Title: A Reinforcement Learning and Decision-Making framework for understanding Mental Disorders

Authors:
Jeung-Hyun Lee[1*], Eunhwi Lee[1*], Jooyeon Jamie Im[1], John P. O'Doherty[2], Woo-Young Ahn[1, 3, 4]

[1]Department of Psychology, Seoul National University, Seoul, Korea
[2]Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, California, USA
[3]Department of Brain and Cognitive Sciences, Seoul National University, Seoul, Korea
[4]AI Institute, Seoul National University, Seoul, Korea

Number of Figures: 4
Number of words in the abstract: 145
Number of words in the main text: 7,508
Number of references: 178

Corresponding author:
Woo-Young Ahn, Ph.D.
Department of Psychology
Seoul National University
Seoul, Korea 08826
E-mail: wahn55@snu.ac.kr

Conflict of Interest: The authors declare no competing financial interests.

Abstract

While mental disorders are complex and characterized by heterogeneous symptoms, a unified framework that fully and mechanistically captures these complexities remains elusive. Reinforcement learning offers a promising way to understand mental health by mathematically modeling the decision-making processes that underlie psychiatric conditions. By breaking decision-making down into key components—such as state representation, valuation, action selection, and outcome evaluation—reinforcement learning provides a structured approach to studying how disruptions in these processes contribute to disorders like depression, anxiety, and addiction. This review explores how reinforcement learning can help clarify the cognitive and neural mechanisms involved in mental disorders and offers insights into their interactions with other psychological and physiological systems. We also discuss the potential of the framework to improve clinical practice. Future directions will focus on expanding and using the reinforcement learning models to naturalistic paradigms and incorporation with advanced technologies like artificial intelligence.

Introduction

Mental disorders are characterized by heterogenetic and multifaceted nature in symptom expression and underlying mechanisms, reflecting the complex interactions among psychological, cognitive, biological, and environmental factors[1]. This complexity presents a fundamental challenge for clinical psychology, making it difficult to pinpoint exact causes and further design effective interventions. For example, major depressive disorder has 227 possible symptom combinations[2], and even individuals with similar symptoms, such as anhedonia, may experience distinct disruptions in cognitive processes like reward prediction error signaling or inhibitory control[3,4]. Moreover, psychiatric conditions often involve comorbidities, and individuals experience overlapping symptoms from multiple disorders. This adds another layer of complexity, making it harder to identify the underlying mechanisms and complicating diagnosis and treatment.

In light of these challenges, one promising and growing field of investigation involves examining how individuals with psychiatric conditions differ from healthy individuals in their decision-making processes. Across different psychiatric conditions, clinical symptoms often reflect disruptions in how individuals make decisions, particularly in how they evaluate rewards, anticipate outcomes, and adjust their behavior accordingly. This idea has positioned the decision-making paradigm as a common axis along which diverse mental disorders and symptoms can be understood.

Reinforcement learning offers a computational framework for modeling how individuals make value-based decisions by integrating feedback from previous outcomes to guide future choices, which are frequently altered in mental disorders[5–7]. The reinforcement learning decision-making framework allows observed behaviors–those associated with clinical symptoms–to be linked to latent cognitive mechanisms through computational parameters, offering insight into the underlying sources of behavioral variability (**Fig.1**). This framework can be broken down into key stages: state representation, valuation, action selection, and outcome evaluation. Each of these processes is interconnected, such that a disturbance in one component can propagate through the stages and manifest as psychiatric conditions. In **Section 1**, we provide an overview of reinforcement learning, exploring the role of each stage in the reinforcement learning framework.

Reinforcement learning models, and generative models more broadly, simulate data based on the mathematical formulations[4,8–10], allowing researchers to quantify psychological constructs–such as reward sensitivity and decision biases–that are not directly observable but are crucial for understanding the underlying cognitive mechanisms in mental disorders[3,5,11].The ability to simulate complex systems and test hypotheses through reinforcement learning models makes them powerful for understanding latent features of decision-making processes. The insights gained from the framework are particularly useful for exploring the interactions between different psychological domains, spanning various cognitive functions and clinical symptoms. In **Section 2**, we discuss how reinforcement learning facilitates the study of these interactions, exploring how cognitive processes, such as memory, and larger systems like the brain and body are interconnected, and how these connections contribute to mental health issues.

The applicability of reinforcement learning extends beyond just the theoretical understanding of mental disorders; it provides empirical evidence into how specific cognitive processes, such as reward sensitivity and learning are altered in conditions like depression, anxiety, and addiction. By modeling decision-making in these contexts, reinforcement learning framework helps identify how impairments in these processes contribute to maladaptive behaviors. In **Section 3**, we explore the empirical evidence demonstrating how reinforcement learning has enhanced our mechanistic understanding of these disorders. Furthermore, reinforcement learning's contributions extend to clinical applications, offering ways to tailor effective interventions considering individual differences. This section will also highlight reinforcement learning's potential in bringing methodological advancements to enable more rapid and reliable estimates in clinical psychology.

Overall, this review aims to introduce reinforcement learning as a framework for advancing our understanding of mental health disorders. This review contributes uniquely to the existing literature by proposing how we can dissect psychopathology into distinct computational components within a reinforcement learning framework—state representation, valuation, action selection, and outcome evaluation. Building upon prior reviews, we aim to integrates neurocognitive domains such as memory, attention, emotion, and interoceptive signals into a cohesive (computational) reinforcement learning framework. At the end of the review, we suggest future directions for applying reinforcement learning in mental health research and personalized treatment, and advocates the adoption of naturalistic tasks and advanced artificial intelligence techniques to address some of the remaining challenges in the field.

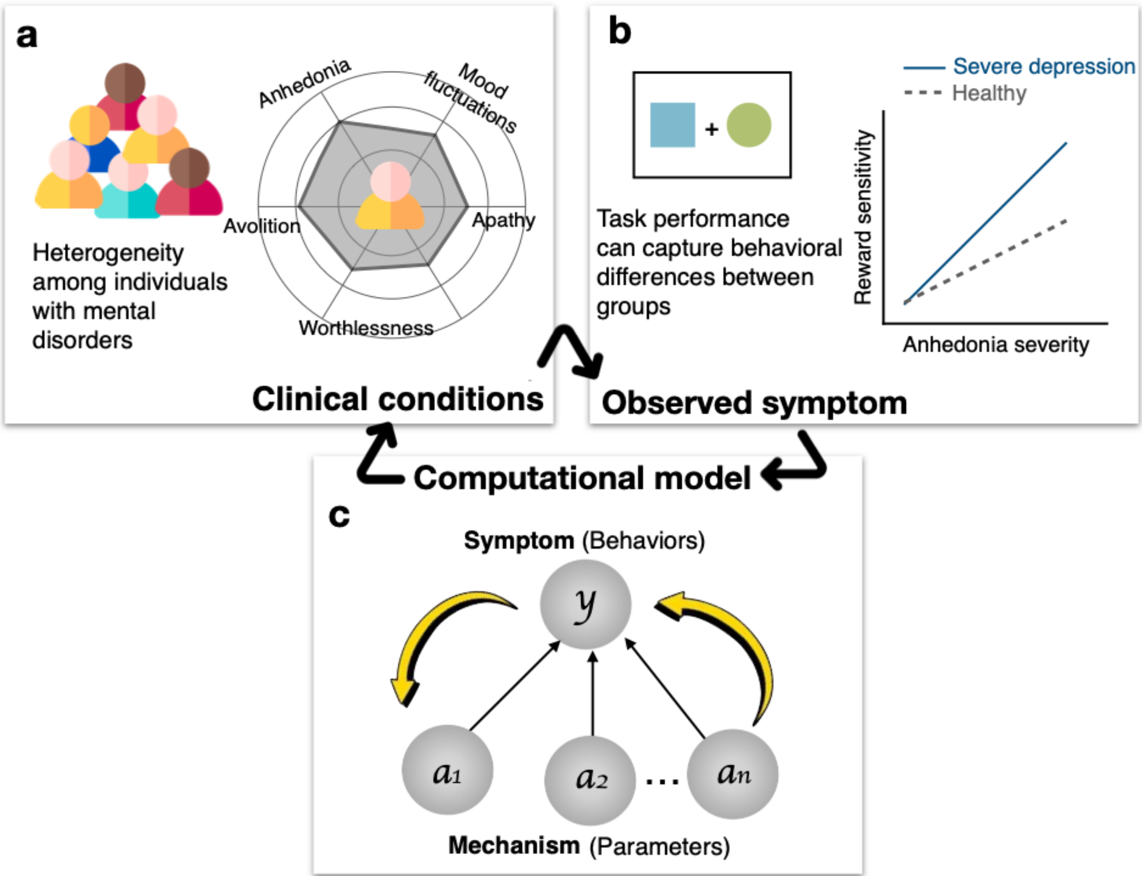

Fig.1: Overview of the reinforcement learning framework: Mental disorders have heterogeneous symptoms. For example, depression is characterized with symptoms such as anhedonia, apathy, and avolition (panel **a**). These symptoms can be manifested as behavioral observations in simple tasks. Hypothetical results of group difference between the healthy control and depressed individuals could differ (panel **b**). Reinforcement learning models can decompose the symptoms by mapping task behaviors to computational components (panel **c**).

Reinforcement learning decision-making

Reinforcement learning is a class of methods in which an agent (for example, a person, animal, or artificial system) learns to make decisions by interacting with its environment. Over time, the agent uses feedback in the form of rewards or penalties to adjust its behavior, optimizing performance based on past experiences. In the context of mental disorders, reinforcement learning provides a valuable framework for understanding psychiatric symptoms, allowing researchers to mechanistically model the maladaptive learning and decision-making process underlying its behavioral manifestations[12]. This approach has been widely adopted in the field of computational psychiatry, which applies mathematical models to elucidate the complex interplay between neural, cognitive, and behavioral processes underlying mental symptoms and disorders[4,8].

Reinforcement learning is governed by several key components: an agent, environment, states (external or internal), and policy[13]. Here, an agent interacts with an environment, which consists of states representing external or internal conditions. The agent selects actions based on a learned policy, which determines the optimal response to a given state. These actions are reinforced through outcomes that either increase (positively reinforce) or decrease (negatively reinforce) the likelihood of repeating certain behaviors. By trial and error, the agent continuously updates its value function, which estimates the expected reward associated with a given state, action, or state-action pair[14].

Building on these principles, reinforcement learning can be understood as a structured sequence of interconnected stages that guide an agent's decision-making process (**Fig.2)**. It begins with a representation of the given state or stimuli, in which the agent encodes both external and internal information to construct a mental model of the environment. Based on the representation, it proceeds to valuation, where subjective values are assigned to different states, actions, or state-action pairs. This informs the next stage, action selection, where the agent chooses an action based on the computed values and action policies. Following action selection, the agent would undergo outcome evaluation, a process where it assesses the consequences of its action by comparing expected and actual outcomes. This comparison often involves computing prediction errors, which signal discrepancies between anticipated and observed rewards. These errors guide the learning stage, updating value functions and refine future decision-making strategies. By continuously cycling through these stages, the agent adapts to its environment, improving its ability to make optimal choices over time.

This reinforcement learning decision-making framework described above can be directly implemented using computational models, allowing researchers to simulate and quantify how specific components of decision-making may be altered in mental disorders. As illustrated in **Fig.2**, this modeling process typically begins with reformulating clinical phenomena of interest into formal reinforcement learning and decision-making processes. A mathematical model is then constructed to represent the hypothesized decision process, followed by parameter estimation based on behavioral data-these parameters reflect latent cognitive mechanisms that shape behavior. The validity of the model can be evaluated through simulation and recovery procedures, which test whether the model can reproduce behavior consistent with observed data and recover known parameter values. By situating clinical symptoms within this structured computational pipeline, researchers can generate mechanistic hypotheses that are both testable and interpretable. In the following sections, we will go over each stage of the reinforcement learning decision-making framework, and outline how each process can be computationally formalized to understand dysfunction in each could lead to psychiatric symptoms.

