Improving the reliability of the Pavlovian go/no-go task for computational psychiatry research

Samuel Zorowitz[1,*], Gili Karni[1], Natalie Paredes[2], Nathaniel Daw[1,3], Yael Niv[1,3]

[1]Princeton Neuroscience Institute, Princeton University, USA
[2]Department of Psychology, University of California, San Diego, USA
[3]Department of Psychology, Princeton University, USA
[*]Corresponding author (szorowi1@gmail.com)

Abstract

Background: The Pavlovian go/no-go task is commonly used to measure individual differences in Pavlovian biases and their interaction with instrumental learning. The task has also been widely used in computational psychiatry research, to correlate Pavlovian biases with mental health symptoms. However, prior research has reported unacceptable reliability for computational model-based performance measures for this task, limiting its usefulness in individual-differences research. Here, we apply several strategies previously shown to enhance task-measure reliability (e.g., task gamification, hierarchical Bayesian modeling for model estimation) to the Pavlovian go/no-go task, to improve the reliability of the task as a tool for future research.

Methods: In two experiments, two independent samples of adult participants (N=103, N=110) completed a novel, gamified version of the Pavlovian go/no-go task multiple times over several weeks. We used hierarchical Bayesian modeling to derive reinforcement learning model-based indices of participants' task performance, and to estimate the reliability of these measures.

Results: In Experiment 1, we observed considerable practice effects, with most participants reaching near-ceiling levels of performance with repeat testing. Consequently, the test-retest reliability of some model parameters was unacceptable (as low as 0.379). In Experiment 2, participants completed a modified version of the task designed to lessen these practice effects. We observed greatly reduced practice effects and improved estimates of the test-retest reliability (range: 0.696–0.989).

Conclusion: The results demonstrate that model-based measures of performance on our modified Pavlovian go/no-go task can reach levels of reliability sufficient for use in individual-differences research. We therefore provide the task code for use by the computational psychiatry community (as well as other researchers). Additional investigation is necessary to validate the modified version of the task in other populations and settings.

Introduction

Humans (and other animals) have an innate tendency to approach rewarding stimuli and shrink from punishing stimuli [1]. Depending on the context, these hardwired Pavlovian biases can either benefit or interfere with instrumental (i.e., action-outcome) learning. This is epitomized in the Pavlovian go/no-go task in which the required action (Go, No-Go) and outcome valence (reward, punishment) are orthogonalized [2, 3]. In the task, participants are typically faster to learn actions that are congruent with Pavlovian response biases (i.e., a "Go" response to receive reward and a "No-Go" response to avoid punishment) as compared to Pavlovian-instrumental incongruent responses (i.e., inhibit action to receive reward, initiate action to avoid punishment).

The Pavlovian go/no-go task has been used in a large number of studies to probe individual differences in reward and punishment learning, of which many have reported changes in Pavlovian biases as a function of psychiatric conditions. For example, an increased tendency towards passive avoidance has been observed in individuals with generalized and social anxiety [4, 5], whereas active avoidance is amplified in individuals with a history of suicidal thoughts or behaviors [6]. Pavlovian biases are larger in individuals with trauma exposure [7] and first-episode psychosis [8], but attenuated in individuals with depression [9] and schizophrenia [10]. Pavlovian biases have also been associated with individual differences in personality (e.g., impulsivity; [11]) and genetics [12, 13]. In developmental and lifespan research, Pavlovian biases have been shown to exhibit a U-shape, decreasing from childhood to young adulthood and increasing again in older age [14, 15]. At a finer temporal scale, Pavlovian biases are also reportedly modulated by state effects including mood [16], anger [17], stress [18], and fear [19].

Despite its extensive use, the Pavlovian go/no-go task has received comparably less psychometric investigation. What studies do exist suggest the task, in its canonical form, is poorly suited for use in individual-differences correlational research. Specifically, three independent studies found that descriptive and computational model-based measures of performance on the Pavlovian go/no-go task exhibited low test-retest reliability over short (two-week) and long (6-, 18-month) retest intervals [20–22]. This is important because the reliability of a measure places an upper bound on the maximum observable correlation between itself and a second measure (e.g., psychiatric symptom scores; [23]). Therefore, as reliability decreases, so too does statistical power; in turn, this increases the possibility of false negatives [24] and bias in correlations that do reach statistical significance [25]. Other studies also reported evidence of practice effects (i.e., improved performance and diminished Pavlovian biases with repeated testing; [20, 22, 26]). Practice effects are of additional concern for individual-differences research because they can imperil reliability (e.g., by minimizing between-participants variability) and obscure effects of interest (e.g., changes in task performance following changes in mood).

