

SUPPLEMENT

Anxiety Impedes Adaptive Social Learning Under Uncertainty

Amrita Lamba, Michael J. Frank, Oriel FeldmanHall

Supplemental Methods

Assessing Anxiety. Subjects were grouped as a function of clinical-significance depending on their scores from the 7-item Generalized-Anxiety Disorder Scale, the GAD-7. Based on clinical guidelines (Spitzer, Kroenke, Williams, & Löwe, 2006), a score of 10+ on the GAD-7 scale reliably predicts the prevalence of an underlying anxiety disorder (89% post-test predictive probability) and is therefore considered clinically-significant. Assessed clinical-significance in our study does not mean that subjects were already clinically-diagnosed based on DSM-5 standards, however, clinically-significant scores strongly suggest that participants experience many of the characteristic symptoms associated with GAD (Löwe et al., 2008). GAD-7 scores are displayed in Figure S1 below.

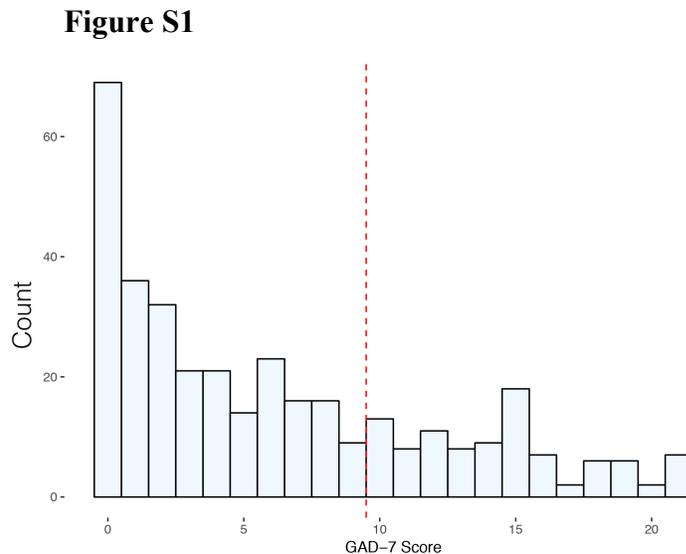


Figure S1. Distribution of GAD-7 scores from study sample ($N = 354$). The dotted red line separates the sample into sub-threshold (healthy) and clinically-significant (anxious) groups with respect to clinical guidelines (score of 10+ considered clinically significant). *Mean GAD-7 score* = 6.16, *SE* = 0.31. $N_{\text{healthy}} = 257$, $N_{\text{anxious}} = 97$.

Treating Anxiety as a Continuous Metric. As noted above, subjects were grouped together depending on whether they were above or below clinically-established thresholds on the GAD-7 scale. Although investigating differences in task performance across groups is a useful way to parse our study hypotheses, we wanted to ensure that our effects also hold when treating anxiety as a continuous construct. In order to do this, we modeled our data using raw GAD-7 scores (i.e. scores shown in Figure S1) and the results still hold. When examining the mean investments across task blocks as a function of GAD-7 score (treated continuously: Investment ~ GAD Continuous * Valence) the anxiety × valence interactions are still significant and follow the same pattern of results observed in the original analyses where GAD-7 is modeled dichotomously. Below are observed statistical effects of linear mixed effects regressions (lme4 package in R) using raw GAD-7 scores (i.e. treated continuously); TG neutral start player, anxiety × valence interaction, $t(351) = 2.39, p = 0.017$; TG positive start player, anxiety × valence interaction, $t(351) = 2.62, p = 0.009$; SM neutral start player, anxiety × valence interaction, $t(351) = 2.53, p = 0.012$; SM positive start player, anxiety × valence interaction, $t(351) = 2.36, p = 0.019$. As observed with dichotomous analyses, no significant anxiety × valence interactions were observed in the TG or SM negative start cases.

When we examine the decay rate difference with GAD as a continuous variable (Decay rate ~ GAD Continuous * Condition), we observe a significant main effect of GAD ($t(687.88) = 2.49, p = 0.013$), and condition ($t(352) = 3.36, p < 0.001$), such that there is an observed Trust Game x GAD effect ($t(352) = 2.64, p = .009$) but no Slot Machine x GAD effect ($p > .05$), mirroring the data from figure 4b. There is not, however, a GAD x condition interaction ($t(352) = -1.39, p = 0.165$) which may be the result of the model being slightly underpowered to detect an effect considering how variance is accounted for in the continuous mixed model regression (there are only two decay rate data points per subject).

Excluded Subjects. Exclusion criteria were twofold: those who demonstrated poorer learning than chance and those who clicked through the entire experiment by indicating the same response on all trials of one or both tasks, which was assessed through model fit values (AICs) that were not better than those derived from chance behavior ($-2 * \log(.5) * (\text{No. trials}) = -116.45$), or those who had very close to 0 values (i.e. perfectly predictable), which were diagnostic of highly deterministic and repetitive behavior (e.g. investing \$1.00 on all 84 trials of the task). These extremely repetitive behaviors could be captured in the model through reaching upper-bound on the inverse temperature parameter (i.e. highly deterministic behavior), together with upper and lower-bounds of the bias parameter (depending on whether the subject invested on every trial), and lower bounds on decay parameters (i.e. no updating of information).

Therefore, to exclude subjects with worse than chance performance and those who simply did not engage in the task as evidenced through repetitious responses on all trials, we restricted our sample to subjects with AIC values > -116.45 (indicating performance at or better than chance) and $<$ the model AIC upper bound (indicating variability in trial by trial responses). Therefore, using AIC values derived from the simplified B-RL model across both trust and slot machine tasks, we excluded data from 58 subjects ($n = 20, \text{AIC}_{\text{B-RL}} < -116.45$ & $n = 38, \text{AIC}_{\text{B-RL}} = -12.00$) for a final sample size of $N = 354$. We chose to exclude subjects based on simplified B-RL model fit because simplified B-RL was identical to DB-RL but had fewer parameters, and

therefore a slightly less strict benchmark for chance performance. Because our experiment requires within-subject analyses across tasks, subjects who exhibited worse-than-chance or extremely repetitive behavior in only one task were still excluded from the study, in order to ensure pairwise comparisons were feasible. Model fits of all subjects prior to exclusion (N = 412) are plotted below.

$$AIC = 2*(\log \text{likelihood}) - 2*(\text{No. model parameters})$$

As a proportion of the total sample, the exclusion criteria removed 5.10% of subjects that were below chance and also had clinically-sig. anxiety, compared to 8.74% that were below chance but were healthy. Therefore, anxious and healthy subjects were removed at similar rates. The excluded group (N = 58) did not have significantly higher anxiety scores compared to the final sample (Mean GAD-7 in final sample: 6.16; Mean GAD-7 of excluded group: 7.63; $t(410) = -1.73, p > 0.05$).