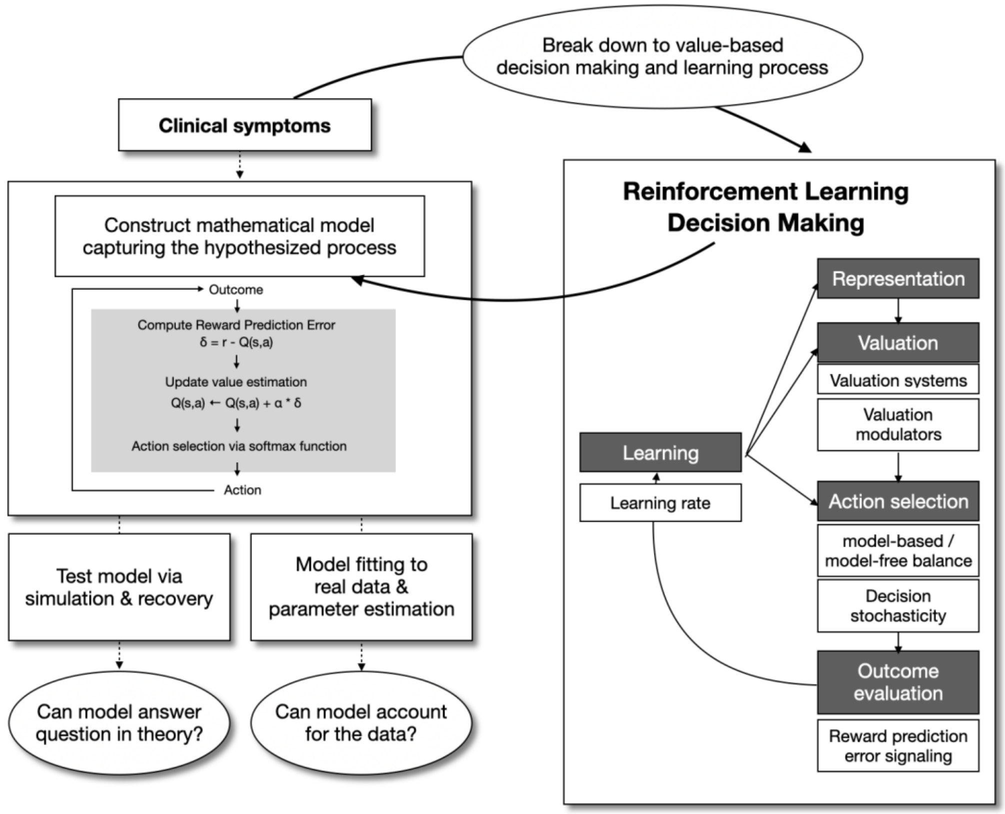

Fig.2: Computational modeling and reinforcement learning framework. Clinical symptoms can be decomposed into underlying cognitive processes using reinforcement learning models, through a typical computational modeling pipeline including model construction, simulation, and model fitting (**left** panel). Letters in the gray box denote: δ (reward prediction error), r (reward), Q(s,a) (action value of action *a* in state *s*), α (learning rate). These components correspond to key stages of the reinforcement learning decision-making framework, including key stages such as representation, valuation, action selection, outcome evaluation, and learning (**right** panel).

State representation

State representation involves constructing a mental state that encodes environmental features and identifies feasible options for the agent[15]. This process transforms external and internal stimuli into goal-relevant state and action spaces, forming a foundation for decision-making. In the reinforcement learning decision-making framework, probabilistic models formalize this process by estimating the agent's current state through observations and prior beliefs (for a detailed description, see ref.[16]). These models enable agents to refine their understanding of the environment by dynamically updating state representations.

Diverse cognitive constructs such as perception, memory, and attention play a central role in state representation. While perception processes sensory input into meaningful task-relevant states[17,18], declarative memory integrates past experiences, allowing agents to anticipate and guide future actions[19,20]. Attention further refines this process by prioritizing goal-relevant information while filtering distractions[21,22]. Computational models operationalize these constructs by mapping noisy sensory inputs and past experiences to probabilistic state estimates and formalizing how attentional biases adapt to emphasize goal-relevant features[23], allowing researchers to quantify how diverse aspects of cognition encode and process environmental states.

Valuation

Next, agents compute values for each state and feasible option by estimating expected returns or costs[1,12]. This valuation process can be affected by multiple control systems[24] : first, goal-directed valuation involves deliberate evaluation of actions based on their expected outcomes, supporting flexible and adaptive behavior. In contrast, habitual system reflects automatic responses shaped by past reinforcement history, often insensitive to outcome devaluation or contingency changes.

These control systems can be computationally formalized as model-free and model-based learning mechanisms. Model-free learning, encompassing habitual behaviors, updates values via prediction errors without maintaining an internal model of the environment[25,26]. This system prioritizes efficiency over flexibility, forming simple stimulus-response associations through reinforcement[27,28]. In contrast, model-based learning, which underlies the goal-directed control system, leverages an internal model to compute expected action values, enabling flexible, context-sensitive decision making[29–31].

Beyond these two instrumental systems, Pavlovian valuation system represents a distinct process in which stimuli are assigned innate value tendencies-such as approaching toward reward-predictive cues or avoidance of aversive stimuli, irrespective of instrumental contingencies[32] . While Pavlovian mechanisms are often viewed as evolutionarily hard-wired, they can interact with both model-based and model-free systems, biasing the valuation process.

These valuation systems interact with cognitive mechanisms such as working memory, which influences one’s capability to manage model-based strategies. Computational studies indicate that increased working memory loads impair model-based learning, suggesting its role in balancing model-based and model-free strategies[33–35]. Additionally, valuation is shaped by factors such as delay and effort, which modulate the subjective value of options, as well as by affective states like fear and anxiety, which bias decision making strategies. Mathematical models incorporate parameters such as a delay discounting rate and risk aversion to quantify these effects[36]. Overall, by formalizing valuation strategies and their modulators, reinforcement learning models provide a unified framework to quantify multiple components affecting the evaluation of stimulus and actions, considering their dynamic interactions.

Action selection

Action selection is the process of translating value computations into concrete decisions and behaviors, resolving conflicts between valuation systems[12]. In reinforcement learning context, this control can encompass orchestrating the switch between model-based and model-free approaches based on contextual demands. Hybrid reinforcement learning models capture this interplay by hierarchically integrating these processes and weighting parameters that quantify their relative contributions[37,38].

Beyond this model-based and model-free balance, action selection is also influenced by the Pavlovian system. Pavlovian biases refer to automatic tendencies to approach reward-predictive cues or avoid aversive ones, irrespective of their instrumental value. These biases can distort action selection by overriding goal-directed or habitual computations under conditions of conflict between systems[39,40] .

In parallel to the competition among multiple control systems, action selection introduces stochasticity and resolves the trade-off between exploration (seeking new information) and exploitation (leveraging known rewards)[29,41]. This trade-off is often quantified using the inverse temperature parameter, where lower values indicate greater exploration and higher values lead to deterministic exploitation of high-value options[42–44].

Sequential sampling models such as drift-diffusion models and race models offer most detailed insights into action selection by modeling its temporal dynamics[45,46]. These models assume that choices emerge from the gradual accumulation of evidence until a decision threshold is reached, providing a mechanistic account of choice accuracy and reaction time as well as the tradeoff between them. Integrating sequential sampling models into the reinforcement learning decision-making framework as a mechanism for translating accumulated evidence into final action selection enables a more mechanistic analysis of value-based deliberation and dynamic evidence accumulation[47–49].

Outcome evaluation and learning

After making a series of decisions, the brain generates value signals to evaluate outcomes and update strategies[12]. Reinforcement learning models capture this through reward prediction error, which refers to the discrepancy between expected and actual outcome. This prediction error drives learning and behavioral adjustment, rather than reward or loss itself. Dopamine neurons play a critical role in encoding this error[25,28,50], with positive prediction errors occurring when outcomes exceed expectations and negative prediction errors when they fall short[50]. Negative outcomes, such as aversive stimuli or losses, also elicit negative prediction errors, leading to avoidance behaviors and behavioral adjustments. These dopaminergic signals update value representations and action-outcome associations, reinforcing future behavior[51,52].

The magnitude of prediction errors also affects the learning rate, quantifying how strongly prediction errors influence subsequent valuation and action selection. A higher learning rate results in rapid adaptation to recent outcomes, while a lower rate reflects reliance on prior expectations, leading to more gradual updates. While classical reinforcement learning models typically assume a fixed learning rate, advanced models such as the Kalman filter and the Hierarchical Gaussian Filter allow for dynamic adjustment of learning rates in response to uncertainty[53–56] . This allows for more flexible and adaptive learning, particularly in volatile or nonstationary environments.

Beyond reward prediction errors, cognitive control and working memory further refine this learning process by integrating feedback and regulating the influence of prediction errors. Cognitive control processes monitor the decision-making process and allow for adjustments in strategy[57]. Working memory, in turn, complements this by modulating prediction error-based learning[58], shaping the extent to which prediction errors drive learning[59]. Together, these mechanisms ensure that learning is not solely dictated by immediate reinforcement but also incorporates broader cognitive resources, and using reinforcement learning decision-making framework, these multiple components can be mapped under a shared computational principle. This phase highlights the unifying role of the reinforcement learning decision-making framework, offering computational insights into how positive and negative outcomes, alongside cognitive resources, drive behavioral adaptation.

Interactions between and within domains

As previously mentioned, mental health problems rarely stem from isolated dysfunctions but rather involve interconnected disruptions across multiple systems. While previous frameworks such as Research Domain Criteria defined diverse constructs such as cognitive control and working memory, they lack a unified structure to formally incorporate these components into a mechanistic model of behavioral manifestations (**Box 1**). However, understanding the dynamic interplay among these constructs requires a structured approach that can formally model their interdependencies.

Here, the reinforcement learning decision-making framework can provide a computational approach to achieve this by offering a structured, quantitative analysis of how different mental processes influence each other within a shared decision-making mechanism. By modeling these interactions mathematically, this framework allows researchers to examine how interconnected disruptions contribute to psychiatric dysfunctions, complementing existing frameworks and providing insights into system-wide dynamics that underlie mental disorders. In this section, we address how reinforcement learning decision-making formalizes these interactions, restructuring various components into an integrated cycle of decision-making processes. By exemplifying cross-domain interactions such as the interplay between reinforcement learning and working memory, as well as empirical studies where reinforcement learning provided insights to comorbidity in multiple mental disorders, we discuss how reinforcement learning decision-making's interactive framework enhances our understanding of mental disorders.

Working memory and reinforcement learning

An example of integrating distinct cognitive processes within the reinforcement learning framework is the working memory and reinforcement learning model[60]. While working memory plays a crucial role in learning and decision-making by maintaining and updating task-relevant information, its interplay with reinforcement learning has rarely been studied in previous studies.

The reinforcement learning decision-making framework can provide a structured approach to show how these processes jointly shape behavior. A probabilistic reward learning task **(Fig.3a)**, where multiple options are associated with stochastic reward, and its variations are widely used to spot distinct decision strategies in maximizing the rewards. During the task, behaviors would be affected by the interplay between reinforcement learning (model-free learning driven by reward prediction errors) and working memory systems (capacity-limited but allows rapid updating of task-relevant information). Mathematical models can easily incorporate working memory construct with learning behavior during the task (**Fig.3b**). In the model, the probability of selecting action a is expressed as a weighted sum of working memory and reinforcement learning (**Fig.3c**). Here, the weighting parameter is adapted to determine how much decision-making is influenced by working memory, dynamically adjusting based on task demands. Under high cognitive load, this weighting parameter decreases, shifting reliance to reinforcement learning strategies. Under low cognitive load, the weighting parameter increases, promoting more model-based decisions.

This comprehensive framework offers insights into psychiatric dysfunctions where cognitive impairments interact with reinforcement learning deficits. Schizophrenia is one example: patients often exhibit impaired working memory but relatively intact reward-based reinforcement learning[61], a dissociation that reinforcement learning and working memory models can formulate by parameterizing the relative contribution of each system. Depression provides another compelling case, where anhedonia, a diminished ability to experience pleasure, may arise from impaired reward valuation and learning processes, which is further exacerbated by deficits in working memory[62]. Reinforcement learning models can disentangle these components, allowing researchers to precisely characterize how cognitive dysfunctions impact learning and decision-making, distinguishing between deficits in maintaining task-relevant information and impairments in value updating.

