a predictor of cognitive behavioral therapy response in anxiety and depression. *Depress Anxiety*. (2016) 33:281–8. doi: 10.1002/da.22482
115. Maia TV, Frank MJ. From reinforcement learning models to psychiatric and neurological disorders. *Nat Neurosci*. (2011) 14:154–62. doi: 10.1038/nn.2723

116. World Health Organization. *Depression and Other Common Mental Disorders: Global Health Estimates*. Geneva: World Health Organization (2017).
117. Culpepper L. Why do you need to move beyond first-line therapy for major depression? *J Clin Psychiatry*. (2010) 71:22466. doi: 10.4088/JCP.9104su1c.01
118. Kennedy SH. Core symptoms of major depressive disorder: relevance to diagnosis and treatment. *Dialogues Clin Neurosci*. (2008) 10:271–7. doi: 10.31887/DCNS.2008.10.3/shkennedy
119. Clark DC, Cavanaugh Sv, Gibbons RD. The core symptoms of depression in medical and psychiatric patients. *J Nerv Mental Dis*. (1983) 171:705–13. doi: 10.1097/00005053-198312000-00001
120. Substance A, Mental Health Services A. *Cbhsq Methodology Report. Dsm-5 Changes: Implications for Child Serious Emotional Disturbance*. Rockville, MD: Substance Abuse and Mental Health Services Administration (US) (2016).
121. Ng TH, Alloy LB, Smith DV. Meta-analysis of reward processing in major depressive disorder reveals distinct abnormalities within the reward circuit. *Transl Psychiatry*. (2019) 9:1–10. doi: 10.1038/s41398-019-0644-x
122. Pizzagalli DA, Holmes AJ, Dillon DG, Goetz EL, Birk JL, Bogdan R, et al. Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder. *Am J Psychiatry*. (2009) 166:702–10. doi: 10.1176/appi.ajp.2008.08081201
123. Rolls ET. The roles of the orbitofrontal cortex via the habenula in non-reward and depression, and in the responses of serotonin and dopamine neurons. *Neurosci Biobehav Rev*. (2017) 75:331–4. doi: 10.1016/j.neubiorev.2017.02.013
124. Hennigan K, D'Ardenne K, McClure SM. Distinct midbrain and habenula pathways are involved in processing aversive events in humans. *J Neurosci*. (2015) 35:198–208. doi: 10.1523/JNEUROSCI.0927-14.2015
125. Garrison J, Erdeniz B, Done J. Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. *Neurosci Biobehav Rev*. (2013) 37:1297–310. doi: 10.1016/j.neubiorev.2013.03.023
126. Shortreed SM, Laber E, Lizotte DJ, Stroup TS, Pineau J, Murphy SA. Inferring sequential clinical decision-making through reinforcement learning: an empirical study. *Mach Learn*. (2011) 84:109–36. doi: 10.1007/s10994-010-5229-0
127. O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ. Temporal difference models and reward-related learning in the human brain. *Neuron*. (2003) 38:329–37. doi: 10.1016/S0896-6273(03)00169-7
128. Frank MJ, Seeberger LC, O'Reilly RC. By carrot or by stick: cognitive reinforcement learning in Parkinsonism. *Science*. (2004) 306:1940–3. doi: 10.1126/science.1102941
129. Tricomi EM, Delgado MR, Fiez JA. Modulation of caudate activity by action contingency. *Neuron*. (2004) 41:281–92. doi: 10.1016/S0896-6273(03)00848-1
130. Kumar P, Goer F, Murray L, Dillon DG, Beltzer ML, Cohen AL, et al. Impaired reward prediction error encoding and striatal-midbrain connectivity in depression. *Neuropsychopharmacology*. (2018) 43:1581–8. doi: 10.1038/s41386-018-0032-x
