



Reinforcement Learning and Decision Making in Depression in Adolescents and Young Adults: Insights from a New Model of the Probabilistic Reward Task

RESEARCH ARTICLE

ZIWEI CHENG 

AMELIA D. MOSER 

JENNA JONES

CHRISTOPHER D. SCHNECK 

DAVID J. MIKLOWITZ 

DANIEL G. DILLON 

ROSELINDE H. KAISER 

][ubiquity press

**Author affiliations can be found in the back matter of this article*

ABSTRACT

Depression is a prevalent psychiatric condition that commonly emerges in adolescence and young adulthood and is associated with reward processing abnormalities. The Probabilistic Reward Task (PRT) is widely used to investigate the impact of depression on reward processing, but prior studies have not comprehensively addressed the reinforcement learning and decision-making mechanisms involved in the task. In 726 adolescents and young adults with varying levels of depression, we collected PRT data and applied a novel computational model with response-outcome learning and evidence accumulation processes to provide new insights into the cognitive processes implicated in depression. Compared to participants with no history of psychopathology, those with depressive disorders showed reduced impact of learned response values on decision bias toward the more frequently rewarded action. In addition, higher levels of anhedonia were associated with slower evidence accumulation during decision-making. Together, these findings improved our understanding of the reinforcement learning and decision-making mechanisms assessed by the PRT and their associations with depression.

CORRESPONDING AUTHOR:

Roselinde H. Kaiser

Research on Affective Disorders and Development Laboratory,
Muenzinger D244, USB 345,
University of Colorado Boulder,
Boulder CO 80309, United States

Roselinde.Kaiser@colorado.edu

KEYWORDS:

computational modeling;
depression; development;
reward; decision making

TO CITE THIS ARTICLE:

Cheng, Z., Moser, A. D., Jones, J., Schneck, C. D., Miklowitz, D. J., Dillon, D. G., & Kaiser, R. H. (2025). Reinforcement Learning and Decision Making in Depression in Adolescents and Young Adults: Insights from a New Model of the Probabilistic Reward Task. *Computational Psychiatry*, 9(1), pp. 268–283. DOI: <https://doi.org/10.5334/cpsy.147>

Depression is associated with abnormalities in reward processing (Diekhof et al., 2008; Zald & Treadway, 2017). Prior research evaluating differences in reward behavior related to clinical diagnoses and symptom severity has provided insights into such abnormalities (Rupprechter et al., 2018; Safra et al., 2019), with the goal of translating research findings into treatment recommendations (Admon & Pizzagalli, 2015; Treadway & Zald, 2013). Investigation of reward processing abnormalities in adolescents and young adults with major depression has particularly important clinical implications. Given that adolescence is characterized by both changes in reward sensitivity and elevated risk of mood disorders, examining their relationship during this developmental period may provide insight into the etiology of early-onset depression (Forbes & Dahl, 2012; Luking et al., 2016).

The current study administered a well-known reward assay, the Probabilistic Reward Task (PRT; Pizzagalli et al., 2005), to a large sample of adolescents and young adults (Table 1; $n = 726$) and used computational modeling to draw inferences about (a) cognitive processes that support PRT performance and (b) the impact of depression on those processes. In the PRT, correct identifications of one (“rich”) stimulus are rewarded more frequently than correct identifications of another (“lean”) stimulus. Because of the reward asymmetry, participants typically develop a response bias such that they make the “rich” response more frequently than the “lean” response, even though the rich and lean stimuli are designed to be ambiguous and are shown with equal frequency. Previous studies have found associations between a blunted rich response bias and depression diagnoses (e.g., Vrieze et al. 2013; Morris et al., 2015), decreased dopaminergic signaling (e.g., Grob et al., 2012; Kaiser et al., 2018), and anhedonia (see Kangas et al., 2022, for review). The PRT has offered valuable insight into behavioral biomarkers associated with reward dysfunction and anhedonia, and consequently it is recognized in the NIH’s RDoC Matrix as a validated measure of reward learning (Insel et al., 2010).

	M (SD)
	N = 726 n (%)
Gender	
Cisgender female	476 (66%)
Cisgender male	225 (31%)
Non-binary, transgender, gender-fluid, or not reported	23 (3.2%)
Ethnicity	
Hispanic or Latino/Latinx/Latine	122 (17%)
Non-Hispanic and non-Latino/Latinx/Latine	602 (83%)
Not reported	2 (0.28%)
Race	
African American	20 (2.8%)
American Indian/Alaskan native	2 (0.3%)
Asian	125 (18%)
Native Hawaiian/other Pacific Islander	1 (0.1%)
White	432 (62%)
More than one race	56 (8.1%)
Other or not reported	62 (8.9%)
Age	19 (2)
Mood Symptoms	
MASQ-LOI	16.71 (6.79)

(Contd.)

Table 1 Demographics and clinical characteristics of the sample.

Notes: MASQ-LOI = Mood and Anxiety Symptom Questionnaire, Loss of Interest Anhedonia subscale.

	N = 422 n (%)
Current Mood Diagnoses	
(Unipolar) Depressive Disorder	121 (29%)
<i>Major depressive disorder</i>	73 (17%)
<i>Persistent depressive disorder</i>	48 (11%)
Bipolar Disorder	26 (6.2%)
<i>Bipolar I disorder</i>	11 (2.6%)
<i>Bipolar II disorder</i>	10 (2.4%)
<i>Bipolar disorder not otherwise specified</i>	5 (1.2%)
Non-Psychiatric Control	275 (65%)

In addition to behavioral analyses, prior studies have applied analytic techniques to PRT data that can isolate component processes to understand dimensions of reward processing dysfunction in depression (e.g., [Pizzagalli et al., 2008](#); [Vrieze et al., 2013](#); [Lawlor et al., 2020](#); [Whitton et al., 2020](#)). In particular, reinforcement learning (RL) models provide a computational perspective on how individuals evaluate and integrate reward information to modify behavior in the PRT. Investigators using the PRT may assume that participants update the values of stimulus-action pairs based on trial-level prediction errors (discrepancy between received and expected reward), such that a bias toward the “rich” response develops over time because of the acquired higher expected value of the rich stimulus. These models have been used to capture the gradual development of response bias across trials and examine relationships between reward learning rate, reward sensitivity, and clinical measures ([Huys et al., 2013](#); [Letkiewicz et al., 2022](#)).

However, recent work characterizing decision-making dynamics in PRT performance shows that updates in stimulus-response associations may not be the main driving mechanism in the task ([Lawlor et al., 2020](#); [Dillon et al., 2022, 2024](#)). Specifically, studies found that the response bias effect is larger for faster compared with slower response times (RTs). This result suggests that, on many trials, participants prepare to make a rich response very quickly. Thus, participants may learn *response-outcome* associations in the PRT, indicative of a true response bias that does not depend heavily on stimulus evaluation (i.e., participants quickly press “rich” regardless of the stimulus shown; [White & Poldrack, 2014](#)). To better understand the underlying decision mechanisms, researchers applied the drift diffusion model (DDM; [Ratcliff & Rouder, 1998](#); [Ratcliff & McKoon, 2018](#); [Wiecki et al., 2013](#)) to capture choice frequencies and RT distributions in the PRT ([Lawlor et al., 2020](#); [Dillon et al. 2022, 2024](#); [Pitliya et al., 2022](#); [Shen et al., 2024](#)). These analyses revealed a strong starting point bias, suggesting that during the task, participants developed a preference towards the rich response *before* evaluating the stimulus.