Beyond the interplay between the working memory system and value-based learning, reinforcement learning models have also been used to dissociate components within the domain of reward processing itself. For example, computational models have been used to dissociate anhedonia-related deficits in reward sensitivity from impairments in reward learning in depression[63]. While traditional descriptive frameworks cannot dissociate the role of distinct components under a single umbrella of reward processing, reinforcement learning models refine understanding of the role of multiple constructs by isolating the parameters capturing each construct and mechanistically linking them in a connected phase. This demonstrates how reinforcement learning helps understanding within-domain interactions using mechanistic approaches; details on these mechanistic approaches and related studies are explained in the following section.

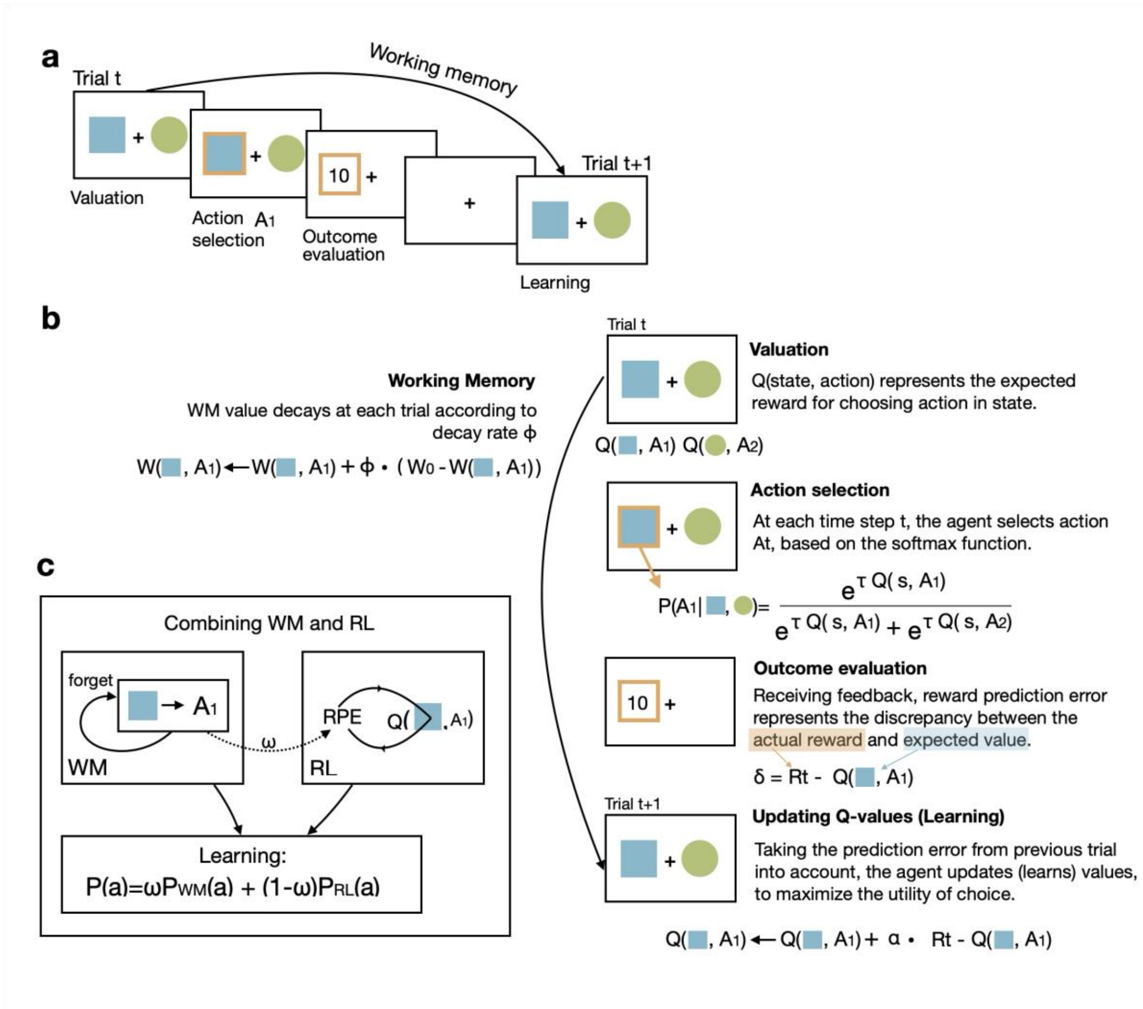

Fig.3: **Schematic of working memory-reinforcement learning model.**

In an exemplar experimental paradigm of a simple 2-armed bandit task, participants choose one stimulus on each trial t to maximize the rewards (panel **a**). On each stage, valuation, action selection, outcome evaluation, and learning occur. Each of the process can be formulated and incorporated in a computational model (panel **b**). A schematic model motivated by[60], depicts a theoretical model where working memory and reinforcement learning constructs are taken into account within a single model (panel **c**).

Understanding comorbidity and latent mechanisms

Reinforcement learning has also been useful in providing insights into comorbidity and shared mechanisms across disorders. For example, anxiety and depression frequently co-occur, with nearly half of individuals diagnosed with one having also received the diagnosis of the other[64]. Reinforcement learning models can simulate this comorbidity by modeling the interaction of dysfunctions across domains, rather than treating them as independent disorders.

A recent study by Zorowitz and colleagues demonstrated the utility of the reinforcement learning approach by showing that the uncertainty associated with potential threats and the pessimistic belief about oneself (inability to handle them) are key to anxiety and avoidance behaviors, which also contributes to depression. By adjusting key parameters using generative models, they simulated pessimistic expectations and maladaptive decision-making patterns observed in anxious individuals. The results supported the idea that clinical anxiety may precede certain types of depression, which, in turn, can lead to missed opportunities for rewards, eventually fostering the belief that rewards are unattainable, and ultimately contributing to the development of depression[65].

Similar modeling approaches have also been applied to understand compulsivity, a symptom shared by both addiction and obsessive-compulsive disorder[66]. Compulsivity is often viewed as an imbalance between goal-directed and habitual learning systems, marked by diminished goal-directed control and excessive habitual responses to immediate rewards or stimuli[67–69]. In substance use disorder, repeated drug intake leads to compulsive use, resulting in a loss of goal-directed control[70].

While both addiction and obsessive-compulsive disorder involve deficits in flexible decision-making, reinforcement learning tasks and Bayesian modeling reveal key differences. For instance, a perseverative behavior in the probabilistic reward learning task refers to a pattern in which the next choice is less affected by the current outcome, but rather is a repetition of a particular action (for example, keep choosing the stimulus presented on the left-hand side), even when it is no longer leading to rewarding outcome. This is often observed in addiction, in a way that positive outcome is less impactful, while in obsessive-compulsive disorder, both positive and negative outcome-driven learning remained intact[71]. Another study found that although both groups were less goal-directed compared to healthy individuals, obsessive-compulsive disorder was uniquely associated with an increased switching between model-free and model-based strategies, indicating more general impairment in action selection[72]. These findings highlight the distinct yet overlapping mechanisms underlying compulsivity in addiction and obsessive-compulsive disorder, emphasizing the role of reinforcement learning models in disentangling these processes.

Brain–Body interaction

Historically, reinforcement learning models have mainly focused on external states, such as reward-related cues and outcomes, but mental disorders often involve disruptions in interoceptive processing[73–75]. The reinforcement learning framework has been expanding its scope and could benefit from incorporating internal states, such as interoceptive signals, into decision-making models. Incorporating internal states into the framework could provide a more comprehensive understanding of how agents integrate bodily signals with external stimuli to drive adaptive decision-making and behavior.

This shift toward incorporating internal states aligns with the growing emphasis in psychology and neuroscience on understanding the dynamic interplay between the brain and body. Advances in this field have linked interoceptive processes to learning theories through dynamic models of bodily states. For instance, predictive coding—a Bayesian framework originally developed to explain how the brain interprets external cues—has been extended to interoception, shedding light on how the brain forms models of its internal environment as it changes over time[76–78]. In homeostatic reinforcement learning, the concept of reward has been redefined to emphasize actions that bring internal states closer to their optimal or desired levels as inherently rewarding. Interoceptive active inference builds on predictive coding by proposing that agents act to minimize prediction errors—or "surprise"—while maintaining homeostasis. This approach suggests that the brain combines information about the body's current state with environmental cues and learned associations to predict future internal states and guide proactive control[79]. Furthermore, a recent proposal has introduced the incorporation of internal states into the reinforcement learning framework, conceptualizing the body as an integral component of decision-making[80].

Although research in this area has been largely theoretical, recent empirical studies have begun to incorporate interoceptive dimensions into computational models. For example, Bayesian models have been applied to cardiac and respiratory data in healthy individuals[81,82], in anxiety-related contexts[83], and among psychiatric patients[75,84]. Notably, findings from Smith et al.[75] suggest that impaired interoceptive processing can perpetuate persistent somatic errors, contributing to emotional and somatovisceral dysregulation commonly observed in many mental disorders. These findings highlight the importance of incorporating internal states in the framework and illustrate how computational modeling can uncover hidden processes that are otherwise difficult to observe, enhancing our understanding of mental disorders.

Mechanistic understanding of mental disorders

Building on the capacity of reinforcement learning models to capture interactions across cognitive domains, we now turn to how this framework enhances a mechanistic insight of psychiatric dysfunction (**Fig.4**). By formally linking observed behaviors to underlying computational components, reinforcement learning enables researchers to identify where specific breakdowns occur in the decision-making process. In this section, we focus on empirical examples from anxiety, depression, and addiction–disorders that have been extensively studied using reinforcement learning frameworks and commonly characterized by significant dysfunctions in reward-related decision-making.

Depression and Reinforcement learning

Major depressive disorder is one of the leading causes of psychiatric conditions worldwide[85] . It is associated with a wide range of symptoms that affect emotional, cognitive, and behavioral functioning, with anhedonia being a hallmark feature. Anhedonia, which is characterized by a reduced capacity to experience pleasure or diminished interest[106,107] in typically rewarding stimuli, has been conceptualized as a dysfunction in reinforcement learning.

While learning is central to motivation and goal-directed behaviors (for a review, see ref.[1,86]), individuals with depression experience a disruption in the mechanisms that drive learning from rewarding experiences, leading to diminished motivation for daily activities.