131. Chen C, Takahashi T, Nakagawa S, Inoue T, Kusumi I. Reinforcement learning in depression: a review of computational research. *Neurosci Biobehav Rev*. (2015) 55:247–67. doi: 10.1016/j.neubiorev.2015.05.005
132. Geugies H, Mocking RJ, Figueroa CA, Groot PF, Marsman J-BC, Servaas MN, et al. Impaired reward-related learning signals in remitted unmedicated patients with recurrent depression. *Brain*. (2019) 142:2510–22. doi: 10.1093/brain/awz167
133. Rothkirch M, Tonn J, Köhler S, Sterzer P. Neural mechanisms of reinforcement learning in unmedicated patients with major depressive disorder. *Brain*. (2017) 140:1147–57. doi: 10.1093/brain/awx025
134. Ubl B, Kuehner C, Kirsch P, Ruttorf M, Diener C, Flor H. Altered neural reward and loss processing and prediction error signalling in depression. *Soc Cogn Affect Neurosci*. (2015) 10:1102–12. doi: 10.1093/scan/nsu158
135. Rupprechter S, Romaniuk L, Series P, Hirose Y, Hawkins E, Sandu A-L, et al. Blunted medial prefrontal cortico-limbic reward-related effective connectivity and depression. *Brain*. (2020) 143:1946–56. doi: 10.1093/brain/awaa106
136. Kahnt T, Park SQ, Cohen MX, Beck A, Heinz A, Wrase J. Dorsal striatal-midbrain connectivity in humans predicts how reinforcements are used to guide decisions. *J Cogn Neurosci*. (2009) 21:1332–45. doi: 10.1162/jocn.2009.21092
137. Lerner TN, Holloway AL, Seiler JL. Dopamine, updated: reward prediction error and beyond. *Curr Opin Neurobiol*. (2021) 67:123–30. doi: 10.1016/j.comb.2020.10.012
138. Rutledge RB, Moutoussis M, Smittenaar P, Zeidman P, Taylor T, Hrynkiewicz L, et al. Association of neural and emotional impacts of reward prediction errors with major depression. *JAMA Psychiatry*. (2017) 74:790–7. doi: 10.1001/jamapsychiatry.2017.1713
139. Forbes EE, Olino TM, Ryan ND, Birmaher B, Axelson D, Moyles DL, et al. Reward-related brain function as a predictor of treatment response in adolescents with major depressive disorder. *Cogn Affect Behav Neurosci*. (2010) 10:107–18. doi: 10.3758/CABN.10.1.107
140. Sheffler ZM, Abdijadid S. *Antidepressants*. Treasure Island, FL: StatPearls Publishing (2022).
141. Kumar P, Waiter G, Ahearn T, Milders M, Reid I, Steele J. Abnormal temporal difference reward-learning signals in major depression. *Brain*. (2008) 131:2084–93. doi: 10.1093/brain/awn136
142. Breuer ME, Groenink L, Oosting RS, Buerger E, Korte M, Ferger B, et al. Antidepressant effects of pramipexole, a dopamine D3/D2 receptor agonist, and 7-Oh-Dpat, a dopamine D3 receptor agonist, in olfactory bulbectomized rats. *Eur J Pharmacol*. (2009) 616:134–40. doi: 10.1016/j.ejphar.2009.06.029
143. Willner P, Hale AS, Argyropoulos S. Dopaminergic mechanism of antidepressant action in depressed patients. *J Affect Disord*. (2005) 86:37–45. doi: 10.1016/j.jad.2004.12.010
144. Bonhomme N, Esposito E. Involvement of serotonin and dopamine in the mechanism of action of novel antidepressant drugs: a review. *J Clin Psychopharmacol*. (1998) 18:447–54. doi: 10.1097/00004714-199812000-00005
145. Stoy M, Schlagenhauf F, Sterzer P, Bermpohl F, Hägele C, Suchtowski K, et al. Hyporeactivity of ventral striatum towards incentive stimuli in unmedicated depressed patients normalizes after treatment with escitalopram. *J Psychopharmacol*. (2012) 26:677–88. doi: 10.1177/0269881111416686