Importantly, although both RL models and the DDM have provided insight into reward processing as indexed by the PRT, a recent comparison highlighted shortcomings in both approaches ([Dillon et al., 2024](#)). RL models capture learning and development of response bias in the PRT, however, they do not account for RTs (or the dependency of response bias on RTs) and psychometric properties of some parameters are weak. By contrast, the DDM captures decision processes in the PRT and parameters generally have strong psychometric properties, but the DDM does not measure learning and cannot account for the development of response bias, which is central to the PRT. Recently developed integrative models combine RL and DDM modeling frameworks (RLDDM; [Pedersen et al., 2017](#); [Pedersen & Frank, 2020](#)), but these existing integrative models do not fit the PRT better than the standard DDM ([Dillon et al., 2024](#)). This lack of improvement may stem from two differences between the PRT and traditional RL tasks. First, we hypothesize that in the PRT reward learning is driven by response-outcome (rather than stimulus-response) associations. Second, the PRT requires a difficult perceptual judgment on each trial; the focus on stimulus-locked prediction errors in most RL models, including the RLDDM, may not be ideal for PRT analyses (because the stimuli used in the PRT are too similar to support robust stimulus-locked prediction errors). In short, although much progress has been made, developing models that can better account for learning

and decision dynamics in the PRT remains a key goal. Such models may complement information provided by reinforcement learning or decision-making approaches alone, yielding new insights into cognitive and behavioral processes implicated in (and affected by) depression.

To address these goals and to identify processes that may confer risk for future depressive illness, the current study applied an adapted reinforcement learning and drift diffusion model to account for trial-level learning and decision dynamics in the PRT. In the new model (Action-DDM), the learning mechanism is driven by response-outcome associations (the values of response options change based on reward feedback) and the choice mechanism is described by DDM (choices are made by continuously sampling and accumulating noisy evidence over time until a decision boundary is reached). Crucially, the reward learning mechanism influences decision-making such that the value differences between actions modulate (1) drift rate (evidence accumulates more quickly towards boundaries that correspond to actions with higher expected values) and (2) starting point bias (there is a pre-stimulus preference for actions associated with higher expected values).

We evaluated the model by assessing posterior predictive accuracy and parameter recovery, and also tested the split-half reliability of model parameters. We then tested associations between model parameters and depression in the full sample; exploratory analyses repeated these tests in adolescent (ages 13–19) or young adult (ages 20–31) subgroups. The analyses used both a dimensional approach focused on anhedonic depression severity and a categorical approach comparing parameters in individuals diagnosed with depressive disorders and those with no psychiatric history. We hypothesized that individuals currently experiencing more severe anhedonic depression and/or with unipolar depression diagnoses would show reduced reward responsiveness (lower response bias), difficulty learning from rewards (lower learning rate), and slower evidence accumulation during decision-making (lower drift rate).

METHODS

PARTICIPANTS

A total of 726 participants, ages 13 to 31 years ($M = 19.10$, $SD = 2.44$), were recruited from the greater Los Angeles, California, and Boulder, Colorado, areas for one of five research protocols with overlapping procedures (Supplement Table 1–5). The goal was to recruit participants who were in a period of high developmental risk for mood disorders. Eligible participants had to speak fluent English, have normal or corrected-to-normal vision, and report no neurological or cognitive impairments. Sources of recruitment included community flyers, school events, local clinics, and electronic postings. Three of the five protocols aimed to recruit samples with high variance in mood disorders and symptoms, and therefore included eligibility criteria requiring that participants either reported a primary mood disorder, first-degree family history of mood disorder, or had no lifetime history of any psychopathology. Participants were excluded for suicidal ideation when deemed necessary by a study clinician. The research protocols were approved by Institutional Review Boards at the University of California Los Angeles (Protocols #16-001894, #17-000065) and the University of Colorado Boulder (Protocols #18-0600, #18-0415, #19-0130, #20-0475). Legal adults (ages 18 years and older) or guardians of legal minors (ages 17 years and younger) provided informed consent to participate, and minors under 18 also provided assent.

PROCEDURES

Participants were recruited for an in-person session that included self-report electronic surveys and cognitive tasks administered via computer by trained staff. Participants in three of the five protocols also completed a clinical diagnostic interview (Structured Clinical Interview for DSM-5, SCID-5; [First et al., 2015](#)). At the end of the session, participants were debriefed and received financial compensation. The present study used data from the PRT, clinical diagnostic information, and self-reported symptoms of anhedonic depression. Additional research procedures and non-overlapping research results are reported elsewhere ([Fassett-Carman et al., 2023](#); [Kaiser et al., 2022](#); [Peterson et al., 2021](#); [Peterson et al., 2022](#)).

Probabilistic Reward Task (PRT)

On each PRT trial (Figure 1A), participants saw a fixation cross (duration: 500 ms), blank cartoon face (500 ms), and a brief presentation (100 ms) of a short (11.5 mm) or long (13 mm) mouth on the face. The cartoon face remained on the screen until participants indicated whether a short or long mouth was shown by pressing the corresponding button on a button box or until the 2 second response deadline was reached. After responding, participants saw either a blank screen (null feedback) or text indicating monetary reward (“Correct! You won 5 cents!”) for 1750 ms. The task consisted of two 100-trial blocks. In each block, an equal number of short and long mouths were presented, but the reinforcement rate was asymmetric: rewards were delivered three times more often for correct identifications of the rich stimulus compared to the lean stimulus (assignment of rich/lean conditions to short/long mouths was counterbalanced across participants). Participants practiced differentiating between the mouth stimuli and were told that not all correct responses would be rewarded, but they were not informed of the asymmetric reinforcement schedule.

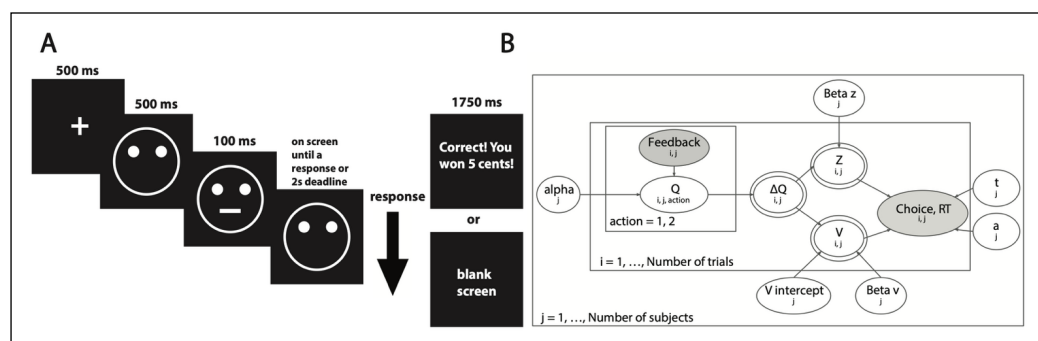


Figure 1 Probabilistic Reward Task and Graphical Illustration of Action-DDM.

(A) On each trial, participants saw a face with a long or short mouth and responded by pressing a button to indicate which mouth length was shown. Rewards were delivered three times more often for correct identifications of the rich vs. lean stimulus; assignment of short/long mouths to the rich/lean conditions was counterbalanced. (B) Graphical illustration of Action-DDM. Shaded nodes represent observed data and unshaded nodes represent parameter estimations. Double-bordered nodes represent trial-wise computed variables. Parameters include α : learning rate; β_v : the degree to which value differences influenced drift rate; β_z : the degree to which value differences influenced starting point bias; $V_{intercept}$: baseline stimulus processing efficiency or drift rate; t : non-decision time; a : decision threshold.

CLINICAL MEASURES

SCID-5

A subset of the sample ($n = 422$ of 726) were interviewed with the SCID-5 (First et al., 2015). Because the subsample of participants who met criteria for bipolar disorders (BIP: $n = 26$) was deemed too small for balanced group analyses, diagnostic group comparisons were restricted to participants with current unipolar depressive disorders (UNI: $n = 121$) compared with a non-psychiatric control group (NC: $n = 275$) (see Supplement Table 6 for detailed clinical characteristics).

MASQ-Loss of Interest

We evaluated current (past week) anhedonic depressive symptoms with the eight-item Loss of Interest Anhedonia subscale of the Mood and Anxiety Symptom Questionnaire (MASQ; Watson et al., 1995a; 1995b). Participants read items describing anhedonic feelings or experiences and indicated the extent to which the item described how they felt in the past week on a scale from 1 (not at all) to 5 (extremely). Sample items include “Felt withdrawn from others” and “Felt like being alone” (Watson et al., 1995a; 1995b). Responses were summed to yield a measure of anhedonic depression severity.