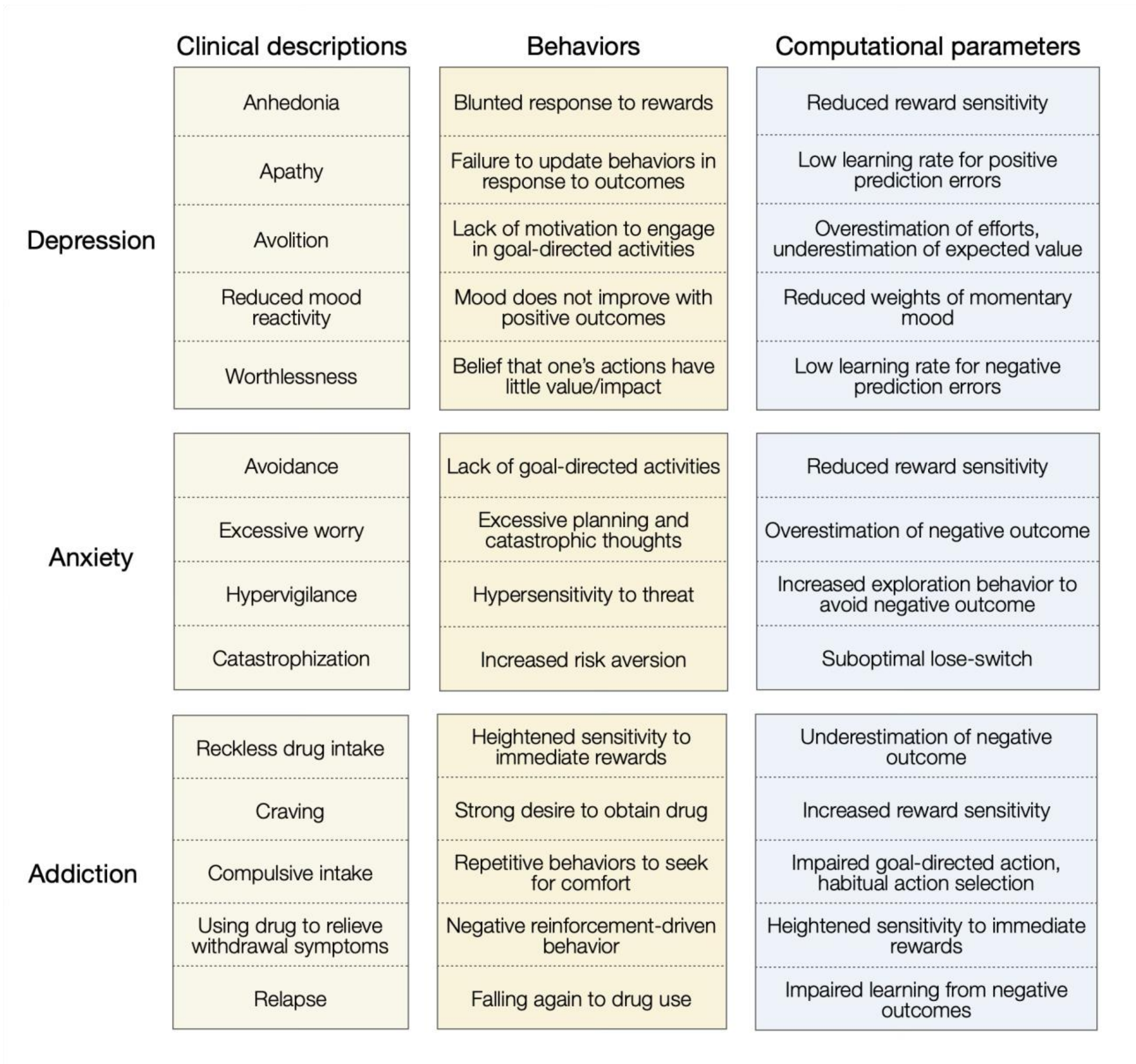

Fig.4: **Behavioral manifestations of symptoms within the framework.** Although there are still mixed or weak findings to establish a definitive mapping between every single clinical symptom and its computational mechanism, we present this hypothetical example to illustrate the potential of operationalizing clinical conditions through computational perspectives.

Empirical evidence consistently demonstrates that individuals with depression exhibit a diminished sensitivity to rewards[87–90], even among those with milder depressive symptoms[113]. Neuroimaging studies have reported reduced activity in key brain areas involved in reward processing[91] , such as the ventral striatum, during tasks designed to measure reward learning and outcome evaluation[92,93] . To mechanistically explain these findings, reinforcement learning models propose that depressive symptoms, including anhedonia, arise primarily from impairments in reward sensitivity and valuation processes[5,94].

Reinforcement learning models distinguish between reward sensitivity (the subjective value assigned to rewards) and learning rates (the speed at which an individual updates their belief based on feedback). For example, Mukherjee et al.[95] have shown that depressed individuals have impaired learning for rewards, but not for losses, compared to healthy controls. Specifically, they showed lower learning rates, which predicted depression more accurately and showed lower value sensitivity (differences in the subjective value of outcomes). Probabilistic reversal learning tasks, commonly used in reinforcement learning, provide a widely-accepted paradigm for investigating how individuals adapt their behavior in response to changing reward contingencies[96]. In a variant of this task, participants typically choose between two options, with the optimal choice associated with a higher probability of receiving a reward. However, this optimal response shifts unpredictably, and individuals are expected to update their behavior and beliefs based on feedback from their choice histories. Huys and colleagues[63] used this task to differentiate the neurocognitive components of reward sensitivity and learning rates. In their study, the reinforcement schedule of each stimulus varied unpredictably, and reward sensitivity was operationalized through a scaling parameter reflecting how strongly individuals valued the rewards they received. The learning rate quantified the extent to which individuals updated their expectations based on new feedback. Anhedonic symptoms were particularly predicted with the smaller reward sensitivity parameters, but not with the learning rates. A low dose of dopamine agonist, which stimulates dopamine receptors in the brain produced the opposite effects. These suggest that while depressed individuals may still be able to learn from feedback with intact dopamine signals, they are less responsive to the rewards they receive, which may explain the lack of motivation and persistent negative affect often observed in depression.

Furthermore, reinforcement learning models underscore the critical role of dopaminergic signaling in reward processing, particularly in the midbrain dopamine system that encodes reward prediction errors[97,98]. These prediction errors are essential for adaptive decision-making and in model-free learning, where positive feedback increases the likelihood for choosing the same action again, and negative feedback facilitates learning from errors. Depressed patients have shown blunted reward prediction error signaling in the striatum to positive outcomes[91,99], whereas this effect is less significant for the negative outcomes[86]. However, some studies report mixed findings, with mild depression with fewer episodes often exhibiting relatively intact dopaminergic signaling[100], suggesting that the severity of depression may influence the degree to which reward processing is impaired[101]. These inconsistencies suggest that this bias in depression may rest more on impaired valuation in the model-based manner, rather than trial-by-trial model-free learning that comes from dopaminergic responses[1] . Furthermore, Pessiglione et al.[102] suggested that multiple phenotypes can exist under the same clinical presentation of apathy, arising from various computational dysfunctions including reduced reward sensitivity, reward-effort tradeoff, and delay discounting for future rewards. Overall, this points to the complexity of reinforcement learning in depression and the need to further explore how different depressive subtypes might affect reward processing differently.

In line with this, depressed individuals frequently show asymmetries in learning, where they tend to undervalue potential rewards and overestimate associated efforts or costs. This is particularly evident in motivational learning, where depressed individuals struggle to integrate positive feedback effectively from past experiences. Such impairments limit subsequent exploratory behaviors, further restricting engagement with rewarding stimuli and actions[17,118]. These difficulties in learning from rewards are compounded by the overemphasis on negative experiences, which could bias future expectations and lead to increased learning from losses compared to gains. For example, discounting future rewards (preferring smaller immediate rewards over larger delayed ones) is typically more pronounced in individuals with depression, reflecting an increased focus on avoiding losses rather than pursuing potential gains[103].

Anxiety and Reinforcement learning

Anxiety is also primarily characterized by excessive fear, worry, and heightened threat perception[104,105]. These symptoms are linked to exaggerated emotional responses to potential threats, often resulting in hyper avoidant reactions. Reinforcement learning provides insight into these anxiety-related decision biases, mainly how value computation for rewards and punishments are distorted to affect action selection, along with altered learning rates in anxiety disorder (for a review, see ref.[106]).

First, distorted value computation plays a significant role in the decision-making process of individuals with anxiety. In particular, anxiety is associated with an overestimation of negative or threatening stimuli, where neutral or ambiguous cues are misinterpreted as dangerous. This bias is captured in reinforcement learning through an increased risk aversion parameter, which leads to a preference for low-risk, low-reward options[107]. A prominent prospect theory[108] of modeling individual risk sensitivity showed that individuals with anxiety are more sensitive to risk in monetary decision-making compared to healthy controls. Essentially, heightened anxiety results in a negative interpretational bias, leading to undervaluation of uncertain outcomes, eventually making individuals withdraw from exploration to avoid potential harms[109] .

The second cognitive mechanism contributing to avoidance behavior in anxiety disorders is maladaptive action selection. Individuals with a high level of anxiety frequently show impairments in model-based reinforcement learning, where their behavior shifts toward model-free strategies[110]. In model-based learning, individuals update their expectations and choices using a structured, goal-directed approach, based on their internal model. However, anxiety often involves excessive simulations and the inability to settle on optimal solutions may hinder the ability to effectively use model-based strategies and settle on optimal solutions[111]. Moreover, Pavlovian conditioning in anxiety disorders leads to strong associations between certain cues and negative outcomes, reinforcing maladaptive avoidance behaviors. Pavlovian biases toward avoidance often override instrumental learning, leading to excessive harm avoidance in volatile environments[112]. Approach-avoidance task is a widely used task paradigm that incorporates threat manipulation to a probabilistic go/no-go reinforcement learning task[112] . In this task, participants are given stimuli options to involve in either approaching or avoiding each stimulus that could be either rewarding or punishing. Using this task, Mkertchian et al.[112] demonstrated an increased tendency for avoidance in action selection (Pavlovian bias to withhold action in the face of punishment), an effect that was exacerbated under stressful conditions specifically in the anxious group.

The third potential mechanism is altered learning for aversive outcomes. Anxiety increases sensitivity to negative feedback, leading to an elevated learning rate for negative prediction errors. In other words, individuals with higher levels of anxiety tend to update their expectations more drastically in response to losses or failures, as opposed to rewards or successes. Indeed, computational studies suggest that anxious individuals exhibit increased weight on negative prediction errors, which could contribute to persistent maladaptive avoidance patterns[113,114]. They often learn more quickly in the punishment domain compared to the reward domain, favoring avoidance behaviors. However, different subtypes of anxiety disorders have mixed findings, where in some anxiety disorders, individuals learn more slowly for both positive and negative reinforcers. For example, socially anxious individuals struggle to update beliefs about social threats, which would lead to persistent fears even when contexts no longer pose danger[115]. In generalized anxiety disorder, individuals showed a slower learning rate over time, regardless of the reinforcement valence[116]. Similarly, individuals who experienced a greater level of task-induced anxiety showed a higher learning rate to punishments relative to rewards[109], which has been replicated with a larger sample[117].

In some cases, a heightened learning rate for negative outcomes may lead to overly rapid adjustments in behavior, manifesting as win-stay, lose-shift–like patterns. Huang and colleagues[118] used a behavioral task to investigate a symptom of anxious individuals' blunted sensitivity in identifying a goal from distractors, remaining hypervigilant to them. Although there was no significant difference across individuals with different anxiety levels in terms of overall performance (how much an individual earned reward during a task), they were able to show that individuals with a greater anxiety level displayed a suboptimal decision strategy, characterized by an excessive lose-shift behavior. This may underlie maladaptive avoidance in anxiety, where individuals are over-sensitive to potential negative outcomes, even when it is no longer adaptive.

Addiction and Reinforcement learning

Addiction, especially substance use disorder, is marked by persistent drug-seeking behavior and an inability to regulate substance use despite negative consequences[119]. While traditionally characterized in behavioral and neurobiological terms, addiction-related dysfunctions have been further dissected mechanistically using reinforcement learning decision-making framework. By decomposing multifaceted symptoms such as craving, withdrawal symptoms, and compulsivity into quantifiable cognitive processes, this approach provides a precise quantification of the latent processes driving addiction-related abnormalities and mechanistic insights beyond descriptive models.

At its core, addiction stems from dysfunctions in the dopamine-mediated reward system, disrupting how individuals process rewards and control their behavior[97]. Within the reinforcement learning framework, dopamine has been strongly linked to the encoding of reward prediction errors, which serve as teaching signals for updating value estimates and refining future behavior[25]. In addiction, this mechanism becomes 'hijacked': pharmacological effects of drugs directly stimulate dopaminergic activity, producing artificially large prediction error signals and causing drug-paired cues to be overvalued and continuously reinforced[120]. This explanation provides a mechanistic view for the exaggerated salience of drug-related cues, offering a computational basis of how value systems become distorted in addiction.

This aberrant reward valuation is further manifested as a phenomenon known as reward narrowing, where drug-related cues become overvalued while natural rewards are devalued[121]. This bias arises from altered dopaminergic signaling, driven by the pharmacological effects of drugs and subsequent neuroadaptive changes, affecting how subjective value is assigned to different stimuli. Computational modeling studies have shown that individuals with substance use disorder exhibit biased value representations, with altered sensitivity to reward and punishment cues and their associated contexts[122–124]. Furthermore, studies originally rooted in behavioral economics-such as delay discounting paradigms-have consistently shown that individuals with addiction exhibit a stronger preference for immediate rewards over long-term gains. Elevated delay discounting has been consistently observed across various substances, including alcohol, cocaine, opioids, cannabis, and nicotine- suggesting this stiff discounting may constitute a core computational risk marker underlying impulsive choice behavior in addiction. In parallel, addicted individuals tend to be more risk-tolerant and less loss-averse than healthy controls, which further bias decision-making strategies. This can modulate utility functions and weighting of prediction errors, dampening sensitivity to negative outcomes during value updating.