146. Tremblay LK, Naranjo CA, Graham SJ, Herrmann N, Mayberg HS, Hevenor S, et al. Functional neuroanatomical substrates of altered reward processing in major depressive disorder revealed by a dopaminergic probe. *Arch Gen Psychiatry*. (2005) 62:1228–36. doi: 10.1001/archpsyc.62.11.1228
147. Huys QJ, Daw ND, Dayan P. Depression: a decision-theoretic analysis. *Annu Rev Neurosci*. (2015) 38:1–23. doi: 10.1146/annurev-neuro-071714-033928
148. Brown VM, Zhu L, Solway A, Wang JM, McCurry KL, King-Casas B, et al. Reinforcement learning disruptions in individuals with depression and sensitivity to symptom change following cognitive behavioral therapy. *JAMA psychiatry*. (2021) 78:1113–22. doi: 10.1001/jamapsychiatry.2021.1844
149. Nair A, Rutledge RB, Mason L. Under the hood: using computational psychiatry to make psychological therapies more mechanism-focused. *Front Psychiatry*. (2020) 11:140. doi: 10.3389/fpsyt.2020.00140
150. O'Reardon JP, Solvason HB, Janicak PG, Sampson S, Isenberg KE, Nahas Z, et al. Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: a multisite randomized controlled trial. *Biol Psychiatry*. (2007) 62:1208–16. doi: 10.1016/j.biopsych.2007.01.018
151. Downar J, Geraci J, Salomons TV, Dunlop K, Wheeler S, McAndrews MP, et al. Anhedonia and reward-circuit connectivity distinguish nonresponders from responders to dorsomedial prefrontal repetitive transcranial magnetic stimulation in major depression. *Biol Psychiatry*. (2014) 76:176–85. doi: 10.1016/j.biopsych.2013.10.026
152. O'Doherty JP, Buchanan TW, Seymour B, Dolan RJ. Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron*. (2006) 49:157–66. doi: 10.1016/j.neuron.2005.11.014
153. Rolls ET, Grabenhorst F, Parris BA. Warm pleasant feelings in the brain. *Neuroimage*. (2008) 41:1504–13. doi: 10.1016/j.neuroimage.2008.03.005
154. Mayberg HS. Modulating dysfunctional limbic-cortical circuits in depression: towards development of brain-based algorithms for diagnosis and optimised treatment. *Br Med Bull*. (2003) 65:193–207. doi: 10.1093/bmb/65.1.193
155. Bracht T, Linden D, Keedwell P. A review of white matter microstructure alterations of pathways of the reward circuit in depression. *J Affect Disord*. (2015) 187:45–53. doi: 10.1016/j.jad.2015.06.041
156. Bracht T, Tüscher O, Schnell S, Kreher B, Rüscher N, Glauche V, et al. Extraction of prefronto-amygdalar pathways by combining probability maps. *Psychiatry Res Neuroimaging*. (2009) 174:217–22. doi: 10.1016/j.pscychresns.2009.05.001
157. Bubb EJ, Metzler-Baddeley C, Aggleton JP. The cingulum bundle: anatomy, function, and dysfunction. *Neurosci Biobehav Rev*. (2018) 92:104–27. doi: 10.1016/j.neubiorev.2018.05.008
158. Henderson SE, Johnson AR, Vallejo AI, Katz L, Wong E, Gabbay V, et al. Preliminary study of white matter in adolescent depression: relationships with illness severity, anhedonia, and irritability. *Front Psychiatry*. (2013) 4:152. doi: 10.3389/fpsyt.2013.00152
159. Ouyang X, Tao HJ, Liu HH, Deng QJ, Sun ZH, Xu L, et al. White matter integrity deficit in treatment-naïve adult patients with major depressive disorder. *East Asian Arch Psychiatry*. (2011) 21:5–9.