DATA ANALYSIS

We followed quality control procedures applied in prior PRT studies (Pizzagalli et al., 2005; Lawlor et al., 2020; Dillon et al., 2022; See Supplement for details) After quality control exclusions, PRT data were available for $n = 671$.

Following prior studies (e.g., [Pizzagalli et al., 2005, 2008](#)), our four key measures were accuracy, discriminability, overall response bias, and change in response bias (block 2 minus block 1). Accuracy referred to the number of correct trials divided by the total number of trials, calculated separately for the rich stimulus and lean stimulus. Response bias and discriminability were calculated with the following equations ([Pizzagalli et al., 2005](#)):

$$responsebias = \frac{1}{2} \log \left(\frac{(Rich_{correct} + 0.5) * (Lean_{incorrect} + 0.5)}{(Rich_{incorrect} + 0.5) * (Lean_{correct} + 0.5)} \right) \quad \text{Equation 1}$$

$$discriminability = \frac{1}{2} \log \left(\frac{(Rich_{correct} + 0.5) * (Lean_{correct} + 0.5)}{(Rich_{incorrect} + 0.5) * (Lean_{incorrect} + 0.5)} \right) \quad \text{Equation 2}$$

Computational Modeling

Action-DDM. We adapted an RL model ([Huys et al., 2013](#)) and a DDM previously applied to the task (e.g., [Lawlor et al., 2020](#); [Dillon et al., 2022](#)) to better characterize learning and decision dynamics in the PRT. In the Action-DDM ([Figure 1B](#)), the upper and lower decision boundaries corresponded to “rich” and “lean” responses, respectively. The expected value of response updates according to the delta learning rule ([Rescorla and Wagner 1972](#); Equation 3). On every trial, the model calculated the difference between the expected value of the responses (Equation 4). The speed of evidence accumulation was described by a drift rate intercept (constant across trials and conceptualized as each participant’s baseline stimulus processing efficiency), the action value difference (ΔQ), and a B_v parameter that captured the degree to which the action value difference influenced an agent’s drift rate (Equation 5). In addition, although prior studies of DDMs with RL components assumed that participants could not develop a decision bias (represented as a change in starting point z) because of randomized symbol-value associations (e.g., [Pedersen et al., 2017](#)), bias toward the rich response boundary is a key behavioral pattern in the PRT. Therefore, to capture the development of preference toward the rich response as a function of reinforcement history, we included a B_z parameter that described the degree to which the acquired response values influenced an agent’s starting point bias on each trial (Equation 6; z is SoftMax-transformed to be bounded between 0 and 1).

$$Q_{t+1}(\text{response}) = Q_t(\text{response}) + \alpha(r - Q_t(\text{response})) \quad \text{Equation 3}$$

$t = \text{trial number}$, $\text{response} = \text{chosen action}$, $r = 1$ when rewarded and 0 otherwise

$$\Delta Q_t = Q_{\text{rich response}} - Q_{\text{lean response}} \quad \text{Equation 4}$$

$$v_t = v_{\text{intercept}} + \Delta Q_t * B_v \quad \text{Equation 5}$$

$$z_t = \frac{e^{B_z * Q_t(\text{rich response})}}{e^{B_z * Q_t(\text{rich response})} + e^{B_z * Q_t(\text{lean response})}} \quad \text{Equation 6}$$

Conceptually, the model implies that at the start of the task ($\Delta Q_t = 0$, $v_t = v_{\text{intercept}}$, $z_t = 0.5$), participants perceive the mouth shown on screen and accumulate perceptual evidence until there is sufficient evidence to cross a response threshold. As the trials continue, participants update the expected value of the actions. As ΔQ_t increases (i.e., $Q_{\text{rich response}} > Q_{\text{lean response}}$), drift rate toward the (correct) “rich” response becomes higher than drift rate for the (correct) “lean” response, leading to higher rich accuracy. In addition, as ΔQ_t increases, the starting point bias shifts towards the rich boundary, leading to faster “rich” responses regardless of the stimulus.

On each trial, we calculated the likelihood of the observed response and RT with the Wiener first-passage time (WFPT) distribution ([Wabersich & Vandekerckhove, 2014](#)). For a rich (lean) stimulus, the v_t was positively (negatively) signed to represent evidence accumulation toward the rich (lean) response, and z_t specified the pre-stimulus bias toward the rich response. The decision threshold (a) indicated how far apart the two response boundaries were and non-decision time (t) accounted for non-decisional processes that added to the RT (i.e., perception and response execution). We applied hierarchical Bayesian modeling to improve parameter estimations at the individual level, and used the RStan package ([Stan Development Team, 2022](#)) for posterior sampling.

Model Convergence. Models were fit using Markov Chain Monte Carlo sampling. For each model, we drew samples from the posterior distribution three times (10,000 samples, 5,000 burn-in, every fifth sample retained) and computed the Gelman Rubin statistic \hat{R} (Gelman & Rubin, 1992) to ensure that the samples converged on a stable solution. \hat{R} values compared between- to within-chain variance.

Model Comparisons and Validations. See Supplement for detailed alternative model specifications and comparisons. To validate the winning models, we ran posterior predictive checks by using the obtained model parameters to simulate 20 datasets per participant with 200 trials per set; we then compared the simulated results to the observed data.

Depression-Related Differences in Behavior

Two-sample *t*-tests compared individuals with unipolar depressive disorders vs. individuals with no psychiatric history for the following parameters: discriminability, response bias, response bias change, $v_{intercept}$, B_v , B_z , α , a , and t . To assess associations between anhedonic depression severity and reward processing, we computed correlations between the above parameters and scores on the MASQ Loss of Interest scale.

To further examine depression-related differences in PRT within developmental subgroups, we divided participants into younger vs. older groups using a median split and repeated the analyses. Sample 1 (ages 13–19) included 393 participants (198 NC, 67 UNI), and Sample 2 (ages 20–31) included 273 participants (55 NC, 44 UNI).

Parameter Validation

To evaluate internal consistency, we split the PRT data by block, fitted the winning model to each block, and quantified across-block parameter agreement with Spearman-Brown correlation coefficients (Brown, 1910; Spearman, 1910). We performed parameter recovery by sampling from the posterior distributions of parameter estimates for each participant, simulating data, fitting the model to the simulated data, and calculating the correlations between the true and recovered parameters.

RESULTS

PERFORMANCE STATISTICS

Overall, participants developed a bias to choose the rich stimulus and showed better accuracy and shorter RTs on rich compared to lean trials (Table 2). Considering that the stimulus effect on accuracy may vary based on RT (Dillon et al., 2022; Lawlor et al., 2020), we classified each participant's fast (slow) RT type as the first 0.1 (last 0.9) quantile of their RT distribution (see Figure S1 for RT visualizations). Results from a multilevel logistic regression model to predict trial-level accuracy revealed a main effect of the rich stimulus ($B = 0.39$, $p < .001$), a main effect of slow RTs ($B = 0.83$, $p < .001$) and an interaction between stimulus type and RT type ($B = -0.08$, $p = .035$). Consistent with prior work, this interaction effect revealed that the rich > lean accuracy effect was larger in trials with fast vs. slow RTs, confirming that the PRT elicited a true response bias (i.e., the rich > lean difference in accuracy was greater on fast trials).

	VARIABLE	MEAN	SD
Block 1	response bias	0.05	0.21
	discriminability	0.67	0.27
	rich acc	0.82	0.10
	lean acc	0.79	0.12
	rich RT	529.64	91.94
	lean RT	530.00	92.43

(Contd.)

Table 2 Summary of PRT performance statistics.