Figure S2

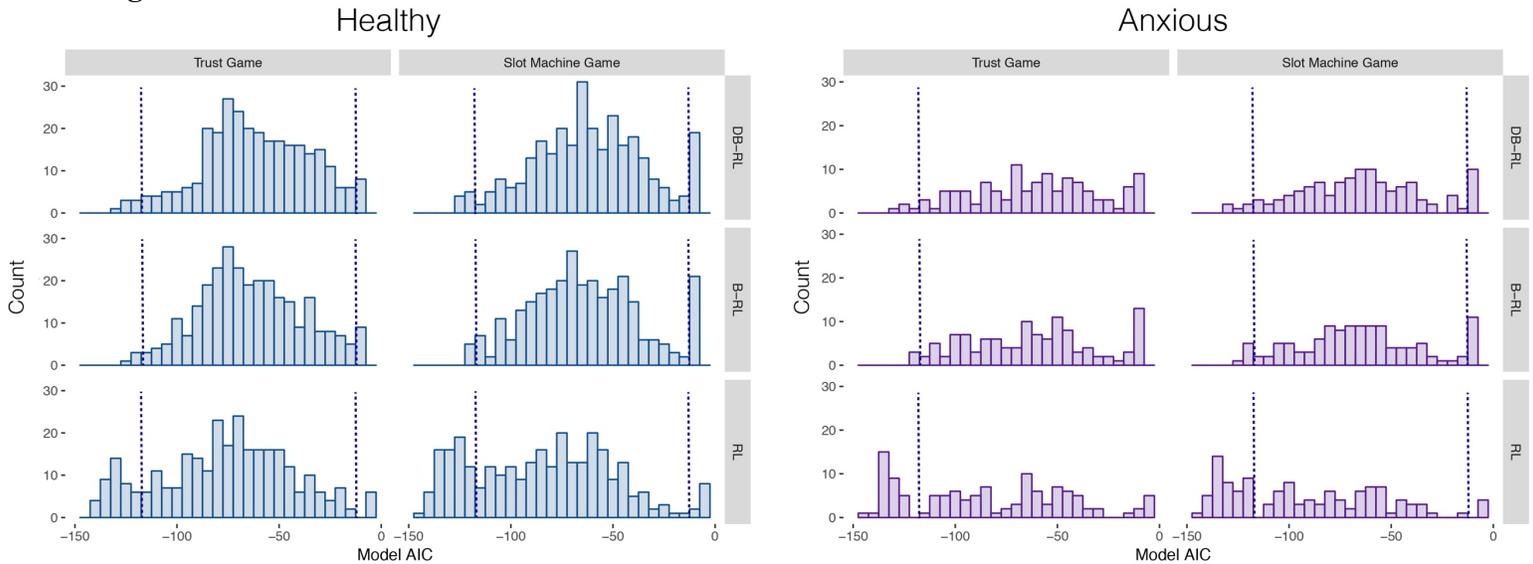


Figure S2. Distribution of AIC values across models and healthy vs. anxious groups. The far-left dotted blue line in each panel demarcates model fit values above and below chance performance. Subjects that fell below the chance threshold corresponding to the simplified B-RL model were excluded from all further analyses. The far-right dotted blue line denotes the upper-bound of model AIC in which subjects were perfectly fit to the model through repetitive behavior, therefore identifying participants that indicated the same response on nearly all trials of the task.

Believability Ratings: We wanted to be extra cautious that subjects were treating the Trust Game as a fundamentally social task and therefore measured the extent to which subjects believed they were interacting with real partners online. At the end of the experiment, subjects were asked to indicate how much they believed they were interacting with real partners. 70% reported no doubt to only some doubt that their partners were real. As a more conservative test, we wanted to ensure that believability ratings (reported on a scale from 1-6) did not impact our observed statistical results. We entered subject-specific believability ratings into our main regressions, which controls for the effects of task suspicion. Entering each subject's believability score into the main models did not quantitatively change any of the relevant statistics reported in the main manuscript or supplement (the betas associated with believability scores for all analyses were non-significant, $p > 0.29$ for all).

Modeling Decision Space. Although subjects could in principle invest any amount from 0.10 to 1.00 on each trial, the optimal payoff maximizing strategy is to invest the full amount (\$1.00) if the trustee returns more than 0.25 the amount (after it is multiplied by 4), or to invest the minimum (\$0.10) if the trustee returns less than that. Indeed, the distribution of investments in both healthy and anxious subjects were bimodal at \$0.10 and \$1.00 (Figure S3), suggesting that most subjects learned to use these anchors to maximize their earnings. We thus binarized the decision space to predict whether subjects would invest or not: when the player was set to return more than 0.25, high investments (at or above \$0.50) were considered optimal, whereas low investments (below \$0.50) were considered approximately optimal when the player-algorithm was set below 0.25. This was further validated by the observation that subjects were best fit by a model in which they would choose \$1.00 if the probability of the other person repaying their trust (i.e. that the return was greater than 0.25) was greater than 50%, as quantified by the bias parameter (see Tables S5 and S6, Bias $\psi \sim 0.5$ in all cases).

Due to the 4% noise boundary, subjects rarely broke exactly even (i.e. the randomly sampled proportion of return always fell on one side of the 0.25 threshold) and subjects generally earned or lost at least \$0.01 on the 0.25 proportion return trials. Although conceptually these trials should be treated as breaking even, the effects produced from the 4% noise boundary essentially cancelled out rewards and losses during the break-even blocks, so value was generally kept stable in these periods.