There are multiple strategies for improving the reliability of cognitive task measures [27]. For example, prior research has found that gamification, or the incorporation of (video) game design elements into cognitive tasks, can promote participant engagement [28] and

improve the reliability of task measures [29, 30]. Moreover, hierarchical Bayesian models – which exert a pooling effect on person-level variables, in effect correcting them for measurement error [31, 32] – have been frequently shown to improve the reliability of task measures [33–35]. Finally, practice effects can be lessened by designing tasks in such a way that prevents participants from discovering and using task-specific knowledge to enhance their performance on subsequent attempts [36].

Here we investigate the reliability and repeatability of a novel version of the Pavlovian go/no-go task, with the aim of designing a variant of the task that is optimized for use in computational psychiatry and other individual differences research. We conducted two experiments involving two independent samples of adult participants who completed a gamified version of the task multiple times over several weeks. We used hierarchical Bayesian models to derive reinforcement-learning model-based indices of their task performance, and additionally to estimate the reliability of these measures. In Experiment 1, using a gamified version of the classic task, participants exhibited large practice effects, which negatively impacted the test-retest reliability of the performance measures. To address this issue, in Experiment 2, participants completed a modified version of the task that reduced practice effects, and led to significant improvements in the test-retest reliability of the reinforcement learning model parameters.

Experiment 1

Methods

Participants

A total of N=148 participants were recruited in May, 2020, from Amazon Mechanical Turk via CloudResearch [37]. Participants were eligible to participate if they were at least 18 years old and resided in the United States. Following best practice recommendations [38], no other inclusion criteria were applied. The study was approved by the Institutional Review Board of Princeton University and all participants provided informed consent. Total study duration was 15–20 minutes. Participants received monetary compensation for their time (rate: USD \$12/hr), plus an incentive-compatible bonus up to \$1.50 based on task performance.

Data from N=45 participants who completed the first session were excluded prior to analysis (see "Exclusion criteria" below), leaving a final sample of N=103 participants. These participants were re-invited to complete follow-up experiments 3, 14, and 28 days later. Once invited, participants were permitted 48 hours to complete each follow-up experiment. Retention was high for each follow-up session (Day 3: N=94 [91.3%]; Day 14: N=92 [89.3%]; Day 28: N=89 [86.4%]). In addition to the performance bonus, participants received a retention bonus of \$1.00 for each completed follow-up session. Detailed demographic information is presented in Table S1. The majority of participants

identified as men (55 men; 47 women; 1 non-binary) and participants were 35.5 years old on average (range: 20–69 years).
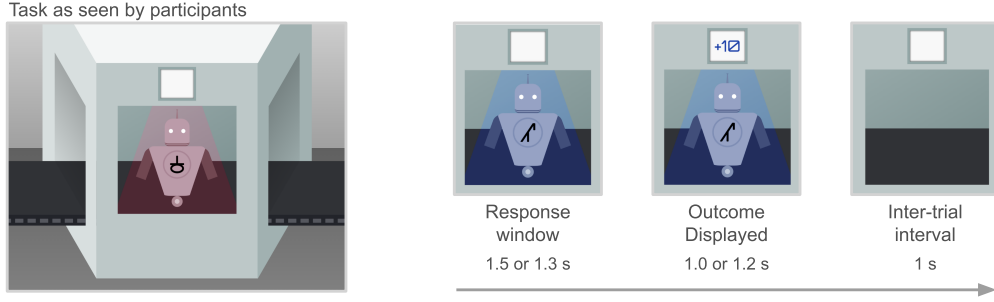
Experimental protocol

In each session, after providing consent, participants started by completing some or all of the following self-report questionnaires: the 7-item generalized anxiety disorder scale (GAD-7; [39]); the 14-item manic and depressive tendencies scale (7-up/7-down; [40]); and the abbreviated 12-item behavioral activation/inhibition scale (BIS/BAS; [41]). Participants also indicated their current mood using an affective slider [42]. Note that participants completed the GAD-7 and mood slider on each session, but the 7-up/7-down and BIS/BAS scales only twice (on Days 0 and 28). These measures were included for exploratory analyses not reported here.

Next, participants completed a gamified version of the Pavlovian go/no-go task. In the task, participants observed different 'robot' stimuli (Figure 1A). On every trial, a robot was shown traveling down a conveyor belt into a 'scanner'. Once inside, participants had 1.5 seconds to decide to either 'repair' the robot by pressing the space bar ("Go" response) or press nothing ("No-Go" response). Participants were told that they would see different types of robots (indicated by a symbol on the robots' chestplates), and that their goal was to learn which types of robots needed repairing based on feedback (points won/lost) following their actions.