160. Seok JH, Choi S, Lim HK, Lee SH, Kim I, Ham BJ. Effect of the Comt Val158met polymorphism on white matter connectivity in patients with major depressive disorder. *Neurosci Lett.* (2013) 545:35–9. doi: 10.1016/j.neulet.2013.04.012
161. de Diego-Adeliño J, Pires P, Gómez-Ansón B, Serra-Blasco M, Vives-Gilbert Y, Puigdemont D, et al. Microstructural white-matter abnormalities associated with treatment resistance, severity and duration of illness in major depression. *Psychol Med.* (2014) 44:1171–82. doi: 10.1017/S003329171300158X
162. Cullen KR, Klimes-Dougan B, Muetzel R, Mueller BA, Camchong J, Hourri A, et al. Altered white matter microstructure in adolescents with major depression: a preliminary study. *J Am Acad Child Adolesc Psychiatry.* (2010) 49:173–83.e1. doi: 10.1016/j.jaac.2009.11.005
163. LeWinn KZ, Connolly CG, Wu J, Drahos M, Hoeft F, Ho TC, et al. White matter correlates of adolescent depression: structural evidence for frontolimbic disconnectivity. *J Am Acad Child Adolesc Psychiatry.* (2014) 53:899–909.e1–7. doi: 10.1016/j.jaac.2014.04.021
164. Carballo A, Amico F, Uguw I, Fagan AJ, Fahey C, Morris D, et al. Reduced fractional anisotropy in the uncinate fasciculus in patients with major depression carrying the met-allele of the val66met brain-derived neurotrophic factor genotype. *Am J Med Genet B Neuropsychiatr Genet.* (2012) 159b:537–48. doi: 10.1002/ajmg.b.32060
165. Huang H, Fan X, Williamson DE, Rao U. White matter changes in healthy adolescents at familial risk for unipolar depression: a diffusion tensor imaging study. *Neuropsychopharmacology.* (2011) 36:684–91. doi: 10.1038/npp.2010.199
166. Keedwell PA, Chapman R, Christiansen K, Richardson H, Evans J, Jones DK. Cingulum white matter in young women at risk of depression: the effect of family history and anhedonia. *Biol Psychiatry.* (2012) 72:296–302. doi: 10.1016/j.biopsych.2012.01.022
167. de Kwaasteniet B, Ruhe E, Caan M, Rive M, Olabarriaga S, Groefsema M, et al. Relation between structural and functional connectivity in major depressive disorder. *Biol Psychiatry.* (2013) 74:40–7. doi: 10.1016/j.biopsych.2012.12.024
168. Zhang A, Leow A, Ajilore O, Lamar M, Yang S, Joseph J, et al. Quantitative tract-specific measures of uncinate and cingulum in major depression using diffusion tensor imaging. *Neuropsychopharmacology.* (2012) 37:959–67. doi: 10.1038/npp.2011.279
169. Aghajani M, Veer IM, van Lang ND, Meens PH, van den Bulk BG, Rombouts SA, et al. Altered white-matter architecture in treatment-naïve adolescents with clinical depression. *Psychol Med.* (2014) 44:2287–98. doi: 10.1017/S0033291713003000
170. Bracht T, Horn H, Strik W, Federspiel A, Schnell S, Höfle O, et al. White matter microstructure alterations of the medial forebrain bundle in melancholic depression. *J Affect Disord.* (2014) 155:186–93. doi: 10.1016/j.jad.2013.10.048
171. Tha KK, Terae S, Nakagawa S, Inoue T, Kitagawa N, Kako Y, et al. Impaired integrity of the brain parenchyma in non-geriatric patients with major depressive disorder revealed by diffusion tensor imaging. *Psychiatry Res.* (2013) 212:208–15. doi: 10.1016/j.psychres.2012.07.004
172. Zou K, Huang X, Li T, Gong Q, Li Z, Ou-yang L, et al. Alterations of white matter integrity in adults with major depressive disorder: a magnetic resonance imaging study. *J Psychiatry Neurosci.* (2008) 33:525–30.