Beyond the biases in the valuation phase, addiction disrupts learning and action selection mechanisms by altering the balance between multiple controllers. Compulsive drug-seeking behavior in addiction has long been thought to result from a shift away from goal-directed control and toward habitual behavior, which can be translated to the alteration of model-based and model-free systems' balance in a computational perspective. Computational models suggest that chronic drug use weakens model-based learning and decision-making, inducing imbalance between the two systems[125,126]. This imbalance is quantified using a weighting parameter, which reflects the weighting of model-based versus model-free processes. Computational analyses indicate that addiction is primarily associated with reduced model-based reinforcement learning rather than overreliance on model-free learning. While addiction is commonly associated with a smaller weighting parameter, reflecting impaired goal-directed control[38,127], findings are mixed, and some evidence suggests excessive model-based valuation of drugs to relieve withdrawal symptoms[128]. Reinforcement learning provides a mechanistic framework to disentangle these competing hypotheses by precisely quantifying individual differences in reward valuation, learning strategies, and action selection biases.

Based on this systematic framework, empirical studies demonstrate that computational modeling can offer insights on what traditional descriptive analyses often overlook. Specifically, by mechanistically decomposing the multidimensional drivers of addiction, reinforcement learning framework allows researchers to identify subtle and latent characteristics underlying behavioral symptoms of addiction and further classify the subtypes based on computationally defined features. For instance, Myers and colleagues[129] demonstrated that while descriptive behavioral measures failed to detect significant differences, computational modeling identified an elevated recency bias in opioid-addicted individuals, meaning they are more likely to repeat previous responses regardless of reinforcement history. Similarly, a recent study[130] found that heroin users exhibited reduced learning rate for punishment, while chronic cannabis users showed greater learning rate for reward, illustrating subtle individual differences in addiction phenotypes that would be difficult to capture through conventional analyses. These findings highlight the heterogeneity within addiction, emphasizing the potential of reinforcement learning-based computational phenotyping.

By leveraging reinforcement learning principles, computational approaches go beyond traditional symptom-based classifications, offering quantifiable, mechanistic insights to understand the neurocomputational basis of addiction. By decomposing addiction-related dysfunctions into precise computational components (reward valuation distortions, learning impairments, and shifts in decision control), reinforcement learning framework enables a more fine-grained analysis of why individuals develop compulsive drug use and how these dysfunctions vary across addiction phenotypes.

Beyond mechanistic understanding

Reinforcement learning's ability to provide mechanistic insights extends beyond theoretical models—it has practical implications for the assessment (**Box 2**) and treatment of mental disorders (for a review, see ref.[8]). Computational phenotyping or fingerprinting, which dissects complex psychiatric symptoms into measurable cognitive components and characterizes individuals based on unique reinforcement learning-derived parameter profiles[122,131], has gained significant interest in psychiatric research. Unlike static trait measures, parameters that are estimated from computational models can capture dynamic behaviors throughout learning (for example, adapting to the environment), thus providing rich insights for computational phenotyping by capturing latent mechanisms that shape observable symptoms. By integrating parameters across tasks, this approach systematically maps individuals onto a multidimensional parameter space, refining the mechanistic understanding of psychiatric heterogeneity. Yechiam et al.[132] used a single reinforcement learning task to distinguish maladaptive decision-making across mental disorders to systematically explain the similarities and differences between them, quantifying their characteristics in a more mechanistic manner. Similarly, Gueguen et al.[122] proposed a multidimensional framework for computational fingerprinting in addiction, demonstrating how reinforcement learning-derived features can uniquely profile individuals, even in the context of shared symptoms. A recent paper further suggested that dynamic computational phenotyping can be refined by integrating state-dependent variability, highlighting its potential on characterizing individual variability[133,134].

Reinforcement learning also contributes to methodology. Rooted in reinforcement learning models, adaptive design optimization dynamically adjusts experimental variables in real-time to guide decision-making. Unlike traditional designs that follow a predetermined fixed structure, adaptive design optimization algorithm adapts over time to optimize task performance, parameter estimation, and model fitting. This contributes to a challenge in the field, which is the poor estimation of individual parameters, particularly when trial numbers are insufficient or when tasks fail to elicit stable behavior patterns[134,135]. Adaptive design optimization[136–138] helps to estimate parameters with a small number of trials, improving efficiency in data collection. By continuously adjusting the design based on real-time participant choices, it maximizes information from each trial, reducing the need for a large number of trials. For example, Ahn et al.[139] demonstrated that the adaptive design optimization led to high test-retest reliability (r = 0.95) of delay discounting parameter within 10-20 trials. Using a similar paradigm, Lee et al. [140] demonstrated computational markers predictive of addictive behaviors and treatment outcome. Furthermore, leveraging high-dimensional datasets incorporating multimodal data would allow researchers to capture a more comprehensive view of psychological constructs, offering the opportunity to enhance the robustness of computational models. Altogether, these systematic efforts ensure that computational approach remains grounded in reproducible and clinically relevant measures, bridging the gap between theoretical modeling and practical application. This not only increases participant compliance but also improves data quality, which is especially valuable in clinical studies where maintaining engagement can be challenging.

Summary and future directions

The reinforcement learning decision-making framework provides a powerful computational approach that decomposes complex psychiatric dysfunctions into mechanistic, quantifiable components. Expanding on traditional descriptive frameworks, reinforcement learning models enable a structured, hypothesis-driven investigation of how latent cognitive and psychological processes drive maladaptive decision-making. By modeling key decision-making mechanisms—including state representation, reward valuation, action selection, and learning—this framework allows researchers to systematically examine how disruptions in these processes contribute to psychiatric conditions.

A major strength of this framework is its use of computational models to foster mechanistic understanding of the heterogeneous and complex nature of mental disorders. Specifically, we showed how reinforcement learning can contribute to understanding and investigating interactions between different cognitive and psychological aspects of mental disorders. Moreover, reinforcement learning leverages generative models to infer latent cognitive processes, offering mechanistic insights that go beyond traditional descriptive methods. Researchers can come up with different hypotheses of which components of learning (for example, reward prediction error and learning rate) can contribute to distinguishing overlapping symptoms by identifying their distinct computational signatures. This structured approach bridges the gap between behavior, neural circuits, and clinical symptoms, refining psychiatric research and advancing personalized treatment strategies.

Application of the reinforcement learning framework in understanding mental disorders is an emerging field that still includes several challenges. Here we suggest some key directions for future research. Despite advancements in reinforcement learning frameworks, most approaches using reinforcement learning models still rely heavily on controlled experimental paradigms, which may not fully capture the complexity of real-world decision-making. Traditional reinforcement learning tasks often simplify the environment and overlook the dynamic, context-dependent nature of decision-making in a naturalistic setting, potentially limiting the ecological validity of the findings. To bridge this gap, future research would greatly benefit from integrating reinforcement learning with naturalistic paradigms, such as movie-watching[141–143] and real-time tasks using games or virtual environments[144–146], have shown promise in capturing complex decision-making behaviors in ecological contexts[147].

Advanced computational algorithms and artificial intelligence, like deep neural networks[148], effectively model dynamic, real-time decision-making, while handling complexity of naturalistic settings is crucial[149]. For instance, Cross et al[150] used these methods to map task representations in the brain during video game play. However, these models often assume optimal performance, overlooking human participants' limited capacity. A recent work by Lee et al.[146] applied inverse reinforcement learning to model real-time driving behaviors and address this by inferring each individual's latent reward functions and capture individual differences in a real-time decision-making task.

Beyond the naturalistic task setting, another promising direction for advancing reinforcement learning research is the study of naturalistic rewards. Traditional reinforcement learning tasks often rely on abstract, task-based reinforcers-most commonly monetary rewards-to model learning and decision-making processes across diverse populations. While such approaches are useful for capturing neurocognitive parameters in controlled and generalized settings, they fail to fully account for how real-world, biologically relevant reinforcers shape reinforcement learning mechanisms.

This issue is particularly crucial in the context of human addiction research, where the neurocognitive mechanisms underlying reinforcement learning are directly tied to substance-specific rewards. Previous studies have demonstrated that chronic drug use blunts dopamine responses to conventional rewards, such as monetary incentives, while preserving or even enhancing responses to cues associated with the addictive substance[151,152]. This suggests that addicted individuals exhibit selectively biased reward processing, prioritizing drug-related reinforcers over other types of rewards. Standard reinforcement learning models that rely on generic task-based reinforcers may overlook these critical alterations in value representation, making it essential to incorporate naturalistic rewards into computational paradigms. Reinforcement learning models that explicitly account for naturalistic rewards will provide a more ecologically valid framework for understanding how addicted individuals respond to substance-specific cues, update reward expectations, and make decisions under conditions of craving or withdrawal.

Finally, future research could expand on the reinforcement learning framework by integrating it with real-life, large-scale behavioral data along with machine learning approaches. The emergence of mobile health tools, passive sensing technologies, and digital phenotyping provides a unique opportunity to apply the reinforcement learning framework to longitudinal, real-world datasets. For instance, reinforcement learning models can be adapted to analyze how individuals engage in real-time decision-making in daily life, tracking patterns of reward sensitivity, habit formation, and cognitive control fluctuations over extended periods. Moreover, integrating the reinforcement learning framework with large-scale datasets or with neuroimaging techniques could contribute to development of neurocognitive biomarkers in mental disorders.

Box 1: Expanding on the other frameworks

The reinforcement learning decision-making framework can complement and expand on previous frameworks such as the Research Domain Criteria. The Research Domain Criteria was introduced by the National Institute of Mental Health as a transdiagnostic framework to provide a comprehensive explanation of mental disorders by breaking down fundamental psychological and neurobiological constructs[153,154]. It has served as a multidimensional approach by focusing on behavioral processes, brain circuits, and biological mechanisms, while spanning various levels of analysis, from brain systems to behavior[155,156]. As an “organizational heuristic”, the Research Domain Criteria has offered a broad outline for mental health research through a two-dimensional matrix.

On the first axis, it defines six major psychobiological domains: Negative Valence systems, Positive Valence systems, Cognitive systems, Social processes, Arousal/Regulatory systems, and Sensorimotor systems. Another axis of the framework consists of units of analysis, such as genes, molecules, cells, circuits, physiology, behavior, and self-reports[157,158]. The Negative Valence systems involve brain processes related to negative emotions, such as fear, anxiety, and loss, while the Positive Valence systems deal with processes related to positive emotions, such as reward, motivation, and pleasure. This includes how the brain responds to enjoyable activities or goals. The Cognitive systems encompass mental processes like attention, memory, and decision-making in general, and the Social Processes include understanding development, sharing social cues, and forming relationships. The Arousal/Regulatory systems and the Sensorimotor systems are related to bodily reactions to external cues, how we cope with external stress, and how we perceive and respond to the world through our senses (for example, sight, sound, or touch) and movement.

While computational approaches[4,8,159] have previously received considerable attention as a method for investigating mental disorders, there has been limited focus on how computational approaches can interface with, complement, and enhance the Research Domain

Criteria framework. The reinforcement learning and decision-making framework not only covers many of the advantages that the previous framework has provided, but also expands on it. For example, the domains and units of analysis introduced in the previous framework are highly compatible with the reinforcement learning decision-making framework, as its core components—such as reward prediction, value computation, and policy selection—align closely with the underlying mechanisms in Positive Valence and Cognitive Systems.

As discussed more in depth in the main text, reinforcement learning models can account for interactions between domains and constructs within an integrated framework[153,156]. For instance, while reward processing deficits in the Positive Valence Systems and heightened threat perception in the Negative Valence Systems can each contribute to maladaptive behaviors, their interaction may drive distinct outcomes[160]. Assuming symptoms arise within a single domain is overly reductive and oversimplifies the mechanisms of mental disorders, ultimately failing to capture the full scope of mental health conditions[161,162]. Understanding these interactions is crucial for advancing psychiatric research and treatment.

Box 2: Psychometric advantages of computational modeling

Reinforcement learning models can be used to enhance the psychometric properties. Studies have shown that computational approaches improve test-retest reliability and provide more robust and precise computational markers compared to traditional psychological constructs and behavioral measures.