Figure S3

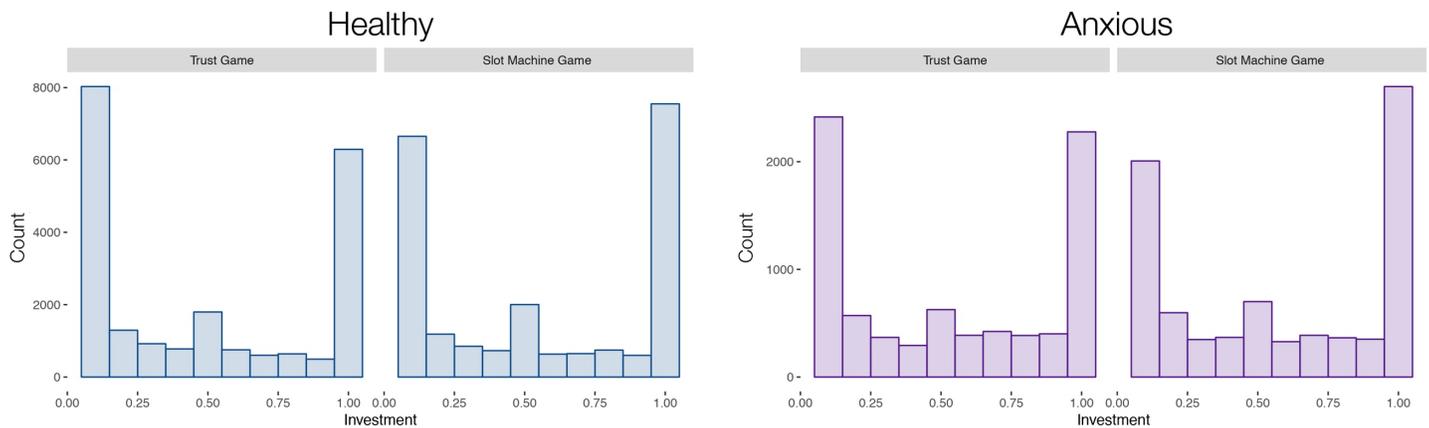


Figure S3. Distribution of investments across the Trust Game and slot machine task, broken out by healthy vs. anxious groups ($N = 354$, study sample). As shown in the above distributions, responses were highly bimodal at \$0.10 and \$1.00. Healthy Group: $\text{Mean}_{\text{TG}} = 0.49$, $\text{SE}_{\text{TG}} = 0.024$, $\text{Mean}_{\text{SM}} = 0.55$, $\text{SE}_{\text{SM}} = 0.024$. Anxious Group: $\text{Mean}_{\text{TG}} = 0.53$, $\text{SE}_{\text{TG}} = 0.038$, $\text{Mean}_{\text{SM}} = 0.58$, $\text{SE}_{\text{SM}} = 0.038$.

Modeling Feedback. Because feedback was presented on a continuous scale in our experiment, we also examined a continuous learning rule in which the value of investing \$1.00 or \$0.10 was updated as a function of scaled prediction errors (i.e. the outcome relative to their actual investment). Continuous versions of both the Bayesian and RL model did not adequately capture the data and resulted in poorer fits to the data, as indexed through the poorer model fits and posterior predictive checks. In other words, subjects were effectively treating the outcomes as more dichotomous in nature, despite the fact that feedback was continuous. This strategy is roughly normative, given that the payoff-maximizing strategy is to fully invest if the outcomes are above the threshold needed to increase the expected value, and to minimally invest otherwise. We therefore simplified the model, such that feedback was coded in terms of the payoff-maximizing strategy (i.e., whether the outcome indicated that investing would be beneficial or not). In other words, if the agent (or machine) was set to return more than 0.25 of the investment, the model iteratively coded this as gain trial, whereas any investment below 0.25 was coded as a loss trial. All subsequent models were fit using the binarized learning rule.

Dynamic Bayesian-RL Model

Model parameters. There were 6 free parameters that were fit to each subject by maximizing the log likelihood across 5 iterations using the *fmincon* function in MATLAB R2017a. Model parameters are described below.

Table S1: Dynamic Bayesian-RL model parameters

<i>Model Parameter</i>	<i>Parameter Description</i>	<i>Upper-bound</i>	<i>Lower-bound</i>
<i>Inverse Temperature</i> ζ	Degree to which subject exploits learned decision-rule associated with rewarding outcomes through deterministic vs. exploratory behavior	20 - highly deterministic behavior (always repeats action previously associated with reward)	1 - always selects at random
<i>Bias</i> ψ	Benchmark for investing (i.e. how much additional value does the subject need to derive from investing \$0.10 in order to select the choice over a preferred strategy to always invest \$1.00)	1.0 - high benchmark for investing (subject is biased towards never investing \$1.00)	0.1 – low benchmark for investing (subject is biased towards always investing \$1.00)
<i>Negative Decay Intercept</i> γ_{0neg}	Degree of decay of negative outcomes, which prevents the posterior distribution from becoming overly confident (certain) from previous experience	1.0 – no decay (posterior distribution is updated from negative outcomes and previous experiences accumulate)	0.1 – maximal decay (added uncertainty prevents posterior distribution from integrating observed negative outcomes over experience)
<i>Positive Decay Intercept</i> γ_{0pos}	Degree of decay or uncertainty over positive outcomes	1.0– no decay (posterior distribution is updated from positive outcomes and previous experiences accumulate)	0.1 – maximal decay (added uncertainty prevents posterior distribution from integrating observed positive outcomes)
<i>Negative Dynamic Decay</i> γ_{1neg}	Extent to which negative outcomes are decayed as a function of change points in uncertainty	0— no modulation of decay rate as a function of changes in entropy (ΔH)	-2.0—maximal adjustment of decay rate as a function of entropy (ΔH)
<i>Positive Dynamic Decay</i> γ_{1pos}	Extent to which positive outcomes are decayed as a function of change points in uncertainty	0 – no modulation of decay rate as a function of changes in entropy (ΔH)	-2.0 —maximal adjustment of decay rate as a function of entropy (ΔH)

Model setup. As noted above we assumed that subjects tracked the probability that it was worth investing in the trustee (i.e., that their proportion of return > 0.25). The optimal behavior for whether to invest only considers whether the outcome is better than the indifference point, and hence we track the probability of the trustee returning amounts that are larger than 0.25, rather than tracking the specific proportion returned (models that considered more continuous representations did not fit as well).

$$p(\theta | \text{return}_{t_1} \dots \text{return}_{t_n}) = \frac{p(\text{return}_{t_1} \dots \text{return}_{t_n} | \theta) \times p(\theta)}{p(\text{return}_{t_1} \dots \text{return}_{t_n})}$$

Given that the likelihood of any observation given the underlying rate of return is Bernoulli, $p(\text{return}_{t_1} \dots \text{return}_{t_n} | \theta)$, we modeled the belief about this probability, $p(\theta)$, as the conjugate, a beta distribution (Daw, Niv, & Dayan, 2005; Doll, Jacobs, Sanfey, & Frank, 2009; Frank, Doll, Oas-Terpstra, & Moreno, 2009; Franklin & Frank, 2015). Thus, each time the proportion return is greater than 0.25 (i.e. there is a positive prediction error) alpha was incremented +1, whereas each time there was a negative prediction error beta was incremented +1. Alpha and beta values were used to update the posterior distribution at the end of each trial. Effectively, this approach tracks the probability of a reward prediction error rather than tracking accumulated rewards over time. Each player (and machine) was individually modeled in separate distributions and all priors were initialized using the beta distribution conjugate prior (beta 1,1; i.e. uniform), as we opted for a more conservative approach in modeling subject-specific priors. Additionally, we did not have any strong predictions regarding subject-specific priors pertaining to the Trust Game, since subjects were told they would be engaging with anonymous others online. Below are the equations for the mean (μ_{jt}) and variance (σ^2_{jt}) of the posterior Beta distribution, where j corresponds to the player type (or machine type) and t denotes the current trial.