	VARIABLE	MEAN	SD
	rich correct RT	530.67	89.50
	lean correct RT	529.14	89.91
	rich error RT	540.25	143.50
	lean error RT	555.30	144.90
Block 2	response bias	0.08	0.23
	discriminability	0.64	0.27
	rich acc	0.82	0.11
	lean acc	0.76	0.14
	rich RT	532.88	94.20
	lean RT	541.15	95.56
	rich correct RT	532.62	91.94
	lean correct RT	543.75	93.81
	rich error RT	550.44	140.57
	lean error RT	554.92	135.46
Average	response bias	0.07	0.19
	discriminability	0.66	0.25
	rich acc	0.82	0.09
	lean acc	0.77	0.12
	rich RT	531.26	88.70
	lean RT	535.57	88.98
	rich correct RT	531.65	86.19
	lean correct RT	536.45	86.99
	rich error RT	544.77	128.20
	lean error RT	554.31	128.11

COMPUTATIONAL MODELING

The across-run maximum Gelman-Rubin statistic for the MCMC sampler was 1.02, under the recommended threshold of 1.1 (Gelman & Rubin, 1992). The Action-DDM outperformed: (1) a model in which action value differences did not influence drift rate; (2) a model in which action value differences did not influence starting point bias; (3) a model in which the influence of action value differences on drift rate was not controlled by B_v ; (4) a model in which the influence of action value differences on starting point bias was not controlled by B_z ; and (5) a model with one B that controlled the extent to which action value differences influenced both starting point bias and drift rate (see Supplement for alternative model details). The behavior simulated by the Action-DDM captured key patterns in participants' choices and response times: it captured larger response bias effects for faster compared to slower RTs, a gradual increase in response bias across trials, and the development of faster RTs when choosing the more frequently rewarded action across trials (Figure 2).

DEPRESSION-RELATED DIFFERENCES IN BEHAVIOR

Compared to controls with no history of psychopathology, depressed individuals showed lower overall response bias ($t = -1.99$, $p = .048$) and reduced adjustments in starting point bias based on learned response values (B_z , $t = -2.51$, $p = .01$). There were no significant group differences on other measures or parameters (Figure 3). Severity of anhedonic symptoms was negatively correlated with drift rate intercept ($v_{intercept}$; $r = -0.09$, $p = .02$). There were no other associations

between PRT measures or parameters and anhedonic symptom severity (Figure 4). These findings were consistent after controlling for age (see Supplement for detailed analyses on associations between age and task parameters).

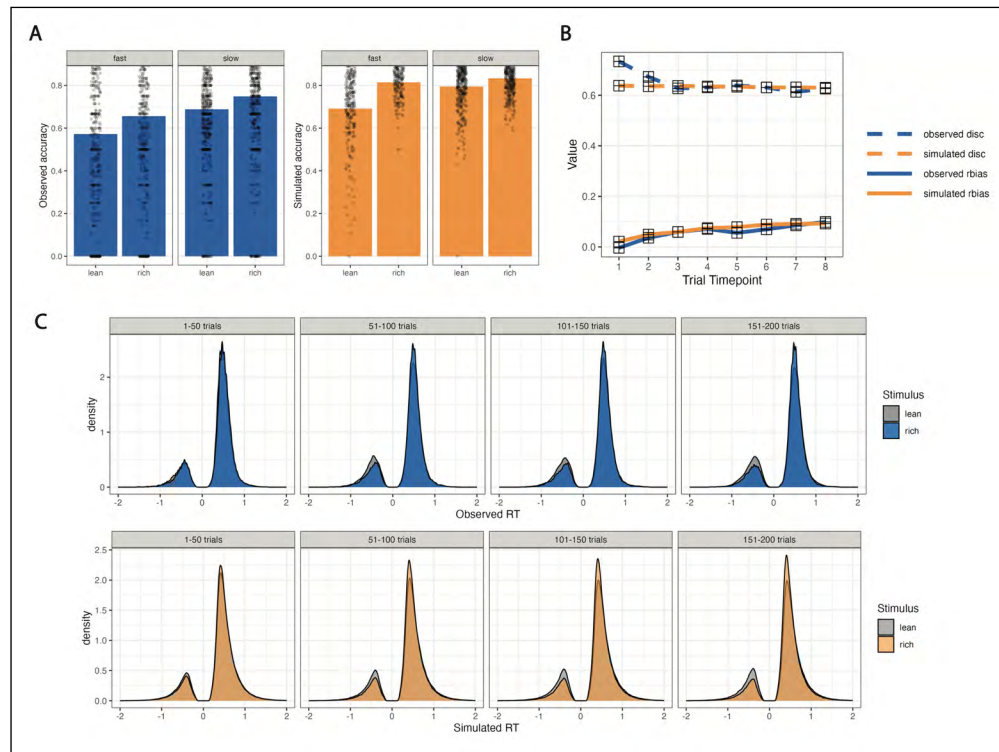


Figure 2 Posterior Predictive Performance. (A) Observed and simulated overall accuracy by stimulus type and response time (fast RT < .1 RT quantile; slow RTs > .9 RT quantile). (B) Observed and simulated changes in response bias and discriminability. The task trials were binned into eight timepoints with 25 trials each and response bias and discriminability were calculated independently for each timepoint. (C) Observed and simulated changes in response time distributions across four timepoints with 50 trials each. For a stimulus (rich or lean), we plotted the RT distributions when participants responded correctly (positively-valued RTs) and incorrectly (flipped to be negatively-valued RTs for illustrations).

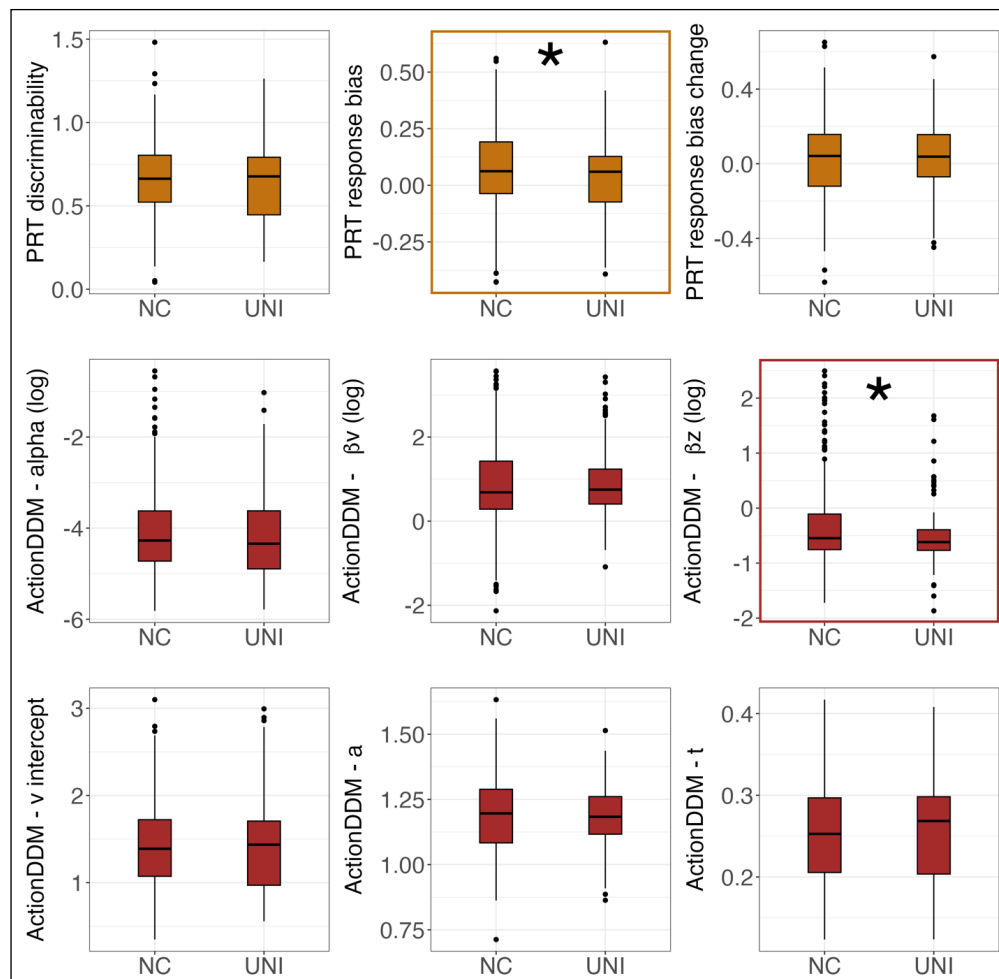


Figure 3 Comparisons of PRT behavioral parameters between non-psychiatric control group (NC) and unipolar depression group (UNI).