Cognitive tasks, defined as experimental paradigms measuring behavior to derive inferences about specific cognitive processes, play a central role in most of the research frameworks, including the reinforcement learning framework[65]. However, despite their utility, challenges remain in ensuring that the psychometric properties of these measures are reliable. The reliability of task measures has long been criticized due to their low reliability in comparison to conventionally established minimum acceptable thresholds and levels consistently attained by self-report instruments[163–165]. This issue is often referred to as the reliability paradox, where tasks designed to produce highly reliable effects at the group level often reduce individual variability, making them less effective at capturing meaningful individual differences[163,166].

This limitation is tied to the reliance on traditional descriptive summary statistics in task data analysis, which often obscure individual variability and the latent mechanisms underlying behavioral responses[167]. As discussed earlier, relying on simple metrics, such as overall accuracy or response time, may overlook confounding factors that drive these behaviors. This underscores the need for more sophisticated analytical techniques that can capture latent cognitive processes and account for subtle individual differences.

Computational approaches, including reinforcement learning, addresses these issues by providing more reliable and precise measures than traditional descriptive metrics. Research indicates that reinforcement learning models can capture the detailed dynamics of decision-making processes, offering more stable and reliable estimates of cognitive traits compared to conventional behavioral measures[168]. For example, recent studies have shown that generative modeling can reliably capture stable latent processes and achieve improved reliability compared over traditional behavioral measures[169,170]. These findings illustrate how the theoretical advantages of computational approaches can lead to practical benefits, such as enhanced reliability in experimental measures.

Computational modeling methods such as hierarchical Bayesian modeling can further help refine parameter estimation by incorporating group-level priors, reducing noise and enhancing stability[170–172]. Compared to standard estimation methods, this strategy provides more consistent parameter recovery in reinforcement learning and decision-making models, making it a valuable tool for improving psychometric robustness[166,173]. Beyond estimation, the parameter recoverability also influences reliability (for a review, see ref.[134]). These considerations ensure that the psychometric advantages of computational approaches are robust, relying on proper model selection, estimation, and validation.

By improving the reliability of task-based measures via computational approaches, this can enhance the discriminatory power of task measures and ultimately improve the accuracy of classification[174]. In a study by Huys et al.[8], simulated behavioral data were first generated using a simple model-free reinforcement learning algorithm, and then fit back to the same model to recover underlying parameters. Classifiers trained on these recovered parameters achieved higher classification accuracy (Area Under Curve = 0.87) than those trained directly on the simulated choice data (Area Under Curve = 0.74), demonstrating that model-derived features can capture clinically relevant information more effectively than raw behavioral outputs.

These advancements in psychometric properties could directly impact clinical practice, offering the potential to improve the precision of diagnosis and treatment. By dynamically quantifying individual differences in cognitive and neural processes, computational approaches such as reinforcement learning could potentially enable personalized treatment monitoring and refinement[175,176]. Empirical studies have shown that reinforcement learning-based parameters, including learning rates and reward prediction error signals, can predict treatment outcomes and guide intervention strategies[177]. For instance, Queirazza et al.[178] demonstrated that striatal and amygdala responses to reward prediction errors distinguished treatment responders in depression. Additionally, computational psychopharmacology provides a framework for mechanism-based treatment approaches, optimizing therapeutic interventions through quantitative modeling[69].

Acknowledgements:
This work was supported by the National Research Foundation of Korea (NRF) (RS-2024-00420674), the Creative-Pioneering Researchers Program through Seoul National University and the National Institute on Drug Abuse of the National Institutes of Health (Grant 5R01DA058038) to W.-Y.A..

References

1. Huys, Q. J. M., Daw, N. D. & Dayan, P. Depression: A Decision-Theoretic Analysis. *Annu. Rev. Neurosci.* **38**, 1–23 (2015).

2. Zimmerman, M., Ellison, W., Young, D., Chelminski, I. & Dalrymple, K. How many different ways do patients meet the diagnostic criteria for major depressive disorder? *Compr. Psychiatry* **56**, 29–34 (2015).

3. Maia, T. V. & Frank, M. J. From reinforcement learning models to psychiatric and neurological disorders. *Nat. Neurosci.* **14**, 154–162 (2011).

4. Montague, P. R., Dolan, R. J., Friston, K. J. & Dayan, P. Computational psychiatry. *Trends Cogn. Sci.* **16**, 72–80 (2012).

5. Huys, Q. J. M., Browning, M., Paulus, M. P. & Frank, M. J. Advances in the computational understanding of mental illness. *Neuropsychopharmacology* **46**, 3–19 (2021).

6. Adams, R. A., Huys, Q. J. M. & Roiser, J. P. Computational Psychiatry: towards a mathematically informed understanding of mental illness. *J. Neurol., Neurosurg. Psychiatry* **87**, 53 (2016).

7. Liebenow, B. *et al.* Computational reinforcement learning, reward (and punishment), and dopamine in psychiatric disorders. *Front. Psychiatry* **13**, 886297 (2022).

8. Huys, Q. J. M., Maia, T. V. & Frank, M. J. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat Neurosci* **19**, 404–413 (2016).

9. Kishida, K. T., King-Casas, B. & Montague, P. R. Neuroeconomic Approaches to Mental Disorders. *Neuron* **67**, 543–554 (2010).

10. Stephan, K. E. & Mathys, C. Computational approaches to psychiatry. *Curr. Opin. Neurobiol.* **25**, 85–92 (2014).

11. Wang, X.-J. & Krystal, J. H. Computational Psychiatry. *Neuron* **84**, 638–654 (2014).

12. Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci* **9**, 545–556 (2008).

13. Sutton, R. S. & Barto, A. G. Reinforcement_learning.pdf. *Journal of Cognitive Neuroscience* 126–134 (1999).

14. Barto. Intrinsically Motivated Learning in Natural and Artificial Systems. (2013) doi:10.1007/978-3-642-32375-1.

15. O'Doherty, J. P., Cockburn, J. & Pauli, W. M. Learning, Reward, and Decision Making. *Annu. Rev. Psychol.* **68**, 73–100 (2016).

16. Schuck, N. W., Wilson, R. & Niv, Y. A State Representation for Reinforcement Learning and Decision-Making in the Orbitofrontal Cortex. *bioRxiv* 210591 (2018) doi:10.1101/210591.

17. McCallum, A. K. Reinforcement learning with selective perception and hidden state. (1996).

18. Pomerantz, J. R. Encyclopedia of Cognitive Science. (2013) doi:10.1002/0470018860.s00589.

19. Schuck, N. W., Cai, M. B., Wilson, R. C. & Niv, Y. Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron* **91**, 1402–1412 (2016).

20. Gershman, S. J. & Daw, N. D. Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annu. Rev. Psychol.* **68**, 1–28 (2015).

21. Wilson, R. C. & Niv, Y. Inferring Relevance in a Changing World. *Front. Hum. Neurosci.* **5**, 189 (2012).

22. Radulescu, A. & Niv, Y. State representation in mental illness. *Curr. Opin. Neurobiol.* **55**, 160–166 (2019).

23. Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V. & Niv, Y. Dynamic Interaction between Reinforcement Learning and Attention in Multidimensional Environments. *Neuron* **93**, 451–463 (2017).

24. Daw, N. D. & O'Doherty, J. P. Neuroeconomics (Second Edition). *Part IV: Neural Mech. Choice* 393–410 (2014) doi:10.1016/b978-0-12-416008-8.00021-8.

25. Schultz, W., Dayan, P. & Montague, P. R. A Neural Substrate of Prediction and Reward. *Science* **275**, 1593–1599 (1997).

26. Balleine, B. W. & O'Doherty, J. P. Human and Rodent Homologies in Action Control: Corticostriatal Determinants of Goal-Directed and Habitual Action. *Neuropsychopharmacol* **35**, 48–69 (2010).

27. Houk, J. C., Adams, J. L. & Barto, A. G. A Model of How the Basal Ganglia Generate and Use Neural Signals That Predict Reinforcement. in *Models of Information Processing in the Basal Ganglia* (1994).

28. Montague, P., Dayan, P. & Sejnowski, T. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* **16**, 1936–1947 (1996).

29. Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).

30. Dolan, R. J. & Dayan, P. Goals and Habits in the Brain. *Neuron* **80**, 312–325 (2013).

31. Balleine, B. W. & Dickinson, A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* **37**, 407–419 (1998).

32. Clark, J. J., Hollon, N. G. & Phillips, P. E. Pavlovian valuation systems in learning and decision making. *Curr. Opin. Neurobiol.* **22**, 1054–1061 (2012).

33. Collins, A. G. E. & Frank, M. J. How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *Eur. J. Neurosci.* **35**, 1024–1035 (2012).

34. Zuo, Z., Yang, L.-Z., Wang, H. & Li, H. Working Memory Guides Action Valuation in Model-based Decision-making Strategy. *J. Cogn. Neurosci.* **37**, 86–96 (2025).

35. Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A. & Daw, N. D. Working-memory capacity protects model-based learning from stress. *Proc. Natl. Acad. Sci.* **110**, 20941–20946 (2013).

36. Peters, J. & Büchel, C. The neural mechanisms of inter-temporal decision-making: understanding variability. *Trends Cogn. Sci.* **15**, 227–239 (2011).

37. Gläscher, J., Daw, N., Dayan, P. & O'Doherty, J. P. States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning. *Neuron* **66**, 585–595 (2010).

38. Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P. & Dolan, R. J. Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron* **69**, 1204–1215 (2011).

39. Talmi, D., Seymour, B., Dayan, P. & Dolan, R. J. Human Pavlovian–Instrumental Transfer. *J. Neurosci.* **28**, 360–368 (2008).

40. Garbusow, M. *et al.* Pavlovian-to-Instrumental Transfer across Mental Disorders: A Review. *Neuropsychobiology* **81**, 418–437 (2022).

41. Cohen, J. D., McClure, S. M. & Yu, A. J. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos. Trans. R. Soc. B: Biol. Sci.* **362**, 933–942 (2007).

42. Ishii, S., Yoshida, W. & Yoshimoto, J. Control of exploitation–exploration meta-parameter in reinforcement learning. *Neural Netw.* **15**, 665–687 (2002).

43. Doya, K. Modulators of decision making. *Nat. Neurosci.* **11**, 410–416 (2008).

44. Luce, R. D. *Individual Choice Behavior : A Theoretical Analysis*. (Greenwood Press, 1979).

45. Bogacz, R., Brown, E., Moehlis, J., Holmes, P. & Cohen, J. D. The Physics of Optimal Decision Making: A Formal Analysis of Models of Performance in Two-Alternative Forced-Choice Tasks. *Psychol. Rev.* **113**, 700–765 (2006).

46. Ratcliff, R. & McKoon, G. The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Comput.* **20**, 873–922 (2008).

47. Pedersen, M. L., Frank, M. J. & Biele, G. The drift diffusion model as the choice rule in reinforcement learning. *Psychon. Bull. Rev.* **24**, 1234–1251 (2017).

48. Miletić, S., Boag, R. J. & Forstmann, B. U. Mutual benefits: Combining reinforcement learning with sequential sampling models. *Neuropsychologia* **136**, 107261 (2020).

49. Fontanesi, L., Gluth, S., Spektor, M. S. & Rieskamp, J. A reinforcement learning diffusion decision model for value-based decisions. *Psychon. Bull. Rev.* **26**, 1099–1121 (2019).

50. O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H. & Dolan, R. J. Temporal Difference Models and Reward-Related Learning in the Human Brain. *Neuron* **38**, 329–337 (2003).

51. Sutton, R. S. Learning to predict by the methods of temporal differences. *Mach. Learn.* **3**, 9–44 (1988).

52. Batten, S. R. *et al.* Dopamine and serotonin in human substantia nigra track social context and value signals during economic exchange. *Nat. Hum. Behav.* **8**, 718–728 (2024).

53. Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).

54. Nassar, M. R., Wilson, R. C., Heasly, B. & Gold, J. I. An Approximately Bayesian Delta-Rule Model Explains the Dynamics of Belief Updating in a Changing Environment. *J. Neurosci.* **30**, 12366–12378 (2010).