$$\mu_{jt} = \left(\frac{\alpha}{\alpha + \beta} \right) \text{ and } \sigma^2_{jt} = \left(\frac{\alpha * \beta}{(\alpha + \beta)^2 * (\alpha + \beta + 1)} \right)$$

The alpha and beta values were therefore used to keep a running track of positive and negative outcomes (α, β), which are used to update the posterior distribution on a trial-by-trial basis, thus approximating the probability that it is worth investing in the trustee.

Choice rule. On each trial of the TG and SM, choice probabilities were modeled through a softmax logistic function, comparing the mean value of the posterior distribution (i.e. the best estimate of whether they are likely to return the investment), where ζ and ψ are inverse temperature and bias parameters, respectively. Note that if the subject is optimal the bias parameter should be 0.5, which would indicate that they only choose to invest \$1.00 when the probability of doing so is greater than 50%. Nevertheless, we allowed for this bias parameter to be freely estimated, to allow for subject-specific biases in overall investment. Here, we will represent the probability of trusting \$1.00 or \$0.10 as $p(\$1.00)$ and $p(\$0.10)$, respectively.

$$p(\$1.00) = \frac{e^{\zeta * \mu_{jt}}}{e^{\zeta * \mu_{jt}} + e^{\zeta * \psi}}$$

$$p(\$0.10) = 1 - p(\text{trust } \$1.00)$$

Entropy. In order to index uncertainty sensitivity, we modeled the entropy (H) in the choice policy to summarize the uncertainty as to whether one should invest or not, as follows:

$$H_t = -[p(\$1.00) \times \log_2(p(\$1.00)) - p(\$0.10) \times \log_2(p(\$0.10))]$$

We then kept track of how this entropy changes from one trial to the next $\Delta H = H_t - H_{t-1}$ to capture whether there is likely to be a change point, such that this change in entropy could be used to further increase uncertainty and thereby speed up learning (Franklin & Frank, 2015).

Decay. When observed outcomes are not decayed ($\gamma = 1$), the beta distribution is maximally updated through the alpha and beta hyperparameters. In other words, the posterior distribution will reflect the entire history of actions taken by the trustee. When observed outcomes are decayed ($\gamma < 1$), the added uncertainty prevents the posterior distribution from becoming overly confident (i.e. it allows for the possibility that earlier observations are no longer relevant) and thus increases opportunity for learning flexibility. The γ_0 parameters correspond to the overall decay of previous outcomes. The γ_1 captures the extent to which decay is modulated on a trial-by-trial basis as a function of changes in the entropy (ΔH) and therefore indexes the extent to which one's learning rule is selectively adjusted through perceived changes in task-level uncertainty (Franklin & Frank, 2015). Critically, both γ_0 and γ_1 parameters jointly impact overall decay (γ) of alpha and beta hyperparameters. All gamma parameters were further partitioned by valence (positive vs. negative outcomes) to measure valence-dependent differences in learning.

$$\begin{aligned} \text{logit}(\gamma_{pos}) &= \gamma_{0_{pos}} + \gamma_{1_{pos}} \cdot \Delta H \\ \text{logit}(\gamma_{neg}) &= \gamma_{0_{neg}} + \gamma_{1_{neg}} \cdot \Delta H \end{aligned}$$

$$\begin{aligned} \alpha_{t+1} &= \alpha_t \cdot \gamma_{pos} \\ \beta_{t+1} &= \beta_t \cdot \gamma_{neg} \end{aligned}$$

Simplified Bayesian-RL Model

Model parameters. A total of 4 free parameters across 5 iterations were fit to each subject. See Table S1 for parameter descriptions. The simplified B-RL model was the same in all respects to DB-RL, except that the model omitted $\gamma_{1_{pos}}$ and $\gamma_{1_{neg}}$ parameters.

Decay. Decay parameters were not modeled as a function of changes in task entropy (ΔH). Thus, only the decay intercepts were used to decay alpha and beta parameters. This model therefore allows for the possibility that reward statistics might change (due to overall decay rates) but does not adaptively alter this perception with changes in uncertainty of the trustee's behavior. Alpha and beta were updated as shown below.

$$\begin{aligned} \alpha_{t+1} &= \alpha_t \cdot \gamma_{pos} \\ \beta_{t+1} &= \beta_t \cdot \gamma_{neg} \end{aligned}$$

Reinforcement Learning Model

Model parameters. In the RL model 3 free parameters across 5 iterations were fit to each subject.

Table S2: RL model parameters

<i>Model parameter</i>	<i>Parameter description</i>	<i>Upper-bound</i>	<i>Lower-bound</i>
<i>Inverse temperature ζ</i>	See description in Table S1	20 - highly deterministic behavior (always repeats action previously associated with reward)	1 - always selects at random
<i>Bias ψ</i>	See description in Table S1	1.0 - high benchmark for investing (subject is biased towards never investing \$1.00)	0.1 – low benchmark for investing (subject is biased towards always investing \$1.00)
<i>Learning rate λ</i>	Degree of to which positive and negative prediction errors update the value function.	0.1 – minimal update of V	1.0 – maximal update of V

As a benchmark-comparison model we constructed an RL model derived from the optimal (i.e. payoff-maximizing) learning rule in the task. The model was set up to separately track the reward-statistics of each player through the value function, where value (V , for each player $_{i;j}$ tracked at each time-step $_{i;t}$) was calculated as the difference between actual reward $_t$ (i.e. return) from player $_j$ and anticipated reward (V_{jt}), weighted by the learning rate (λ), which was fit to each subject. Value was initialized at 0.5, as we assumed participants would begin the task treating players indiscriminately.

$$V_{jt} = \lambda * (\text{reward}_t - V_{jt})$$

For the choice function, denoted below as the probability of trusting, we again leveraged the optimal choice rule for the task. Because the payoff maximizing response on a given trial was always either \$0.10 or \$1.00, subject choices were modeled using these values. Here, the probability of investing the full \$1.00 was indexed by $p(\$1.00)$ and the probability of investing the minimum \$0.10 was indexed as $p(\$0.10)$. Subject choices on a given trial were modeled in the following choice function, where V_{jt} is the derived predicted value associated with trusting a specific player, and ζ and ψ , respectively, are inverse-temperature and bias parameters, fit to each subject.