Note: PRT summary statistics and Action-DDM parameters are indicated in gold and red respectively.

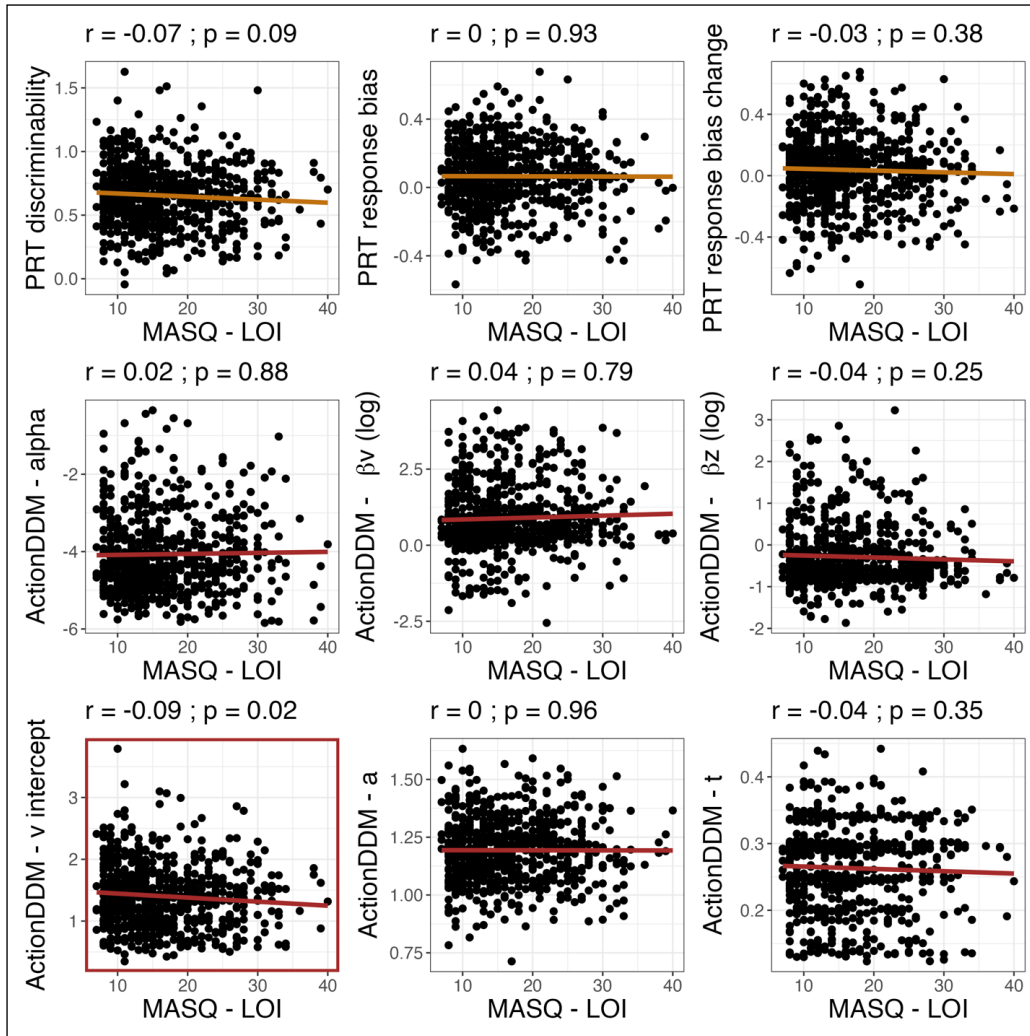


Figure 4 Correlations between anhedonic symptoms and PRT parameters.

Note: PRT summary statistics and Action-DDM parameters are indicated in gold and red respectively.

Exploratory subgroup analyses using a median split by age found that depressed adolescents (Sample 1) showed trending lower response bias ($t = -1.67$, $p = .097$), trending lower reward learning rate (α , $t = -1.67$, $p = .098$), and reduced adjustments in starting point bias (B_z , $t = -2.27$, $p = .02$) compared to control adolescents with no history of psychopathology. Among young adults (Sample 2), severity of anhedonic symptoms was negatively correlated with adjustments in starting point bias (B_z ; $r = -0.12$, $p = .04$) and no group-level differences emerged. In both samples, higher anhedonia severity was marginally associated with lower drift-rate intercept ($r_{\text{sample1}} = -0.09$, $r_{\text{sample2}} = -0.10$, $ps < .10$).

PARAMETER VALIDATION

Analyses calculated block1-block2 Spearman-Brown coefficients to quantify split-half parameter reliability. For PRT performance measures, Spearman-Brown coefficients were 0.61 (response bias), 0.81 (discriminability), and 0.82 (accuracy). For Action-DDM model parameters, parameter reliability was moderate for learning rate ($r_{\text{sb}} = 0.65$), B_v ($r_{\text{sb}} = 0.46$), and B_z ($r_{\text{sb}} = 0.67$), and strong for $v_{\text{intercept}}$ ($r_{\text{sb}} = 0.85$), decision threshold ($r_{\text{sb}} = 0.74$), and non-decision time ($r_{\text{sb}} = 0.86$). Intraclass correlations yielded comparable reliability estimates: 0.64 (learning rate), 0.35 (B_v), 0.64 (B_z), 0.86 ($v_{\text{intercept}}$), 0.75 (decision threshold), 0.88 (non-decision time).

Parameter recovery was excellent for $v_{\text{intercept}}$ ($r = 0.95$), decision threshold ($r = 0.95$), and non-decision time ($r = 0.99$), and adequate for learning rate ($r = 0.48$), B_v ($r = 0.52$), and B_z ($r = 0.54$).

This study developed and applied a new computational model to examine dimensions of reward processing using the PRT, and also tested associations between reward parameters and the diagnosis of depression or the severity of anhedonic symptoms. The model successfully captured response bias in both choices and response times, and provided a more integrated account of learning and decision-making processes in the PRT. Youth diagnosed with unipolar depression (especially adolescents) were characterized by blunted biases towards reward, evident in lower response biases and reduced influence of reward values on the starting point bias during decision making. In addition, adolescents and young adults with more severe anhedonia showed slower evidence accumulation. Together, these findings provide insights into the dynamics of learning and decision making in the PRT, and identify distinctive abnormalities associated with depression diagnoses and anhedonia severity.

The finding of reduced response bias, especially among adolescents, replicates past results that demonstrated deficits in reward responsiveness in youth with psychiatric disorders (Morris et al., 2015; Shen et al., 2024). Although recent studies have found that depression-related differences in response bias may not always emerge, they have nonetheless revealed differences in the underlying decision-making parameters (e.g., Lawlor et al., 2020; Dillon et al., 2022, 2024). Importantly, our computational model captured both reward learning and decision-making processes in the task, allowing us to more precisely examine how depression may differentially affect each component. The computational analysis revealed that reward hyposensitivity reflected weaker trial-by-trial adjustments to starting point bias. Such reward hyposensitivity was evident among adolescents diagnosed with unipolar depression, and transdiagnostically associated with higher anhedonic symptoms among young adults. Therefore, all youth learned from feedback and updated action values during the task, but those *without* depression developed a stronger predisposition toward the rich response, leading to faster and more frequent rich responses (i.e., stronger response bias).

Results also showed that more anhedonic youth showed slower evidence accumulation (drift rate intercept). Given the convergence between drift rate intercept and discriminability (level of task difficulty for an individual; see Supplement), this finding may reflect an association between symptom severity and goal-directed information processing. This interpretation is consistent with prior research linking depression with reduced evidence accumulation efficiency (Pitliya et al., 2022; Shen et al., 2024; Sripada & Weigard, 2021). Notably, we did not observe associations between depression or anhedonia and adjustments in drift rate based on reinforcement history. Therefore, *changes* in evidence accumulation speed based on differences in expected values were similar across individuals, but those reporting elevated anhedonia showed slower *overall* evidence accumulation.