55. Mathys, C., Daunizeau, J., Friston, K. J. & Stephan, K. E. A Bayesian Foundation for Individual Learning Under Uncertainty. *Front. Hum. Neurosci.* **5**, 39 (2011).

56. Gershman, S. J. Dopamine, Inference, and Uncertainty. *Neural Comput.* **29**, 3311–3326 (2017).

57. Frömer, R. & Shenhav, A. Filling the gaps: Cognitive control as a critical lens for understanding mechanisms of value-based decision-making. *Neurosci. Biobehav. Rev.* **134**, 104483 (2022).

58. Yoo, A. H. & Collins, A. G. E. How Working Memory and Reinforcement Learning Are Intertwined: A Cognitive, Neural, and Computational Perspective. *J. Cogn. Neurosci.* **34**, 551–568 (2022).

59. Collins, A. G. E., Ciullo, B., Frank, M. J. & Badre, D. Working Memory Load Strengthens Reward Prediction Errors. *J. Neurosci.* **37**, 4332–4342 (2017).

60. Collins, A. G. E. The Tortoise and the Hare: Interactions between Reinforcement Learning and Working Memory. *J. Cogn. Neurosci.* **30**, 1422–1432 (2018).

61. Collins, A. G. E., Albrecht, M. A., Waltz, J. A., Gold, J. M. & Frank, M. J. Interactions Among Working Memory, Reinforcement Learning, and Effort in Value-Based Choice: A New Paradigm and Selective Deficits in Schizophrenia. *Biol. Psychiatry* **82**, 431–439 (2017).

62. Rutherford, A. V., McDougle, S. D. & Joormann, J. “Don’t [ruminate], be happy”: A cognitive perspective linking depression and anhedonia. *Clin. Psychol. Rev.* **101**, 102255 (2023).

63. Huys, Q. J., Pizzagalli, D. A., Bogdan, R. & Dayan, P. Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biol. Mood Anxiety Disord.* **3**, 12 (2013).

64. Kessler, R. C. *et al.* Anxious and non-anxious major depressive disorder in the World Health Organization World Mental Health Surveys. *Epidemiology Psychiatr. Sci.* **24**, 210–226 (2015).

65. Zorowitz, S., Momennejad, I. & Daw, N. D. Anxiety, Avoidance, and Sequential Evaluation. *Comput. Psychiatry* **4**, 1–17 (2020).

66. Robbins, T. W., Gillan, C. M., Smith, D. G., Wit, S. de & Ersche, K. D. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends Cogn Sci* **16**, 81–91 (2012).

67. Figee, M. *et al.* Compulsivity in obsessive–compulsive disorder and addictions. *Eur. Neuropsychopharmacol.* **26**, 856–868 (2016).

68. Gillan, C. M. & Robbins, T. W. Goal-directed learning and obsessive–compulsive disorder. *Philos. Trans. R. Soc. B: Biol. Sci.* **369**, 20130475 (2014).

69. Robbins, T. W. & Cardinal, R. N. Computational psychopharmacology: a translational and pragmatic approach. *Psychopharmacology* **236**, 2295–2305 (2019).

70. Lüscher, C., Robbins, T. W. & Everitt, B. J. The transition to compulsion in addiction. *Nat Rev Neurosci* **21**, 247–263 (2020).

71. Kanen, J. W., Ersche, K. D., Fineberg, N. A., Robbins, T. W. & Cardinal, R. N. 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* **236**, 2337–2358 (2019).

72. Ruan, Z., Liu, S., Liu, Y. an, Yang, Q. & Peng, Z. Disorders of compulsivity: Deficits in arbitrating learning strategies. *Addict. Biol.* **29**, e13433 (2024).

73. Paulus, M. P. & Stein, M. B. Interoception in anxiety and depression. *Brain Struct. Funct.* **214**, 451–463 (2010).

74. Khalsa, S. S. *et al.* Interoception and Mental Health: A Roadmap. *Biol. Psychiatry: Cogn. Neurosci. Neuroimaging* **3**, 501–513 (2018).

75. Smith, R. *et al.* A Bayesian computational model reveals a failure to adapt interoceptive precision estimates across depression, anxiety, eating, and substance use disorders. *PLoS Comput. Biol.* **16**, e1008484 (2020).

76. Seth, A. K. Interoceptive inference, emotion, and the embodied self. *Trends Cogn. Sci.* **17**, 565–573 (2013).

77. Barrett, L. F. & Simmons, W. K. Interoceptive predictions in the brain. *Nat. Rev. Neurosci.* **16**, 419–429 (2015).

78. Allen, M., Levy, A., Parr, T. & Friston, K. J. In the Body's Eye: The computational anatomy of interoceptive inference. *PLoS Comput. Biol.* **18**, e1010490 (2022).

79. Petzschner, F. H., Garfinkel, S. N., Paulus, M. P., Koch, C. & Khalsa, S. S. Computational Models of Interoception and Body Regulation. *Trends Neurosci.* **44**, 63–76 (2021).

80. Weber, L., Yee, D. M., Small, D. M. & Petzschner, F. H. Rethinking reinforcement learning: The interoceptive origin of reward. (2024) doi:10.31234/osf.io/be6nv.

81. Legrand, N. *et al.* The heart rate discrimination task: A psychophysical method to estimate the accuracy and precision of interoceptive beliefs. *Biol. Psychol.* **168**, 108239 (2022).

82. Nikolova, N. *et al.* The respiratory resistance sensitivity task: An automated method for quantifying respiratory interoception and metacognition. *Biol. Psychol.* **170**, 108325 (2022).

83. Harrison, O. K. *et al.* Interoception of breathing and its relationship with anxiety. *Neuron* **109**, 4080-4093.e8 (2021).

84. Lavalley, C. A. *et al.* Transdiagnostic failure to adapt interoceptive precision estimates across affective, substance use, and eating disorders: A replication and extension of previous results. *Biol. Psychol.* **191**, 108825 (2024).

85. Organization, W. H. Depression and Other Common Mental Disorders: Global Health Estimates. (2017).

86. Cooper, J. A., Arulpragasam, A. R. & Treadway, M. T. Anhedonia in depression: biological mechanisms and computational models. *Curr. Opin. Behav. Sci.* **22**, 128–135 (2018).

87. Shankman, S. A. *et al.* A Psychophysiological Investigation of Threat and Reward Sensitivity in Individuals With Panic Disorder and/or Major Depressive Disorder. *J. Abnorm. Psychol.* **122**, 322–338 (2013).

88. Alloy, L. B., Olino, T., Freed, R. D. & Nusslock, R. Role of Reward Sensitivity and Processing in Major Depressive and Bipolar Spectrum Disorders. *Behav. Ther.* **47**, 600–621 (2016).

89. Pizzagalli, D. A., Jahn, A. L. & O'Shea, J. P. Toward an objective characterization of an anhedonic phenotype: A signal-detection approach. *Biol. Psychiatry* **57**, 319–327 (2005).

90. Vrieze, E. *et al.* Reduced Reward Learning Predicts Outcome in Major Depressive Disorder. *Biol. Psychiatry* **73**, 639–645 (2013).

91. Kumar, P. *et al.* Impaired reward prediction error encoding and striatal-midbrain connectivity in depression. *Neuropsychopharmacology* **43**, 1581–1588 (2018).

92. Foell, J. *et al.* Ventral striatal activation during reward differs between major depression with and without impaired mood reactivity. *Psychiatry Res.: Neuroimaging* **313**, 111298 (2021).

93. Gotlib, I. H. & Joormann, J. Cognition and Depression: Current Status and Future Directions. *Annu. Rev. Clin. Psychol.* **6**, 285–312 (2010).

94. Cooper, J. A., Arulpragasam, A. R. & Treadway, M. T. Anhedonia in depression: biological mechanisms and computational models. *Curr. Opin. Behav. Sci.* **22**, 128–135 (2018).

95. Mukherjee, D., Filipowicz, A. L. S., Vo, K., Satterthwaite, T. D. & Kable, J. W. Reward and Punishment Reversal-Learning in Major Depressive Disorder. *J. Abnorm. Psychol.* **129**, 810–823 (2020).

96. Cools, R., Clark, L., Owen, A. M. & Robbins, T. W. Defining the Neural Mechanisms of Probabilistic Reversal Learning Using Event-Related Functional Magnetic Resonance Imaging. *J. Neurosci.* **22**, 4563–4567 (2002).

97. Dayan, P. Dopamine, Reinforcement Learning, and Addiction. *Pharmacopsychiatry* **42**, S56–S65 (2009).

98. Friedman, N. P. & Robbins, T. W. The role of prefrontal cortex in cognitive control and executive function. *Neuropsychopharmacology* **47**, 72–89 (2022).

99. Whitton, A. E., Treadway, M. T. & Pizzagalli, D. A. Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. *Curr Opin Psychiatr* **28**, 7–12 (2015).

100. Rutledge, R. B. *et al.* Association of Neural and Emotional Impacts of Reward Prediction Errors With Major Depression. *JAMA Psychiatry* **74**, 790–797 (2017).

101. Treadway, M. T. & Zald, D. H. Reconsidering anhedonia in depression: Lessons from translational neuroscience. *Neurosci. Biobehav. Rev.* **35**, 537–555 (2011).

102. Pessiglione, M., Bouc, R. L. & Vinckier, F. When decisions talk: computational phenotyping of motivation disorders. *Curr. Opin. Behav. Sci.* **22**, 50–58 (2018).

103. Loh, E. *et al.* Parsing the Role of the Hippocampus in Approach–Avoidance Conflict. *Cereb. Cortex* **27**, 201–215 (2017).

104. Sussman, T. J., Jin, J. & Mohanty, A. Top-down and bottom-up factors in threat-related perception and attention in anxiety. *Biol. Psychol.* **121**, 160–172 (2016).

105. Dillon, D. G. *et al.* PERIL AND PLEASURE: AN RDOC-INSPIRED EXAMINATION OF THREAT RESPONSES AND REWARD PROCESSING IN ANXIETY AND DEPRESSION. *Dépress. Anxiety* **31**, 233–249 (2014).

106. Yamamori, Y. & Robinson, O. J. Computational perspectives on human fear and anxiety. *Neurosci. Biobehav. Rev.* **144**, 104959 (2023).

107. Hartley, C. A. & Phelps, E. A. Anxiety and Decision-Making. *Biol. Psychiatry* **72**, 113–118 (2012).

108. Kahneman, D. & Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica* **47**, 263–292 (1979).

109. Aylward, J. *et al.* Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nat Hum Behav* **3**, 1116–1123 (2019).

110. Park, H., Lee, D. & Chey, J. Stress enhances model-free reinforcement learning only after negative outcome. *Plos One* **12**, e0180588 (2017).

111. Bishop, S. J. & Gagne, C. Anxiety, Depression, and Decision Making: A Computational Perspective. *Annu. Rev. Neurosci.* **41**, 1–18 (2018).

112. Mkrtchian, A., Aylward, J., Dayan, P., Roiser, J. P. & Robinson, O. J. Modeling Avoidance in Mood and Anxiety Disorders Using Reinforcement Learning. *Biol. Psychiatry* **82**, 532–539 (2017).

113. Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X. & Bishop, S. J. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat. Neurosci.* **18**, 590–596 (2015).

114. Pike, A. C. & Robinson, O. J. Reinforcement Learning in Patients With Mood and Anxiety Disorders vs Control Individuals. *JAMA Psychiatry* **79**, 313–322 (2022).

115. Beltzer, M. L., Adams, S., Beling, P. A. & Teachman, B. A. Social Anxiety and Dynamic Social Reinforcement Learning in a Volatile Environment. *Clin. Psychol. Sci.* **7**, 1372–1388 (2019).

116. LaFreniere, L. S. & Newman, M. G. Probabilistic Learning by Positive and Negative Reinforcement in Generalized Anxiety Disorder. *Clin. Psychol. Sci.* **7**, 502–515 (2018).

117. Yamamori, Y., Robinson, O. J. & Roiser, J. P. Approach-avoidance reinforcement learning as a translational and computational model of anxiety-related avoidance. *eLife* **12**, RP87720 (2023).