$$p(\$1.00) = \frac{e^{\zeta * V_{jt}}}{\sum_{i=1}^t e^{\zeta * V_{jt}} + e^{\zeta * \psi}}$$

Supplemental Results

Model Fits and Model Comparison. All model fits were calculated based on Bayes Loglikelihood, which was used to compute AIC (Akaike information criterion) using the following formula.

$$AIC = 2*(\log \text{likelihood}) - 2*(\text{No. model parameters})$$

To determine the best-fitting model, we applied Bayesian model selection over group AICs using the `spm_BMS` function (Stephan, Penny, Daunizeau, Moran, & Friston, 2009). Bayesian model selection evaluates the exceedance probability that a given model is more likely than others, given the full set of AIC values for each model and each subject. Bayesian model selection is particularly robust to outliers than a simple comparison of mean AIC fits across groups (Stephan et al., 2009).

Table S3: Model fit estimates across groups, tasks, and computational models.

Group	Condition	Model	AIC	SE _{AIC}	Bayes Loglikelihood	SE _{BayesLL}
Healthy	Trust Game	DB-RL	-62.08	0.99	25.04	0.49
		B-RL	-64.67	0.984	28.34	0.49
		RL	-74.97	1.27	34.48	0.63
	Slot Machine	DB-RL	-64.61	1.13	26.31	0.56
		B-RL	-67.26	1.117	29.63	0.56
		RL	-83.51	1.357	34.48	0.63
Anxious	Trust Game	DB-RL	-64.76	1.80	26.38	0.90
		B-RL	-66.24	1.84	29.12	0.92
		RL	-85.46	2.46	39.73	1.23
	Slot Machine	DB-RL	-66.26	1.71	27.13	0.86
		B-RL	-67.48	1.76	29.74	0.88
		RL	-90.01	2.37	42.01	1.18

Figure S4

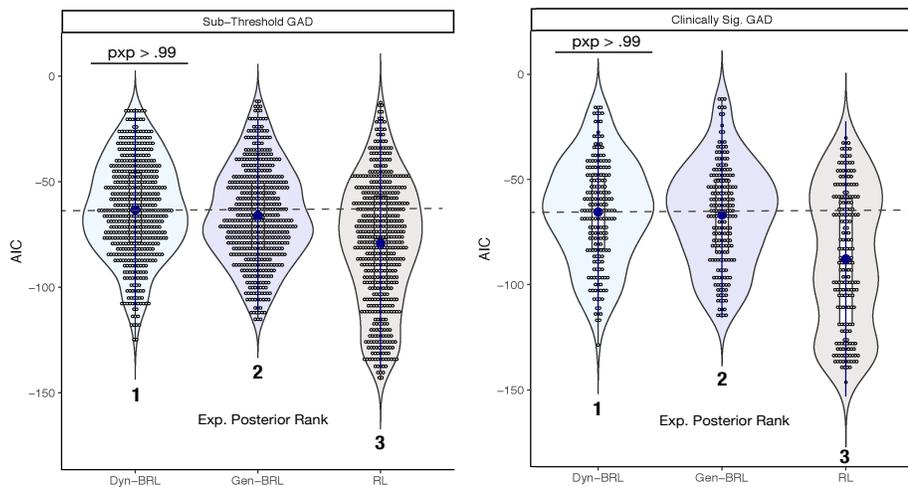


Figure S4. Distribution of model AICs with Bayesian Model Selection rank. Model fits are broken out by healthy and anxious groups.

Dynamic Bayesian-RL Model Parameter Fit

Table S4: Mean and standard error estimates of DB-RL model parameters.

Group	Cond.	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
		γ_{1pos}	γ_{1pos}	γ_{1neg}	γ_{1neg}	γ_{0pos}	γ_{0pos}	γ_{0neg}	γ_{0neg}	ζ	ζ	ψ	ψ
Healthy	TG	-0.87	0.04	-0.95	0.04	0.44	0.02	0.52	0.02	7.31	0.26	0.52	0.01
	SM	-0.88	0.04	-0.97	0.04	0.43	0.02	0.48	0.02	7.26	0.26	0.50	0.01
Anxious	TG	-0.82	0.06	-0.83	0.06	0.54	0.03	0.49	0.02	7.17	0.38	0.52	0.02
	SM	-0.79	0.06	-0.96	0.06	0.53	0.03	0.48	0.02	6.40	0.38	0.47	0.02