173. Guo WB, Liu F, Chen JD, Xu XJ, Wu RR, Ma CQ, et al. Altered white matter integrity of forebrain in treatment-resistant depression: a diffusion tensor imaging study with tract-based spatial statistics. *Prog Neuropsychopharmacol Biol Psychiatry.* (2012) 38:201–6. doi: 10.1016/j.pnpbp.2012.03.012
174. Bewernick BH, Kayser S, Gippert SM, Switala C, Coenen VA, Schlaepfer TE. Deep brain stimulation to the medial forebrain bundle for depression- long-term outcomes and a novel data analysis strategy. *Brain Stimul.* (2017) 10:664–71. doi: 10.1016/j.brs.2017.01.581
175. Schlaepfer TE, Bewernick BH, Kayser S, Mädler B, Coenen VA. Rapid effects of deep brain stimulation for treatment-resistant major depression. *Biol Psychiatry.* (2013) 73:1204–12. doi: 10.1016/j.biopsych.2013.01.034
176. Chase HW, Kumar P, Eickhoff SB, Dombrowski AY. Reinforcement learning models and their neural correlates: an activation likelihood estimation meta-analysis. *Cogn Affect Behav Neurosci.* (2015) 15:435–59. doi: 10.3758/s13415-015-0338-7
177. Lawson RP, Nord CL, Seymour B, Thomas DL, Dayan P, Pilling S, et al. Disrupted habenula function in major depression. *Mol Psychiatry.* (2017) 22:202–8. doi: 10.1038/mp.2016.81
178. Liu W-H, Valton V, Wang L-Z, Zhu Y-H, Roiser JP. Association between habenula dysfunction and motivational symptoms in unmedicated major depressive disorder. *Soc Cogn Affect Neurosci.* (2017) 12:1520–33. doi: 10.1093/scan/nsx074
179. Lammell S, Lim BK, Ran C, Huang KW, Betley MJ, Tye KM, et al. Input-specific control of reward and aversion in the ventral tegmental area. *Nature.* (2012) 491:212–7. doi: 10.1038/nature11527
180. McCabe C, Woffindale C, Harmer CJ, Cowen PJ. Neural processing of reward and punishment in young people at increased familial risk of depression. *Biol Psychiatry.* (2012) 72:588–94. doi: 10.1016/j.biopsych.2012.04.034
181. Shalev A, Liberzon I, Marmar C. Post-traumatic stress disorder. *N Engl J Med.* (2017) 376:2459–69. doi: 10.1056/NEJMra1612499
182. Fonzo GA. Diminished positive affect and traumatic stress: a biobehavioral review and commentary on trauma affective neuroscience. *Neurobiol Stress.* (2018) 9:214–30. doi: 10.1016/j.ynstr.2018.10.002
183. Lissek S, van Meurs B. Learning models of PTSD: theoretical accounts and psychobiological evidence. *Int J Psychophysiol.* (2015) 98:594–605. doi: 10.1016/j.ijpsycho.2014.11.006
184. Jovanovic T, Kazama A, Bachevalier J, Davis M. Impaired safety signal learning may be a biomarker of Ptsd. *Neuropharmacology.* (2012) 62:695–704. doi: 10.1016/j.neuropharm.2011.02.023
185. Jovanovic T, Keyes M, Fiallos A, Myers KM, Davis M, Duncan EJ. Fear potentiation and fear inhibition in a human fear-potentiated startle paradigm. *Biol Psychiatry.* (2005) 57:1559–64. doi: 10.1016/j.biopsych.2005.02.025
186. Jovanovic T, Norrholm SD, Fennell JE, Keyes M, Fiallos AM, Myers KM, et al. Posttraumatic stress disorder may be associated with impaired fear inhibition: relation to symptom severity. *Psychiatry Res.* (2009) 167:151–60. doi: 10.1016/j.psychres.2007.12.014
187. Lee JC, Wang LP, Tsien JZ. Dopamine rebound-excitation theory: putting brakes on Ptsd. *Front Psychiatry.* (2016) 7:163. doi: 10.3389/fpsy.2016.00163
188. Haaker J, Gaburro S, Sah A, Gartmann N, Lonsdorf TB, Meier K, et al. Single dose of L-dopa makes extinction memories context-independent and prevents the return of fear. *Proc Natl Acad Sci U S A.* (2013) 110:E2428–36. doi: 10.1073/pnas.1303061110
189. The National Institute of Mental Health. *Definitions of the RDoC Domains and Constructs.* (2015). Available online at: www.nimh.nih.gov (accessed February 28, 2022).