118. Huang, H., Thompson, W. & Paulus, M. P. Computational Dysfunctions in Anxiety: Failure to Differentiate Signal From Noise. *Biol. Psychiatry* **82**, 440–446 (2017).

119. APA. *Diagnostic And Statistical Manual Of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR)*. (American Psychiatric Association, 2022).

120. Redish, A. D., Jensen, S. & Johnson, A. A unified framework for addiction: Vulnerabilities in the decision process. *Behav Brain Sci* **31**, 415–437 (2008).

121. Redish, A. D. Addiction as a Computational Process Gone Awry. *Science* **306**, 1944–1947 (2004).

122. Gueguen, M. C. M. *et al.* Recent Opioid Use Impedes Range Adaptation in Reinforcement Learning in Human Addiction. *Biol. Psychiatry* **95**, 974–984 (2024).

123. Ahn, W. Y., Dai, J., Vassileva, J., Busemeyer, J. R. & Stout, J. C. Chapter 3 Computational modeling for addiction medicine From cognitive models to clinical applications. *Prog. Brain Res.* **224**, 53–65 (2016).

124. Vassileva, J., Lee, J. H., Psederska, E. & Ahn, W.-Y. Utility of computational approaches for precision psychiatry: Applications to substance use disorders. (2022) doi:10.31234/osf.io/uyzeg.

125. Voon, V. *et al.* Disorders of compulsivity: a common bias towards learning habits. *Mol. Psychiatry* **20**, 345–352 (2015).

126. Reiter, A. M. F. *et al.* Behavioral and Neural Signatures of Reduced Updating of Alternative Options in Alcohol-Dependent Patients during Flexible Decision-Making. *J. Neurosci.* **36**, 10935–10948 (2016).

127. Groman, S. M. *et al.* Orbitofrontal Circuits Control Multiple Reinforcement-Learning Processes. *Neuron* **103**, 734-746.e3 (2019).

128. Hogarth, L. Addiction is driven by excessive goal-directed drug choice under negative affect: translational critique of habit and compulsion theory. *Neuropsychopharmacol* **45**, 720–735 (2020).

129. Myers, C. E. *et al.* Probabilistic reward- and punishment-based learning in opioid addiction: Experimental and computational data. *Behav. Brain Res.* **296**, 240–248 (2016).

130. Haines, N., Vassileva, J. & Ahn, W. The Outcome-Representation Learning Model: A Novel Reinforcement Learning Model of the Iowa Gambling Task. *Cogn. Sci.* **42**, 2534–2561 (2018).

131. Tiego, J. *et al.* Precision behavioral phenotyping as a strategy for uncovering the biological correlates of psychopathology. *Nat. Ment. Heal.* **1**, 304–315 (2023).

132. Yechiam, E., Busemeyer, J. R., Stout, J. C. & Bechara, A. Using Cognitive Models to Map Relations Between Neuropsychological Disorders and Human Decision-Making Deficits. *Psychol. Sci.* **16**, 973–978 (2005).

133. Schurr, R., Reznik, D., Hillman, H., Bhui, R. & Gershman, S. J. Dynamic computational phenotyping of human cognition. *Nat. Hum. Behav.* **8**, 917–931 (2024).

134. Karvelis, P., Paulus, M. P. & Diaconescu, A. O. Individual differences in computational psychiatry: A review of current challenges. *Neurosci. Biobehav. Rev.* **148**, 105137 (2023).

135. Rouder, J. N., Kumar, A. & Haaf, J. M. Why many studies of individual differences with inhibition tasks may not localize correlations. *Psychon. Bull. Rev.* **30**, 2049–2066 (2023).

136. Myung, J. I. & Pitt, M. A. Optimal Experimental Design for Model Discrimination. *Psychol. Rev.* **116**, 499–518 (2009).

137. Myung, J. I., Cavagnaro, D. R. & Pitt, M. A. A tutorial on adaptive design optimization. *J. Math. Psychol.* **57**, 53–67 (2013).

138. Kwon, M., Lee, S. H. & Ahn, W.-Y. Adaptive Design Optimization as a Promising Tool for Reliable and Efficient Computational Fingerprinting. *Biol. Psychiatry: Cogn. Neurosci. Neuroimaging* **8**, 798–804 (2023).

139. Ahn, W.-Y. *et al.* Rapid, precise, and reliable measurement of delay discounting using a Bayesian learning algorithm. *Sci Rep-uk* **10**, 12091 (2020).

140. Lee, J.-H. *et al.* Rapid and Reliable Computational Markers for Predicting Daily Smoking Behavior and Smoking Cessation Treatment Outcomes. *PsyArxiv* (2025) doi:10.31234/osf.io/bj7fd_v4.

141. Sonkusare, S., Breakspear, M. & Guo, C. Naturalistic Stimuli in Neuroscience: Critically Acclaimed. *Trends Cogn. Sci.* **23**, 699–714 (2019).

142. Finn, E. S. *et al.* Idiosynchrony: From shared responses to individual differences during naturalistic neuroimaging. *NeuroImage* **215**, 116828 (2020).

143. Finn, E. S. & Bandettini, P. A. Movie-watching outperforms rest for functional connectivity-based prediction of behavior. *NeuroImage* **235**, 117963 (2021).

144. Datta, S. R., Anderson, D. J., Branson, K., Perona, P. & Leifer, A. Computational Neuroethology: A Call to Action. *Neuron* **104**, 11–24 (2019).

145. Mobbs, D. *et al.* Promises and challenges of human computational ethology. *Neuron* **109**, 2224–2238 (2021).

146. Lee, S. H., Song, M. S., Oh, M. & Ahn, W.-Y. Bridging the Gap Between Self-Report and Behavioral Laboratory Measures: A Real-Time Driving Task With Inverse Reinforcement Learning. *Psychol. Sci.* **35**, 345–357 (2023).

147. Wise, T., Emery, K. & Radulescu, A. Naturalistic reinforcement learning. *Trends Cogn. Sci.* **28**, 144–158 (2024).

148. Mnih, V. *et al.* Human-level control through deep reinforcement learning. *Nature* **518**, 529–533 (2015).

149. Gelly, S. & Silver, D. Combining online and offline knowledge in UCT. *Proc. 24th Int. Conf. Mach. Learn.* 273–280 (2007) doi:10.1145/1273496.1273531.

150. Cross, L., Cockburn, J., Yue, Y. & O'Doherty, J. P. Using deep reinforcement learning to reveal how the brain encodes abstract state-space representations in high-dimensional environments. *Neuron* (2020) doi:10.1016/j.neuron.2020.11.021.

151. Volkow, N. D., Fowler, J. S., Wang, G. J., Baler, R. & Telang, F. Imaging dopamine's role in drug abuse and addiction. *Neuropharmacology* **56**, 3–8 (2009).

152. Volkow, N. D., Fowler, J. S., Wang, G.-J. & Goldstein, R. Z. Role of Dopamine, the Frontal Cortex and Memory Circuits in Drug Addiction: Insight from Imaging Studies. *Neurobiol. Learn. Mem.* **78**, 610–624 (2002).

153. Cuthbert, B. N. & Insel, T. R. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Med.* **11**, 126 (2013).

154. Insel, T. R. The NIMH Research Domain Criteria (RDoC) Project: Precision Medicine for Psychiatry. *Am. J. Psychiatry* **171**, 395–397 (2014).

155. Kozak, M. J. & Cuthbert, B. N. The NIMH Research Domain Criteria Initiative: Background, Issues, and Pragmatics. *Psychophysiology* **53**, 286–297 (2016).

156. Lilienfeld, S. O. The Research Domain Criteria (RDoC): An analysis of methodological and conceptual challenges. *Behav. Res. Ther.* **62**, 129–139 (2014).

157. Morris, S. E. & Cuthbert, B. N. Research Domain Criteria: cognitive systems, neural circuits, and dimensions of behavior. *Dialogues Clin. Neurosci.* **14**, 29–37 (2012).

158. Weinberger, D. R. & Goldberg, T. E. RDoCs redux. *World Psychiatry* **13**, 36–38 (2014).

159. Adams, R. A., Huys, Q. J. M. & Roiser, J. P. Computational Psychiatry: towards a mathematically informed understanding of mental illness. *J. Neurol., Neurosurg. Psychiatry* **87**, 53 (2016).

160. Hu, K. Neural activity to threat in ventromedial prefrontal cortex correlates with individual differences in anxiety and reward processing. *Neuropsychologia* **117**, 566–573 (2018).

161. Beauchaine, T. P. & Hinshaw, S. P. RDoC and Psychopathology among Youth: Misplaced Assumptions and an Agenda for Future Research. *J. Clin. Child Adolesc. Psychol.* **49**, 322–340 (2020).

162. Ross, C. A. & Margolis, R. L. Research Domain Criteria: Strengths, Weaknesses, and Potential Alternatives for Future Psychiatric Research. *Mol. Neuropsychiatry* **5**, 218–236 (2019).

163. Hedge, C., Powell, G. & Sumner, P. The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behav. Res. Methods* **50**, 1166–1186 (2018).

164. Enkavi, A. Z. *et al.* Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proc. Natl. Acad. Sci.* **116**, 5472–5477 (2019).

165. Nitsch, F. J., Lüpken, L. M., Lüschow, N. & Kalenscher, T. On the reliability of individual economic rationality measurements. *Proc. Natl. Acad. Sci.* **119**, e2202070119 (2022).

166. Haines, N. *et al.* Theoretically Informed Generative Models Can Advance the Psychological and Brain Sciences: Lessons from the Reliability Paradox. (2020) doi:10.31234/osf.io/xr7y3.

167. Rouder, J. N. & Haaf, J. M. A psychometrics of individual differences in experimental tasks. *Psychon. Bull. Rev.* **26**, 452–467 (2019).

168. Mkrtchian, A., Valton, V. & Roiser, J. P. Reliability of Decision-Making and Reinforcement Learning Computational Parameters. *Comput. Psychiatry* **7**, 30 (2023).

169. Price, R. B., Brown, V. & Siegle, G. J. Computational Modeling Applied to the Dot-Probe Task Yields Improved Reliability and Mechanistic Insights. *Biol. Psychiatry* **85**, 606–612 (2019).

170. Brown, V. M., Chen, J., Gillan, C. M. & Price, R. B. Improving the Reliability of Computational Analyses: Model-Based Planning and Its Relationship With Compulsivity. *Biol. Psychiatry: Cogn. Neurosci. Neuroimaging* **5**, 601–609 (2020).

171. Ahn, W.-Y. *et al.* Decision-making in stimulant and opiate addicts in protracted abstinence: evidence from computational modeling with pure users. *Front. Psychol.* **5**, 849 (2014).

172. Wiecki, T. V., Sofer, I. & Frank, M. J. HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Front. Neuroinformatics* **7**, 14 (2013).

173. Waltmann, M., Schlagenhauf, F. & Deserno, L. Sufficient reliability of the behavioral and computational readouts of a probabilistic reversal learning task. *Behav. Res. Methods* **54**, 2993–3014 (2022).

174. Wiecki, T. V., Poland, J. & Frank, M. J. Model-Based Cognitive Neuroscience Approaches to Computational Psychiatry. *Clin. Psychol. Sci.* **3**, 378–399 (2014).

175. Dhamala, E., Yeo, B. T. T. & Holmes, A. J. One Size Does Not Fit All: Methodological Considerations for Brain-Based Predictive Modeling in Psychiatry. *Biol. Psychiatry* **93**, 717–728 (2023).

176. Monosov, I. E., Zimmermann, J., Frank, M. J., Mathis, M. W. & Baker, J. T. Ethological computational psychiatry: Challenges and opportunities. *Curr. Opin. Neurobiol.* **86**, 102881 (2024).

177. Berwian, I., Hitchcock, P., Pisupati, S., Schoen, G. & Niv, Y. Using computational models of learning to advance cognitive behavioral therapy. (2023) doi:10.31234/osf.io/8snbq.

178. Queirazza, F., Fouragnan, E., Steele, J. D., Cavanagh, J. & Philiastides, M. G. Neural correlates of weighted reward prediction error during reinforcement learning classify response to cognitive behavioral therapy in depression. *Sci. Adv.* **5**, eaav4962 (2019).