Figure S5

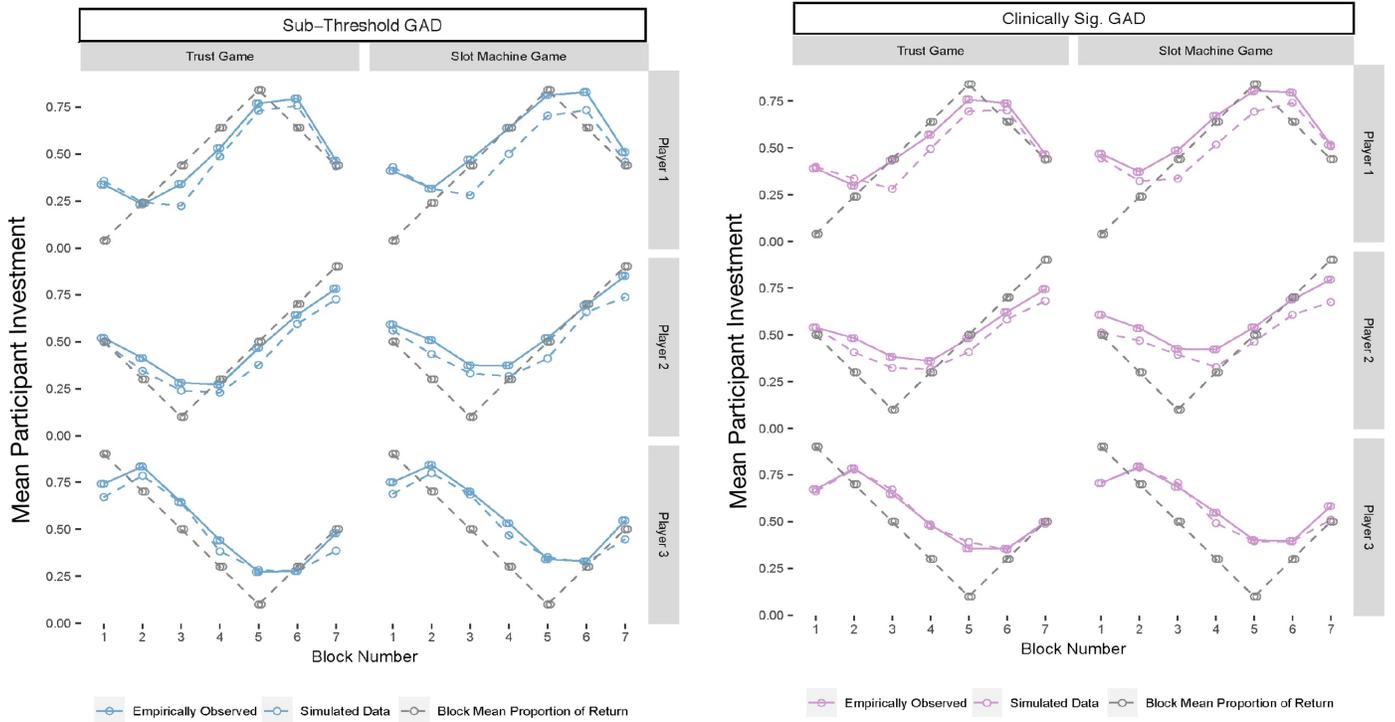
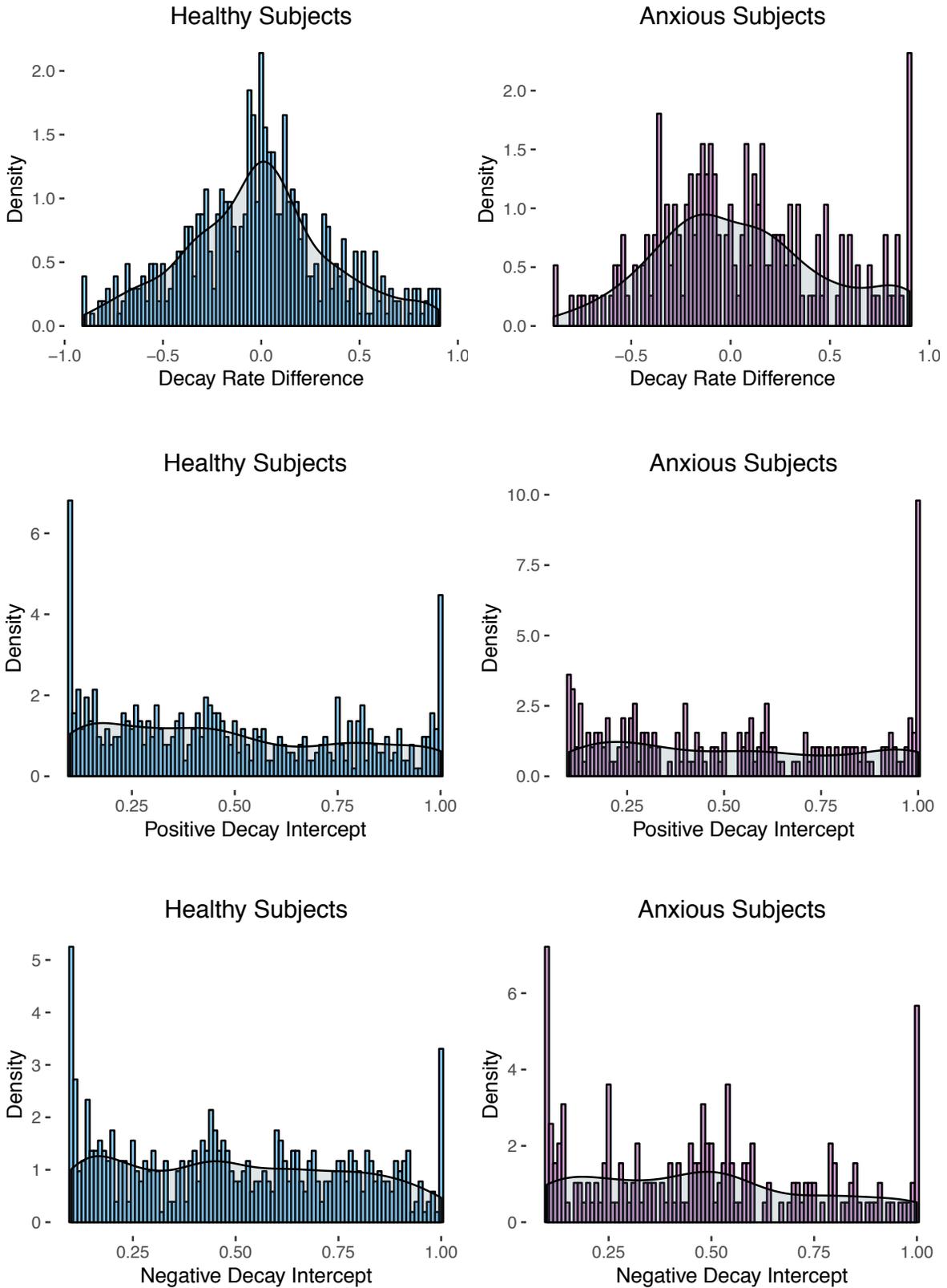
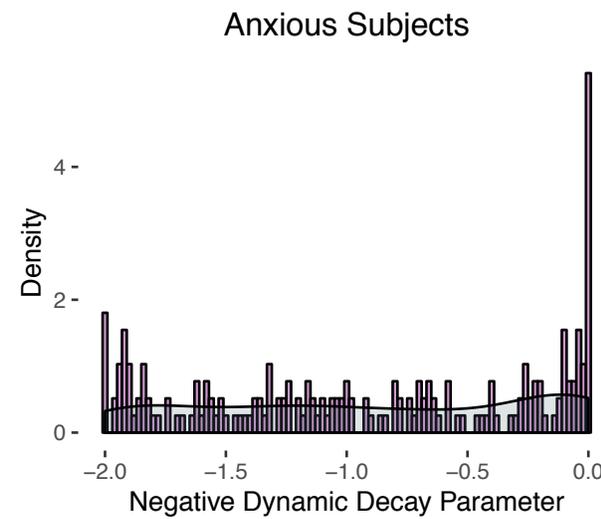
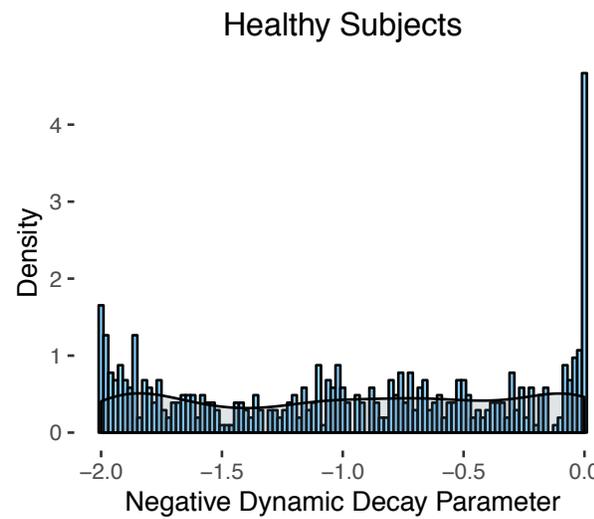
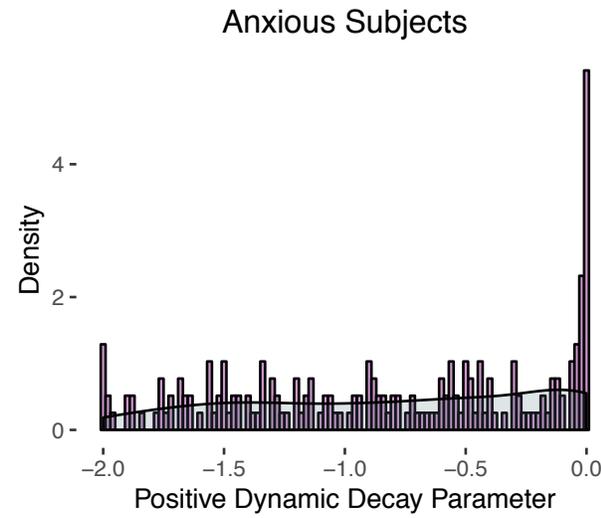
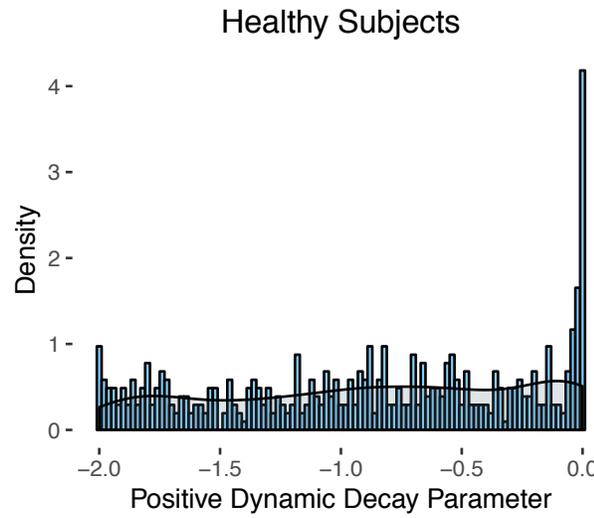
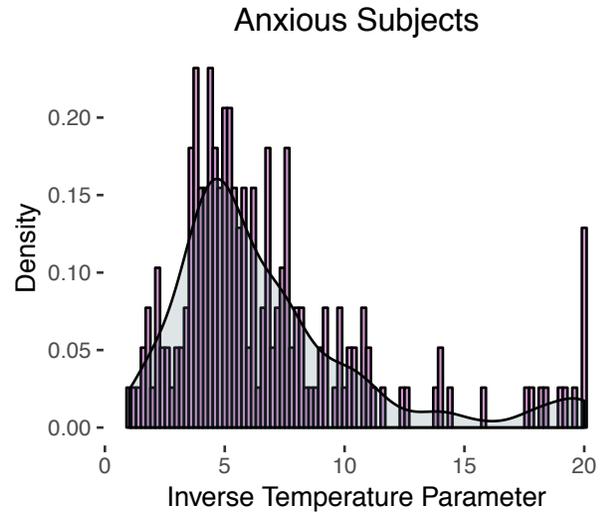
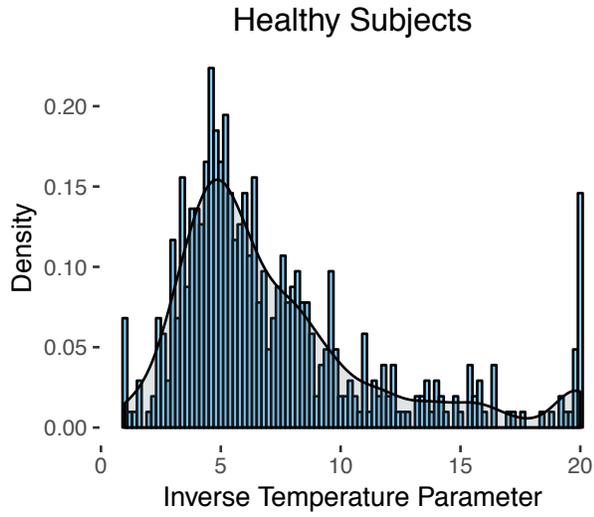