The task involved four trial types that differed by their correct action (Go, No-Go) and outcome domain (reward, punishment; Figure 1B). Specifically, the four trial types were: go to win points (GW); no-go to win points (NGW); go to avoid losing points (GAL); and no-go to avoid losing points (NGAL). Note that GW and NGAL trials are Pavlovian-instrumental 'congruent' because there is a match between the correct response and the expected approach/avoidance bias due to winning or losing points for each. In contrast, NGW and GAL trials are Pavlovian-instrumental 'incongruent'. In rewarding trials (GW, NGW), the possible outcomes were +10 or +1 points where a correct action was rewarded with +10 on 80% of the trials and +1 otherwise; in turn, an incorrect action was rewarded with +1 on 80% of the trials and +10 otherwise. In punishing trials (GAL, NGAL), outcomes were -1 or -10 points, where the correct action led to -1 on 80% of trials and the incorrect action led to -10 on 80% of trials (Figure 1C). The outcome domain of each robot was explicitly signaled to participants by a blue or orange 'scanner light' (one color signaling reward domain and the other punishment domain, randomized within participants across sessions).

Participants saw eight unique robots in each session of the task. Each individual robot was presented for 30 trials (240 trials total; Figure 1D). Trials were divided into two blocks with four robots per block (one of each trial type). Prior to task start, participants were required to review instructions, correctly answer five comprehension questions that touched on all essential parts of the instructions, and complete several practice trials. Failing to correctly answer all comprehension questions forced the participant to

A. Task schematic

Task as seen by participants

Response window
1.5 or 1.3 s

Outcome Displayed
1.0 or 1.2 s

Inter-trial interval
1 s

B. Trial types

		Outcome domain	
		Reward	Punishment
Correct response	Go	**Go to Win (GW)**	**Go to Avoid Losing (GAL)**
	No-go	**No-go to Win (NGW)**	**No-go to Avoid Losing (NGAL)**

C. Outcome probabilities

		Outcome domain		
		Reward		Punishment
Correct Response		**+10** 80%	**-1** 80%	
		+1 20%	**-10** 20%	
Incorrect Response		**+1** 80%	**-10** 80%	
		+10 20%	**-1** 20%	

D. Trial composition

Experiment 1	Experiment 2
2 x **GW** (30 trials)	6 x **GW** (8-12 trials)
2 x **NGW** (30 trials)	6 x **NGW** (8-12 trials)
2 x **GAL** (30 trials)	6 x **GAL** (8-12 trials)
2 x **NGAL** (30 trials)	6 x **NGAL** (8-12 trials)
Total: 240 trials	**Total:** 240 trials

Figure 1: **(A) Schematic of the Pavlovian go/no-go task.** On each trial, a robot entered the 'scanner' from the left of screen, prompting a response (go or no-go) from the participant during a response window (Experiment 1: 1.5 seconds; Experiment 2: 1.3 seconds). The outcome (number of points won or lost) was subsequently presented on the scanner display (Experiment 1: 1.0 seconds; Experiment 2: 1.2 seconds), followed by an inter-trial interval animation (1 second) in which the conveyor belt carried the old robot out of view and a new robot into the scanner. The color of the scanner light denoted outcome domain (e.g., blue denoting reward and red denoting punishment). **(B) The four trial types**, produced by a factorial combination of outcome domain (rewarding, punishing) and correct action (go, no-go). **C) Outcome probabilities** for each outcome domain following a correct or incorrect response. Correct responses yielded the better of the two possible outcomes with 80% chance. **(D) Trial composition.** In Experiment 1, participants saw 8 total robots (two of each trial type), each presented for 30 trials (240 total trials). In Experiment 2, participants saw 24 total robots (6 of each trial type), each for 8, 10, or 12 trials (240 total trials).

reread sections of the instructions. Participants were required to complete the instructions and comprehension questions in each session. Participants were provided a break between blocks. After completing the task, participants appraised the task along three dimensions: difficulty, fun, and clarity of instructions (see Table S2). The task was programmed in jsPsych [43] and distributed using custom web-application software (see Code Availability).

Exclusion criteria

To ensure data quality, data from multiple participants from the initial session were excluded prior to analysis for one or both of the following reasons: failing more than one attention check embedded in the self-report measures (N=13; [44]) or exhibiting chance-

level performance (<55% correct responses) on go-to-win trials (N=43). In total, data from N=45 participants who completed the first session were excluded based on these criteria, leaving a final sample of N=103 participants. No exclusions were applied to subsequent session data.