Together, findings provided insight into reward processing abnormalities associated with depression during a developmental period of vulnerability to mood disorders. Results suggest that some forms of reward processing abnormality may be more closely related to diagnostic status, whereas others are more closely related to the severity of anhedonic depression. These findings have significant potential clinical implications, as they may help distinguish distinct profiles of reward processing variations and guide the development and targeting of early intervention programs tailored to specific cognitive profiles (Adams et al., 2016).

STUDY LIMITATIONS

First, the PRT is well-validated and widely used, but other tasks can capture additional dimensions of reward processing. In particular, previous research that reported slower reward learning in depressed individuals often employed traditional learning paradigms based on visual discrimination instead of signal detection (e.g., bandit tasks; Brown et al., 2021; Dombrovski et al., 2010). In addition, learning processes may also differ depending on the outcome types (Pike & Robinson, 2022); however, the current study cannot parse behavior based on outcome valence since the PRT only delivers rewards. Future studies may apply a battery of tasks to more

comprehensively and reliably assess aspects of reinforcement learning affected (or unaffected) by depression. Relatedly, the models assumed a fixed learning rate within an individual, but the cognitive strategies underlying task behavior may evolve across trials (Xiong et al., 2025). Such adaptations may reflect meta-learning mechanisms and future studies may model trial-by-trial fluctuations in learning parameters to better characterize how reinforcement learning processes evolve during the PRT.

Second, the current study yielded effect sizes in the small or small-medium range, consistent with prior work reporting on reward processing anomalies in depression (see systematic reviews of Keren et al., 2018 and Halahakoon et al., 2020). Although modest effect sizes are common in clinical research linking depression to neural (Story et al., 2008; Goldstein-Piekarski et al., 2022), genetic (Wichers et al., 2007; Tyrrell et al., 2019) and behavioral (Walsh et al., 2018; Hobbs et al., 2023) variables, combining information from multiple variables (e.g., polygenic risk scoring) may enhance clinical utility and more robustly predict depressive symptoms (Serretti, 2022). Future research may adopt similar approaches, using computational parameters to generate predictive models with improved clinical utility.

Third, it is unclear whether the observed reward processing abnormalities would fluctuate with depressive episode onsets or recoveries, or over the course of youth development. Exploratory analyses revealed that depressed adolescents showed marginally slower reward learning, an effect not observed in young adults with depression. Reduced adjustments in starting point bias were associated with diagnostic status in adolescents and with symptom severity in young adults. Although developmental hypotheses were beyond scope of the present study, results suggest that by isolating specific processes that contribute to behavior in the PRT, the computational model holds potential for refining early identification of reward processing deficits (Adams et al., 2016; Huys et al., 2016; Hauser et al., 2019). Future longitudinal research should investigate the changes of symptom-related or diagnosis-related reward processing abnormalities over the course of adolescent development.

CONCLUSION

This study developed a new computational model of RL and decision making in the PRT, and identified different dimensions of reward processing associated with depression diagnoses and anhedonic symptom severity. Future investigations may apply multi-task approaches to more comprehensively assess how depression affects reward processing and explore how reward processing abnormalities change throughout adolescent development.

DATA ACCESSIBILITY STATEMENT

With the consent/assent of our subjects, data from study sample 3 are available in the National Institute of Mental Health Data Archive (NDA) collection C3598. Participant permissions for data sharing were not achieved for other samples. Software for data analysis were documented in the methods section.

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Supplementary file.** Supplementary methods and Supplementary Results. DOI: <https://doi.org/10.5334/cpsy.147.s1>

FUNDING INFORMATION

Research supported by NARSAD Young Investigator Award 24879 (R.H.K.), NIMH grant MH117131 (R.H.K.), and NIHM grant MH111676 (D.G.D.).

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

Conceptualization: Z.C.

Methodology: Z.C., D.G.D., R.H.K., consultation by C.D.S and D.J.M.

Validation: Z.C.

Formal Analysis: Z.C.

Investigation: Z.C., A.D.M., J.J., under R.H.K.'s supervision, consultation by C.D.S and D.J.M.

Data Curation: Z.C., A.D.M., J.J., under R.H.K.'s supervision

Writing – Original Draft Preparation: Z.C.

Writing – Review & Editing: all co-authors

Supervision: R.H.K.

Project Administration: A.D.M., J.J.

Funding: R.H.K.

AUTHOR AFFILIATIONS

Ziwei Cheng  orcid.org/0009-0008-9247-0427

Department of Psychology, University of California Berkeley, Berkeley, CA, United States

Amelia D. Moser  orcid.org/0000-0002-9889-6102

Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, United States;
Institute of Cognitive Science, University of Colorado Boulder, Boulder, CO, United States

Jenna Jones

Department of Psychology, Virginia Polytechnic Institute and State University, Blacksburg, VA, United States

Christopher D. Schneck  orcid.org/0000-0002-7422-4707

Department of Psychiatry, University of Colorado Anschutz Medical Campus, Anschutz, CO, United States

David J. Miklowitz  orcid.org/0000-0002-9647-6147

Department of Psychiatry, University of California Los Angeles, Los Angeles, CA, United States

Daniel G. Dillon  orcid.org/0000-0002-1977-700X

Center for Depression, Anxiety and Stress Research, McLean Hospital, Belmont, MA, United States;
Department of Psychiatry, Harvard Medical School, Boston, MA, United States

Roselinde H. Kaiser  orcid.org/0000-0002-0369-0583

Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, United States;
Institute of Cognitive Science, University of Colorado Boulder, Boulder, CO, United States

REFERENCES

- Adams, R. A., Huys, Q. J., & Roiser, J. P.** (2016). Computational psychiatry: towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery & Psychiatry*, 87(1), 53–63.
- Admon, R., & Pizzagalli, D. A.** (2015). Dysfunctional reward processing in depression. *Current opinion in psychology*, 4, 114–118. <https://doi.org/10.1016/j.copsyc.2014.12.011>
- Brown, V. M., Zhu, L., Solway, A., Wang, J. M., McCurry, K. L., King-Casas, B., & Chiu, P. H.** (2021). Reinforcement Learning Disruptions in Individuals With Depression and Sensitivity to Symptom Change Following Cognitive Behavioral Therapy. *JAMA psychiatry*, 78(10), 1113–1122. <https://doi.org/10.1001/jamapsychiatry.2021.1844>
- Brown, W.** (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, 3(3), 296–322. <https://doi.org/10.1111/j.2044-8295.1910.tb00207.x>
- Diekhof, E. K., Falkai, P., & Gruber, O.** (2008). Functional neuroimaging of reward processing and decision-making: a review of aberrant motivational and affective processing in addiction and mood disorders. *Brain research reviews*, 59(1), 164–184. <https://doi.org/10.1016/j.brainresrev.2008.07.004>