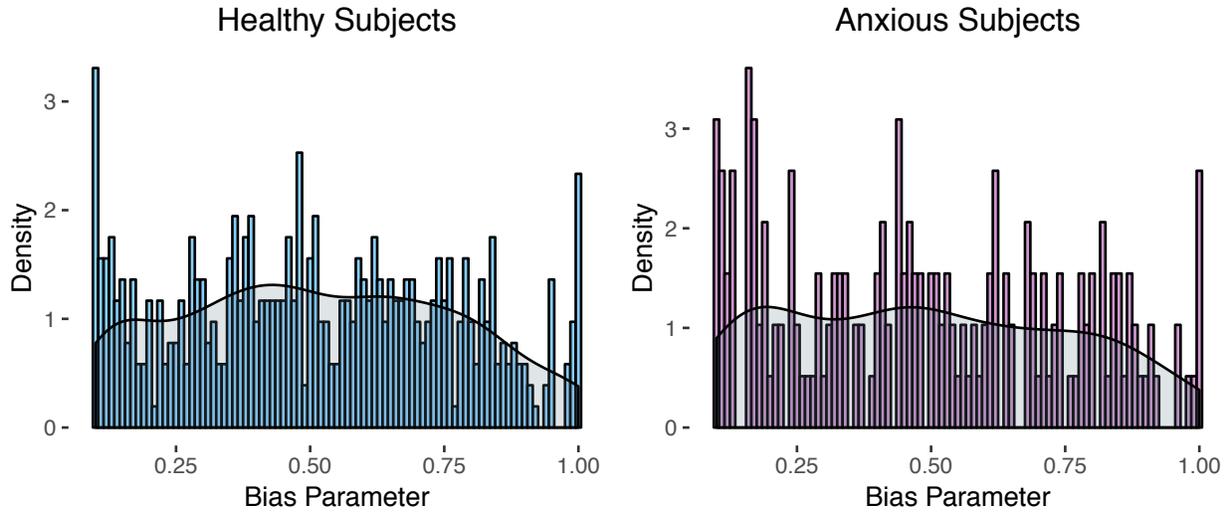
Figure S5. Plot of posterior predictive simulation of DB-RL model. Posterior predictive check was conducted by simulating data from subject-specific parameter fits.