Descriptive analyses

We first evaluated participants' choice behavior using five performance measures: overall percent correct responses; go bias, calculated as the difference in correct responses between Go and No-Go trials; valence bias, calculated as the difference in correct responses between rewarding and punishing trials; Pavlovian bias, which was the difference in correct responses between Pavlovian-instrumental congruent and incongruent trials; and feedback sensitivity, calculated as the difference in correct responses between trials following veridical or sham feedback (that is, following 80% of the trials where feedback aligned with the correctness of the response, and the 20% of trials with feedback matching the alternative response, respectively). Consistent with previous research [2, 22], only small or nonsignificant valence biases were observed. As such, these statistics are reported only in the Supplementary Materials (Table S4).

For each session and measure, we tested if the median value across participants was significantly different than zero (or 50% for overall percent correct responses). We used the median due to skew in the performance measures. We also tested if the median value of each measure was significantly different between each pair of sessions. P-values were derived via permutation testing, where a null distribution of values was obtained by permuting the condition labels (for within-session tests) or session labels (for between-session tests) 5,000 times. Within-session tests were not corrected for multiple comparisons as each test constituted an individual hypothesis test; however, between-session tests were corrected using the family-wise error rate correction [45] because they constituted a disjunctive test [46].

Reinforcement learning models

To more precisely characterize participants' performance on the Pavlovian go/no-go task, we fit a nested set of reinforcement learning models to the choice data. All models were variants of the Rescorla-Wagner model and have previously been used to predict choice behavior on this task [2, 4, 20, 47]. Under the most complex model (M7), the probability that a participant makes a go response following stimulus k was defined as:

$$p(y = \text{go}) = (1 - \xi) \cdot \text{logit}^{-1} \left(\beta_{v_k} \cdot [Q_k(\text{Go}) - Q_k(\text{NoGo})] + \tau_{v_k} \right) + \frac{\xi}{2} \tag{1}$$

where β_{v_k} was the reward sensitivity (if the valence v of stimulus k was rewarding) or the punishment sensitiivty (if stimulus k was punishing), $Q_k(\text{go})$ and $Q_k(\text{no-go})$ were learned stimulus-action values for the go and no-go responses for stimulus k, respectively, τ_{v_k} was an approach bias (if stimulus k was rewarding) or avoidance bias parameter (if stimulus

5

k was punishing), and ξ was the lapse rate (i.e., the rate of choosing actions randomly due to lapse of attention). The Q values were learned through feedback according to a learning rule:

$$Q_k(\text{action}) \leftarrow \eta_{v_k} \cdot [r - Q_k(\text{action})] \qquad (2)$$

where r was the observed outcome on this trial and η_{v_k} was the learning rate or step-size parameter (η_+ if stimulus k was rewarding, η_- if it was punishing). In all models, we encoded rewards as follows: the better of the two possible outcomes in each domain (i.e., 10 and -1) was encoded as $r = 1$ whereas the worse of the two possible outcomes (i.e., 1 and -10) was encoded as $r = 0$. This was done for convenience only. As reward magnitudes were instructed, the reward/punishment domain signaled, and appetitive and aversive rewards were never intermixed, only the relative reward within condition was germane. Q-values were accordingly initialized to $Q = 0.5$.

Simplifications of this model involved either fixing parameters to be equal to zero (e.g., no lapse rate) or fixing parameters to be equal for reward and punishment domains. Specifically, the base model (M1) had only two free parameters: a single outcome sensitivity parameter and a single learning rate, both shared across outcome domains (i.e., $\beta_+ = \beta_-$; $\eta_+ = \eta_-$; $\tau_+ = \tau_- = 0$, $\xi = 0$). Model 2 added a static action bias parameter that was shared across outcome domains (i.e., $\tau_+ = \tau_-$). Model 3 added to M2 independent approach (τ_+) and avoidance (τ_-) parameters. Models 4 and 5 respectively added to M3 independent outcome sensitivity (β_+, β_-; M4) or learning rate (η_+, η_-; M5) parameters by outcome domain. Model 6 included both independent outcome sensitivity and learning rate parameters. Finally, Model 7, the most complex model, added to M6 a potentially non-zero lapse rate (ξ).

All models were estimated within a hierarchical Bayesian modeling framework using Hamiltonian Monte Carlo sampling as implemented in Stan (v2.30; [48]). For each model, four separate chains with randomized start values each took 7,500 samples from the posterior. The first 5,000 samples, and every other subsequent sample from each chain were discarded, retaining a total of 5,000 post-warmup samples from the joint posteriors of all four chains. The $\hat{R}$ values for all parameters were ≤ 1.01, indicating acceptable convergence between chains, and there were no divergent transitions in any chain. For all models, we specified diffuse, uninformative priors in order to avoid biasing parameter estimation (Table S3).