- Dillon, D. G., Belleau, E. L., Origlio, J., McKee, M., Jahan, A., Meyer, A., Souther, M. K., Brunner, D., Kuhn, M., Ang, Y. S., Cusin, C., Fava, M., & Pizzagalli, D. A. (2024). Using Drift Diffusion and RL Models to Disentangle Effects of Depression On Decision-Making vs. Learning in the Probabilistic Reward Task. *Computational psychiatry* (Cambridge, Mass.), 8(1), 46–69. <https://doi.org/10.5334/cpsy.108>
- Dillon, D. G., Lazarov, A., Dolan, S., Bar-Haim, Y., Pizzagalli, D. A., & Schneier, F. R. (2022). Fast evidence accumulation in social anxiety disorder enhances decision making in a probabilistic reward task. *Emotion*, 22(1), 1–18. <https://doi.org/10.1037/emo0001053>
- Dombrowski, A. Y., Clark, L., Siegle, G. J., Butters, M. A., Ichikawa, N., Sahakian, B. J., & Szanto, K. (2010). Reward/Punishment reversal learning in older suicide attempters. *The American journal of psychiatry*, 167(6), 699–707. <https://doi.org/10.1176/appi.ajp.2009.09030407>
- Fassett-Carman, A. N., Moser, A. D., Ruzic, L., Neilson, C., Jones, J., Barnes-Horowitz, S., Schneck, C. D., & Kaiser, R. H. (2023). Amygdala and nucleus accumbens activation during reward anticipation moderates the association between life stressor frequency and depressive symptoms. *Journal of Affective Disorders*, 330, 309–318. <https://www.sciencedirect.com/science/article/pii/S0165032723003208>
- First, M. B., Williams, J. B., Karg, R. S., & Spitzer, R. L. (2015). *Structured clinical interview for DSM-5—Research version (SCID-5 for DSM-5, research version; SCID-5-RV)*. American Psychiatric Association (pp. 1–94).
- Forbes, E. E., & Dahl, R. E. (2012). Research Review: altered reward function in adolescent depression: what, when and how? *Journal of Child Psychology and Psychiatry*, 53(1), 3–15. <https://doi.org/10.1111/j.1469-7610.2011.02477.x>
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 457–472. <https://doi.org/10.1214/ss/1177011136>
- Goldstein-Piekarski, A. N., Ball, T. M., Samara, Z., Staveland, B. R., Keller, A. S., Fleming, S. L., Grisanzio, K. A., Holt-Gosselin, B., Stetz, P., Ma, J., & Williams, L. M. (2022). Mapping Neural Circuit Biotypes to Symptoms and Behavioral Dimensions of Depression and Anxiety. *Biological psychiatry*, 91(6), 561–571. <https://doi.org/10.1016/j.biopsych.2021.06.024>
- Grob, S., Pizzagalli, D. A., Dutra, S. J., Stern, J., Mörgeli, H., Milos, G., Schnyder, U., & Hasler, G. (2012). Dopamine-Related Deficit in Reward Learning After Catecholamine Depletion in Unmedicated, Remitted Subjects with Bulimia Nervosa. *Neuropsychopharmacology* (New York, N.Y.), 37(8), 1945–1952. <https://doi.org/10.1038/npp.2012.41>
- Halahakoon, D. C., Kieslich, K., O'Driscoll, C., Nair, A., Lewis, G., & Roiser, J. P. (2020). Reward-Processing Behavior in Depressed Participants Relative to Healthy Volunteers: A Systematic Review and Meta-analysis. *JAMA psychiatry*, 77(12), 1286–1295. <https://doi.org/10.1001/jamapsychiatry.2020.2139>
- Hauser, T. U., Will, G. J., Dubois, M., & Dolan, R. J. (2019). Annual research review: developmental computational psychiatry. *Journal of Child psychology and Psychiatry*, 60(4), 412–426. <https://doi.org/10.1111/jcpp.12964>
- Hobbs, C., Sui, J., Kessler, D., Munafò, M. R., & Button, K. S. (2023). Self-processing in relation to emotion and reward processing in depression. *Psychological medicine*, 53(5), 1924–1936. <https://doi.org/10.1017/S0033291721003597>
- Huys, Q. J., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature neuroscience*, 19(3), 404–413. <https://doi.org/10.1038/nn.4238>
- Huys, Q. J., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biology of mood & anxiety disorders*, 3(1), 1–16. <https://doi.org/10.1186/2045-5380-3-12>
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *The American journal of psychiatry*, 167(7), 748–751. <https://doi.org/10.1176/appi.ajp.2010.09091379>
- Kaiser, R. H., Moser, A. D., Neilson, C., Peterson, E. C., Jones, J., Hough, C. M., Rosenberg, B. M., Sandman, C. F., Schneck, C. D., Miklowitz, D. J., & Friedman, N. P. (2022). Mood Symptom Dimensions and Developmental Differences in Neurocognition in Adolescence. *Clinical Psychological Science*, 11(2), 308–325. <https://doi.org/10.1177/21677026221111389>
- Kaiser, R. H., Treadway, M. T., Wooten, D. W., Kumar, P., Goer, F., Murray, L., Beltzer, M., Pechtel, P., Whitton, A., Cohen, A. L., Alpert, N. M., El Fakhri, G., Normandin, M. D., & Pizzagalli, D. A. (2018). Frontostriatal and Dopamine Markers of Individual Differences in Reinforcement Learning: A Multi-modal Investigation. *Cerebral cortex* (New York, N.Y.: 1991), 28(12), 4281–4290. <https://doi.org/10.1093/cercor/bhx281>
- Kangas, B. D., Der-Avakian, A., & Pizzagalli, D. A. (2022). Probabilistic reinforcement learning and anhedonia. In *Anhedonia: Preclinical, Translational, and Clinical Integration* (pp. 355–377). Springer International Publishing. https://doi.org/10.1007/7854_2022_349

- Keren, H., O'Callaghan, G., Vidal-Ribas, P., Buzzell, G. A., Brotman, M. A., Leibenluft, E., Pan, P. M., Meffert, L., Kaiser, A., Wolke, S., Pine, D. S., & Stringaris, A. (2018). Reward Processing in Depression: A Conceptual and Meta-Analytic Review Across fMRI and EEG Studies. *The American journal of psychiatry*, 175(11), 1111–1120. <https://doi.org/10.1176/appi.ajp.2018.17101124>
- Lawlor, V. M., Webb, C. A., Wiecki, T. V., Frank, M. J., Trivedi, M., Pizzagalli, D. A., & Dillon, D. G. (2020). Dissecting the impact of depression on decision-making. *Psychological medicine*, 50(10), 1613–1622. <https://doi.org/10.1017/S0033291719001570>
- Letkiewicz, A. M., Cochran, A. L., Mittal, V. A., Walther, S., & Shankman, S. A. (2022). Reward-based reinforcement learning is altered among individuals with a history of major depressive disorder and psychomotor retardation symptoms. *Journal of psychiatric research*, 152, 175–181. <https://doi.org/10.1016/j.jpsychires.2022.06.032>
- Luking, K. R., Pagliaccio, D., Luby, J. L., & Barch, D. M. (2016). Reward processing and risk for depression across development. *Trends in cognitive sciences*, 20(6), 456–468. <https://doi.org/10.1016/j.tics.2016.04.002>
- Morris, B. H., Bylsma, L. M., Yaroslavsky, I., Kovacs, M., & Rottenberg, J. (2015). Reward learning in pediatric depression and anxiety: preliminary findings in a high-risk sample. *Depression and anxiety*, 32(5), 373–381. <https://doi.org/10.1002/da.22358>
- Pedersen, M. L., & Frank, M. J. (2020). Simultaneous hierarchical bayesian parameter estimation for reinforcement learning and drift diffusion models: a tutorial and links to neural data. *Computational Brain & Behavior*, 3(4), 458–471. <https://doi.org/10.1007/s42113-020-00084-w>
- Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic bulletin & review*, 24(4), 1234–1251. <https://doi.org/10.3758/s13423-016-1199-y>
- Peterson, E. C., Rosenberg, B. M., Hough, C. M., Sandman, C. F., Neilson, C., Miklowitz, D. J., & Kaiser, R. H. (2021). Behavioral mediators of stress-related mood symptoms in adolescence & young adulthood. *Journal of Affective Disorders*, 294, 94–102. <https://doi.org/10.1016/j.jad.2021.06.079>
- Peterson, E. C., Snyder, H. R., Neilson, C., Rosenberg, B. M., Hough, C. M., Sandman, C. F., Ohanian, L., Garcia, S., Kotz, J., Finegan, J., Ryan, C. A., Gyimah, A., Sileo, S., Miklowitz, D. J., Friedman, N. P., & Kaiser, R. H. (2022). General and Specific Dimensions of Mood Symptoms Are Associated With Impairments in Common Executive Function in Adolescence and Young Adulthood. *Frontiers in human neuroscience*, 16, 838645. <https://doi.org/10.3389/fnhum.2022.838645>
- Pike, A. C., & Robinson, O. J. (2022). Reinforcement learning in patients with mood and anxiety disorders vs control individuals: A systematic review and meta-analysis. *JAMA psychiatry*, 79(4), 313–322. <https://doi.org/10.1001/jamapsychiatry.2022.0051>
- Pitliya, R. J., Nelson, B. D., Hajcak, G., & Jin, J. (2022). Drift-Diffusion Model Reveals Impaired Reward-Based Perceptual Decision-Making Processes Associated with Depression in Late Childhood and Early Adolescent Girls. *Research on Child and Adolescent Psychopathology*, 50(11), 1515–1528. <https://doi.org/10.1007/s10802-022-00936-y>
- Pizzagalli, D. A., Iosifescu, D., Hallett, L. A., Ratner, K. G., & Fava, M. (2008). Reduced hedonic capacity in major depressive disorder: evidence from a probabilistic reward task. *Journal of psychiatric research*, 43(1), 76–87. <https://doi.org/10.1016/j.jpsychires.2008.03.001>
- Pizzagalli, D. A., Jahn, A. L., & O'Shea, J. P. (2005). Toward an objective characterization of an anhedonic phenotype: A signal-detection approach. *Biological Psychiatry*, 57(4), 319–327. <https://doi.org/10.1016/j.biopsych.2004.11.026>
- Ratcliff, R., & McKoon, G. (2018). Modeling numerosity representation with an integrated diffusion model. *Psychological review*, 125(2), 183. <https://doi.org/10.1037/rev0000085>
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological science*, 9(5), 347–356. <https://doi.org/10.1111/1467-9280.00067>
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. K. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.
- Rupprechter, S., Stankevicius, A., Huys, Q. J., Steele, J. D., & Seriès, P. (2018). Major depression impairs the use of reward values for decision-making. *Scientific reports*, 8(1), 1–8. <https://doi.org/10.1038/s41598-018-31730-w>
- Safra, L., Chevallier, C., & Palminteri, S. (2019). Depressive symptoms are associated with blunted reward learning in social contexts. *PLoS computational biology*, 15(7), e1007224. <https://doi.org/10.1371/journal.pcbi.1007224>
- Serretti, A. (2022). Clinical Utility of Fluid Biomarker in Depressive Disorder. *Clinical psychopharmacology and neuroscience: the official scientific journal of the Korean College of Neuropsychopharmacology*, 20(4), 585–591. <https://doi.org/10.9758/cpn.2022.20.4.585>