Distribution of Parameter Fits.

Figure S6

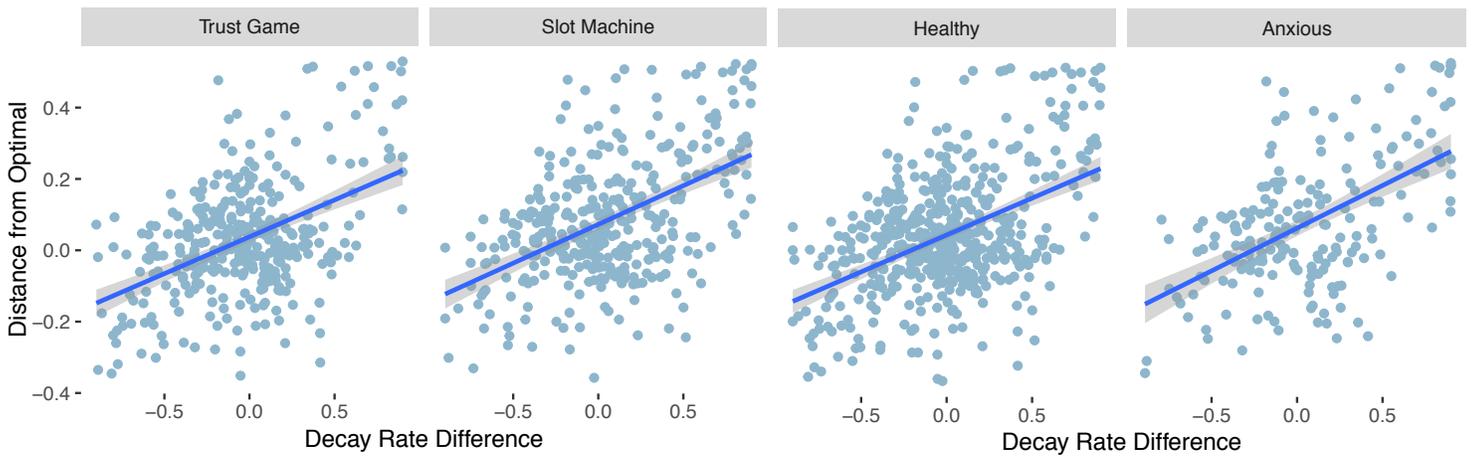






Model Performance Check. In order to ensure that our parameters of interest were adequately capturing task performance, we modeled the decay rate difference as a function of distance from the optimal response on the previous trial which is a rough approximation of learning from feedback (subject investment – optimal investment_{t-1}). As shown below, the decay rate difference was significantly correlated with distance from optimal ($p < 0.001$ for all), validating that our measure of interest was capturing relevant task behavior. As shown below, a higher decay rate difference ($\gamma_{0_{pos}} - \gamma_{0_{neg}}$) was associated with over-investing, whereas a lower negative rate difference ($\gamma_{0_{neg}} > \gamma_{0_{pos}}$) was indicative of under-invested. Conversely, a decay rate difference of 0, ($\gamma_{0_{neg}} = \gamma_{0_{pos}}$) meaning equal decay of rewards and losses, was associated with optimal learning.

Figure S7.



Bayesian-RL Model

Table S5: Mean and standard error estimates of B-RL model parameters.

<i>Group</i>	<i>Cond.</i>	<i>Mean</i>	<i>SE</i>	<i>Mean</i>	<i>SE</i>	<i>Mean</i>	<i>SE</i>	<i>Mean</i>	<i>SE</i>
		γ_{pos}	γ_{pos}	γ_{neg}	γ_{neg}	ζ	ζ	ψ	ψ
<i>Healthy</i>	<i>TG</i>	0.50	0.01	0.55	0.01	8.29	0.30	0.52	0.01
	<i>SM</i>	0.54	0.01	0.50	0.01	7.83	0.30	0.50	0.01
<i>Anxious</i>	<i>TG</i>	0.56	0.03	0.51	0.02	7.61	0.44	0.53	0.02
	<i>SM</i>	0.56	0.03	0.46	0.02	7.86	0.44	0.48	0.02

RL Model

Table S6: Mean and standard error estimates of RL model parameters.

<i>Group</i>	<i>Cond.</i>	<i>Mean</i>	<i>SE</i>	<i>Mean</i>	<i>SE</i>	<i>Mean</i>	<i>SE</i>
		λ	λ	ζ	ζ	ψ	ψ
<i>Healthy</i>	<i>TG</i>	0.44	0.01	3.49	0.15	0.37	0.01
	<i>SM</i>	0.42	0.01	2.89	0.15	0.31	0.01
<i>Anxious</i>	<i>TG</i>	0.37	0.02	2.64	0.21	0.33	0.02
	<i>SM</i>	0.39	0.02	2.62	0.21	0.30	0.02

References

- Daw, N., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci*, *8*(12), 1704-1711. doi:10.1038/nn1560
- Doll, B. B., Jacobs, W. J., Sanfey, A. G., & Frank, M. J. (2009). Instructional control of reinforcement learning: a behavioral and neurocomputational investigation. *Brain Res*, *1299*, 74-94. doi:10.1016/j.brainres.2009.07.007
- Frank, M. J., Doll, B. B., Oas-Terpstra, J., & Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat Neurosci*, *12*(8), 1062-1068. doi:10.1038/nn.2342
- Franklin, N. T., & Frank, M. J. (2015). A cholinergic feedback circuit to regulate striatal population uncertainty and optimize reinforcement learning. *Elife*, *4*. doi:10.7554/eLife.12029
- Löwe, B., Decker, O., Muller, S., Brahler, E., Schellberg, D., Herzog, W., & Herzberg, P. Y. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Med Care*, *46*(3), 266-274. doi:10.1097/MLR.0b013e318160d093
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *JAMA Internal Medicine*, *166*(10), 1092-1097. doi:10.1001/archinte.166.10.1092
- Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *Neuroimage*, *46*(4), 1004-1017. doi:10.1016/j.neuroimage.2009.03.025