Fits of the models to behavioral data were assessed using posterior predictive checks. Specifically, we inspected each model's ability to reproduce both group-averaged learning curves by trial type and each participant's proportion of go responses by trial type. Model fits were compared using approximate leave-one-trial-out cross-validation via Pareto smoothed importance sampling (PSIS-LOO; [49]). (Note this may, in principle, differ from cross-validation at the participant level, which has been argued to be a relevant unit of exchangeability at which to compare models [50].) We considered a difference in PSIS-LOO values that is four times larger than the mean PSIS-LOO standard error as a significant improvement in model fit due to additional parameters [51].

We also investigated the reliability of the model parameters for the best-fitting model

using a Bayesian hierarchical modeling framework, in which data were pooled within and across participants [32]. Briefly, the best-fitting reinforcement learning model chosen using the above procedure was then re-fit separately per participant and session (in the case of test-retest reliability) or task block (in the case of split-half reliability). Specifically, each parameter $\theta \in \{\beta_+, \beta_-, \eta_+, \tau_+, \tau_-, \xi\}$ was estimated as follows:

$$\theta_{i1} = \mu_1 + \theta_{ic} - \theta_{id}$$
$$\theta_{i2} = \mu_2 + \theta_{ic} + \theta_{id} \tag{3}$$

where θ_{i1} and θ_{i2} are a given parameter (e.g., reward sensitivity, β_+) for participant i in sessions or blocks 1 and 2, respectively; μ_1 and μ_2 are the group-averaged parameters for sessions or blocks 1 and 2; θ_{ic} is the common effect for participant i (i.e., the component of a participant-level parameter that is different from the group mean and stable across sessions or blocks); and θ_{id} is the difference effect for participant i (i.e., the parameter component that is variable across sessions or blocks). The collection of θ_{ic} parameters constituted between-participants variability, whereas the collection of θ_{id} parameters constituted within-participants variability. Both θ_{ic} and θ_{id} were assumed to be normally distributed with zero means and independent estimated variances. Split-half and test-retest reliability estimates were calculated by taking the Pearson correlation of θ_{i1} and θ_{i2} across task blocks and sessions, respectively [21, 34]. We used the Pearson correlation because we were primarily interested in the consistency of rank ordering of participants' parameter estimates over time. Although arbitrary, we followed convention and defined $\rho \geq 0.7$ as the threshold for "acceptable" reliability [52].

Results

Descriptive analyses

Trial-by-trial choice behavior for each session is presented in Figure 2A. Performance in the first session qualitatively conformed to the expected pattern of results (i.e., worse performance on Pavlovian-instrumental incongruent trials [GAL, NGW]). However, this effect seemed diminished in all follow-up sessions. Indeed, group-averaged performance measures by session (Figure 2B; complete descriptive statistics are reported in Table S4) showed that participants made the correct response on 85.0% of trials on the first session (Day 0), which increased to near-ceiling levels in all subsequent sessions. Pairwise comparisons confirmed that performance was indeed worse on Day 0 compared to each follow-up session (all $p < 0.001$); no other comparisons were significant. Participants' self-reported mood and anxiety were largely stable over the same period (Figure S1), indicating this shift in performance more likely reflects practice effects rather than changes in participants' state.