- Shen, L., Hu, Y. X., Lv, Q. Y., Yi, Z. H., Gong, J. B., & Yan, C. (2024). Using hierarchical drift diffusion models to elucidate computational mechanisms of reduced reward sensitivity in adolescent major depressive disorder. *BMC Psychiatry*, 24, 933. <https://doi.org/10.1186/s12888-024-06353-3>
- Spearman, C. C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 3(3), 271–295. <https://doi.org/10.1111/j.2044-8295.1910.tb00206.x>
- Sripada, C., & Weigard, A. (2021). Impaired evidence accumulation as a transdiagnostic vulnerability factor in psychopathology. *Frontiers in psychiatry*, 12, 627179. <https://doi.org/10.3389/fpsy.2021.627179>
- Stan Development Team. (2022). “RStan: the R interface to Stan.” R package version 2.21.7. <https://mc-stan.org/>.
- Story, T. J., Potter, G. G., Attix, D. K., Welsh-Bohmer, K. A., & Steffens, D. C. (2008). Neurocognitive Correlates of Response to Treatment in Late-Life Depression. *The American Journal of Geriatric Psychiatry*, 16(9), 752–759. <https://doi.org/10.1097/JGP.0b013e31817e739a>
- Treadway, M. T., & Zald, D. H. (2013). Parsing anhedonia: translational models of reward-processing deficits in psychopathology. *Current directions in psychological science*, 22(3), 244–249. <https://doi.org/10.1177/0963721412474460>
- Tyrrell, J., Mulugeta, A., Wood, A. R., Zhou, A., Beaumont, R. N., Tuke, M. A., Jones, S. E., Ruth, K. S., Yaghootkar, H., Sharp, S., Thompson, W. D., Ji, Y., Harrison, J., Freathy, R. M., Murray, A., Weedon, M. N., Lewis, C., Frayling, T. M., & Hyppönen, E. (2019). Using genetics to understand the causal influence of higher BMI on depression. *International journal of epidemiology*, 48(3), 834–848. <https://doi.org/10.1093/ije/dyy223>
- Vrieze, E., Pizzagalli, D. A., Demyttenaere, K., Hompes, T., Sienaert, P., de Boer, P., Schmidt, M., & Claes, S. (2013). Reduced Reward Learning Predicts Outcome in Major Depressive Disorder. *Biological Psychiatry*, 73(7), 639–645. <https://doi.org/10.1016/j.biopsych.2012.10.014>
- Wabersich, D., & Vandekerckhove, J. (2014). The RWiener package: An R package providing distribution functions for the Wiener diffusion model.
- Walsh, A. E. L., Browning, M., Drevets, W. C., Furey, M., & Harmer, C. J. (2018). Dissociable temporal effects of bupropion on behavioural measures of emotional and reward processing in depression. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 373(1742), 20170030. <https://doi.org/10.1098/rstb.2017.0030>
- Watson, D., Clark, L. A., Weber, K., Assenheimer, J. S., Strauss, M. E., & McCormick, R. A. (1995b). Testing a tripartite model: II. Exploring the symptom structure of anxiety and depression in student, adult, and patient samples. *J. Abnorm. Psychol.* 104(1), 15–25. <https://doi.org/10.1037//0021-843x.104.1.15>
- Watson, D., Weber, K., Assenheimer, J. S., Clark, L. A., Strauss, M. E., & McCormick, R. A. (1995a). Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. *J. Abnorm. Psychol.* 104(1), 3–14. <https://doi.org/10.1037//0021-843x.104.1.3>
- White, C. N., & Poldrack, R. A. (2014). Decomposing bias in different types of simple decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(2), 385–398. <https://doi.org/10.1037/a0034851>
- Whitton, A. E., Reinen, J. M., Slifstein, M., Ang, Y.-S., McGrath, P. J., Iosifescu, D. V., Abi-Dargham, A., Pizzagalli, D. A., & Schneier, F. R. (2020). Baseline reward processing and ventrostriatal dopamine function are associated with pramipexole response in depression. *Brain (London, England: 1878)*, 143(2), 701–710. <https://doi.org/10.1093/brain/awaa002>
- Wichers, M., Myin-Germeyns, I., Jacobs, N., Peeters, F., Kenis, G., Derom, C., Vlietinck, R., Delespaul, P., & van Os, J. (2007). Genetic risk of depression and stress-induced negative affect in daily life. *The British Journal of Psychiatry*, 191(3), 218–223. <https://doi.org/10.1192/bjp.bp.106.032201>
- Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the drift-diffusion model in Python. *Frontiers in neuroinformatics*, 14. <https://doi.org/10.3389/fninf.2013.00014>
- Xiong, H. D., LI, J., Mattar, M. G., & Wilson, R. (2025). DynamicRL: Data-Driven Estimation of Trial-by-Trial Reinforcement Learning Parameters. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 47). https://doi.org/10.31234/osf.io/4xumc_v2
- Zald, D. H., & Treadway, M. T. (2017). Reward processing, neuroeconomics, and psychopathology. *Annual review of clinical psychology*, 13, 471–495. <https://doi.org/10.1146/annurev-clinpsy-032816-044957>

TO CITE THIS ARTICLE:

Cheng, Z., Moser, A. D., Jones, J., Schneck, C. D., Miklowitz, D. J., Dillon, D. G., & Kaiser, R. H. (2025). Reinforcement Learning and Decision Making in Depression in Adolescents and Young Adults: Insights from a New Model of the Probabilistic Reward Task. *Computational Psychiatry*, 9(1), pp. 268–283. DOI: <https://doi.org/10.5334/cpsy.147>

Submitted: 19 April 2025

Accepted: 24 November 2025

Published: 30 December 2025

COPYRIGHT:

© 2025 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Computational Psychiatry is a peer-reviewed open access journal published by Ubiquity Press.