Across sessions, participants made more correct responses on Go trials than on No-Go trials. However, this "Go bias" was significantly reduced in all follow-up sessions compared to Day 0 (all $p < 0.001$); so too was it on Day 28 compared to Day 3 ($p < 0.001$).
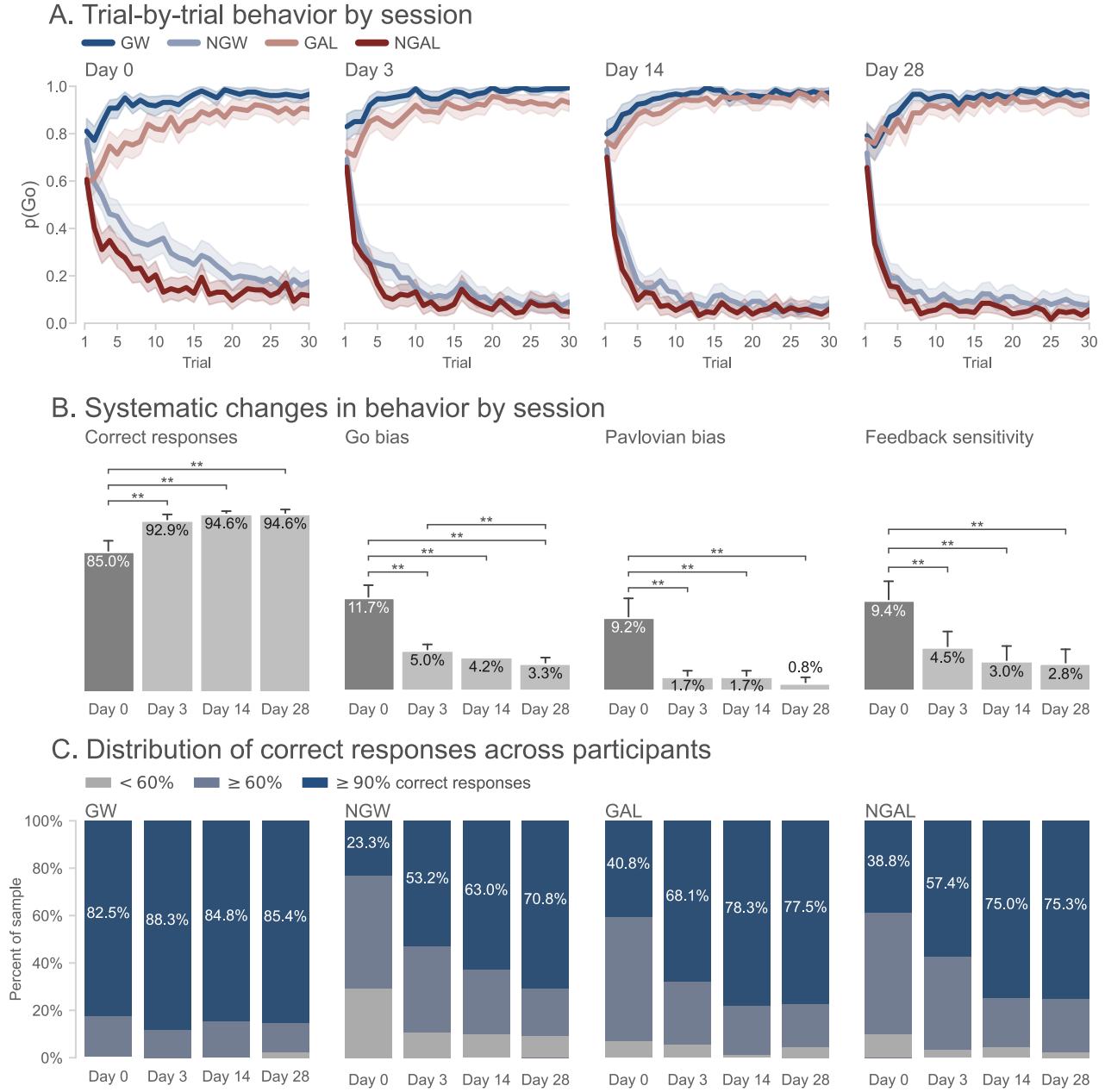
Figure 2: **Large practice effects on the standard Pavlovian go/no-go task in Experiment 1. (A) Group-averaged learning curves** for each trial type and session. Shaded regions indicate 95% bootstrapped confidence intervals. **(B) Group-averaged performance for each session.** Performance measures from left-to-right: Correct responses, or overall accuracy; Go bias, or difference in accuracy between Go and No-Go trials; Congruence effect, or difference in accuracy between congruent (GW, NGAL) and incongruent (NGW, GAL) trials; and Feedback sensitivity, or the difference in accuracy on trials following veridical and sham feedback. Behavior on the first session was significantly different from all other sessions on all measures. ** Denotes significant pairwise difference ($p < 0.05$, corrected for multiple comparisons). **(C) Distribution of correct responses across sessions by trial type.** Percentage of participants, for each session and trial type, exhibiting at- or below-chance performance ($< 60\%$ response accuracy; grey), intermediate performance ($\geq 60\%$ response accuracy; light blue), or near-perfect performance ($\geq 90\%$ response accuracy; dark blue). Across sessions, performance improved on all trial types that were not already close to ceiling on the first session.

Similarly, participants made more correct responses on congruent than incongruent trials. As with the Go bias, this "Pavlovian bias" was significantly reduced in all follow-up sessions compared to Day 0 (all $p < 0.001$; no other comparisons were significant).

Feedback sensitivity also diminished from the first to later sessions. Across sessions, participants made more correct responses following veridical compared to sham feedback (all $p < 0.001$). However, feedback sensitivity was significantly reduced in all follow-up sessions compared to Day 0 (all $p < 0.001$; no other comparisons were significant) suggesting that feedback had less of an effect on choice in later sessions. This is consistent with participants' learning curves which show, in all days except Day 0, that participants quickly learned the correct action for each stimulus and maintained this policy despite the 20% sham feedback (Figure 2A).

These results summarize group-averaged performance. To gain insight into individual differences, Figure 2C shows the proportion of participants who exhibited chance-level ($<60\%$ correct responses), intermediate ($\geq 60\%$ and $<90\%$), or near-ceiling performance ($\geq 90\%$) by session and trial type. Excepting GW trials, where performance of over 80% of participants was close to ceiling already in the first session, the percentage of participants nearing ceiling-level performance increases from a minority on Day 0 to the majority of participants in all follow-up sessions. Two-way chi-squared tests confirmed this trend (GW: $\chi^2(6) = 8.149$, $p = 0.227$; NGW: $\chi^2(6) = 55.458$, $p < 0.001$; GAL: $\chi^2(6) = 42.191$, $p < 0.001$; NGAL: $\chi^2(6) = 39.287$, $p < 0.001$). In sum, the improvements in task performance (and accompanying reductions in choice biases) with repeat testing observed at the group-level extended to the majority of participants.

Model comparison

The results of the model comparison are summarized in Table 1. Collapsing across sessions, the best-fitting model was the most complex one (i.e., the model including independent reward sensitivity, learning rate and approach/avoidance bias parameters per outcome domain, plus a lapse rate; M7). Importantly, this was also the best-fitting model within each session (Table S5). Posterior predictive checks indicated that this model provided excellent fits to the choice data from each session (Figure S3).

Model parameters

Figure 3A shows the estimated group-level parameters from the best-fitting model. Consistent with the descriptive analyses above, large shifts in parameter values were observed following Day 0. The reward and punishment sensitivity parameters (β_+, β_-) exhibited an almost threefold increase between Days 0 and 3, and stabilized thereafter. The inverse pattern was observed for the positive learning rate (η_+). Crucially, the approach/avoidance bias parameters followed a similar pattern. The approach bias (τ_+) decreased significantly between Days 0 and 3, and qualitatively declined thereafter. In turn, the avoidance bias (τ_-) increased significantly between Days 0 and 3, and stabi-

lized thereafter. That is, Pavlovian biases diminished in absolute and relative terms (i.e., compared to the outcome sensitivity parameters) with repeat testing.

Model	Parameters	Accuracy	PSIS-LOO	ΔPSIS-LOO (se)
M1	β, η	87.5%	-151457.9	-5602.6 (68.3)
M2	β, τ, η	89.0%	-154011.9	-3048.6 (51.2)
M3	$\beta, \tau_+, \tau_-, \eta$	89.8%	-155817.8	-1242.7 (31.3)
M4	$\beta_+, \beta_-, \tau_+, \tau_-, \eta$	89.8%	-156261.6	-798.8 (22.6)
M5	$\beta, \tau_+, \tau_-, \eta_+, \eta_-$	89.9%	-156265.9	-794.6 (20.7)
M6	$\beta_+, \beta_- \tau_+, \tau_-, \eta_+, \eta_-$	89.9%	-156401.8	-658.6 (18.8)
M7	$\beta_+, \beta_- \tau_+, \tau_-, \eta_+, \eta_-, \xi$	90.1%	-157060.5	-

Table 1: **Model comparison collapsing across sessions.** Accuracy = trial-level choice prediction accuracy between observed and model-predicted Go responses. PSIS-LOO = approximate leave-one-out cross-validation scores presented in deviance scale (smaller numbers indicate better fit). ΔPSIS-LOO = difference in PSIS-LOO values between each model and the best-fitting model (M7).
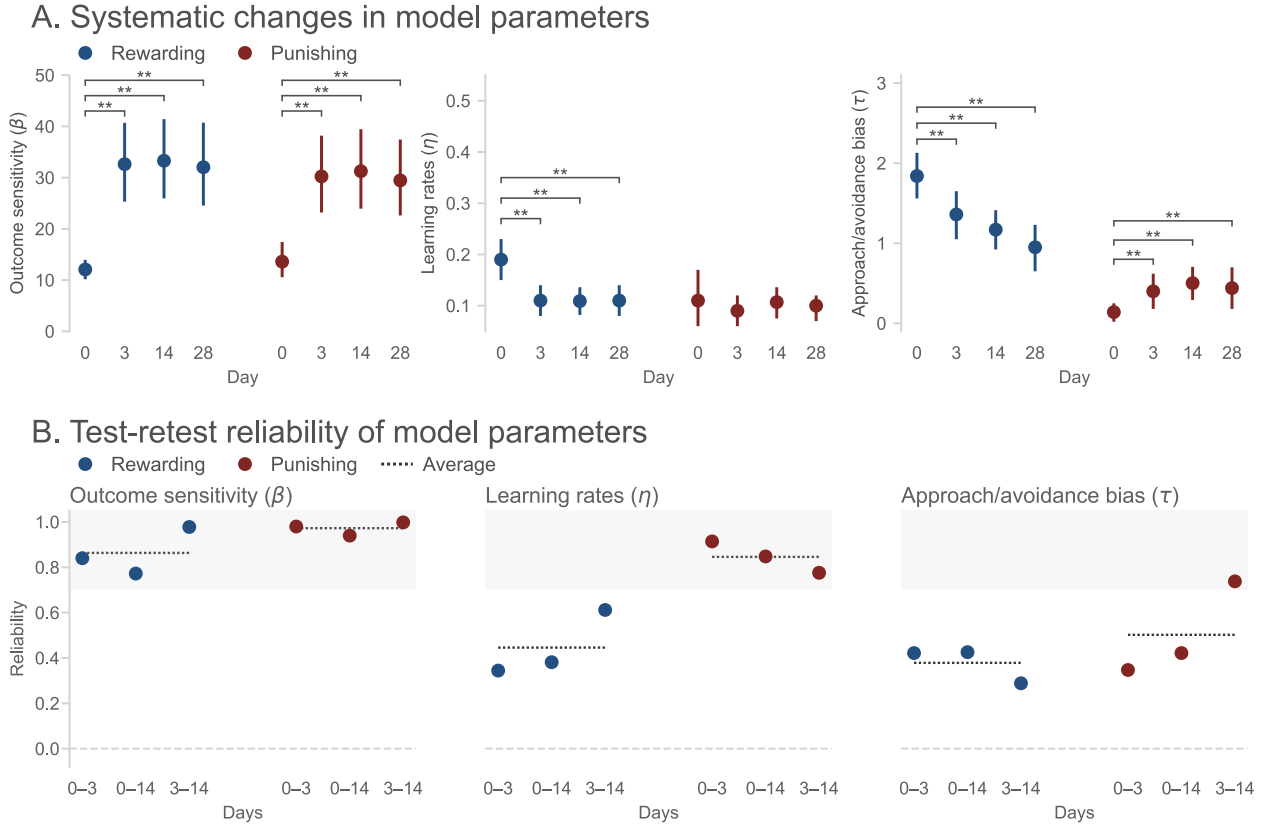

Figure 3: **Reinforcement learning model parameters in Experiment 1 show evidence of practice effects and low reliability.** (A) Group-level model parameters for each session. Error bars indicate 95% Bayesian confidence intervals (CIs). ** Denotes pairwise comparison where 95% CI of the difference excludes zero. (B) Test-retest reliability estimates for each model parameter. Dotted lines indicate average across pairs of sessions. Shaded region indicates conventional range of acceptable reliability ($\rho \geq 0.7$).

The test-retest reliability estimates for each model parameter is presented in Figure 3B. The results are mixed. Averaging across session-pairs, acceptable test-retest reliability was observed for the reward and punishment sensitivity parameters β (average $\rho = 0.863$, 95% CI = [0.837, 0.889] and $\rho = 0.973$, 95% CI = [0.963, 0.981], respectively), and the negative learning rate parameter η_- (average $\rho = 0.846$, 95% CI = [0.764, 0.886]). Conversely, test-retest reliability was unacceptable for the approach and avoidance bias parameters τ (average $\rho = 0.379$, 95% CI = [0.270, 0.486] and $\rho = 0.502$, 95% CI = [0.412, 0.583], respectively), and positive learning rate η_+ (average $\rho = 0.446$, 95% CI = [0.327, 0.555]). A similarly mixed pattern was observed for the split-half reliability estimates (Figure S5A).

Discussion

Our goal was to evaluate the stability and reliability of individual differences in performance on a gamified version of the popular Pavlovian go/no-go task. At both the group and participant levels, we observed significant practice effects following the first session. An increasing majority of participants exhibited near-ceiling performance, across trial types, with each additional task administration. Consequently, the magnitude of group-averaged behavioral effects including the go bias, Pavlovian bias, and feedback sensitivity were diminished by half or more after the first session. This was reflected in the group-level parameters of a reinforcement learning model fit to participants' choice data, which indicated that Pavlovian biases were significantly attenuated in follow-up sessions. Consequently, we found that the Pavlovian bias parameters exhibited poor-to-moderate test-retest reliability. This last result is perhaps unsurprising insofar that low between-participants variability diminishes reliability [27].

The results of Experiment 1 raise two questions: what underlies these practice effects and what can be done to mitigate or prevent them? With respect to the first question, one possibility is that, after the initial session, participants rely on the already learned structure of the task to solve it more effectively. Specifically, in the canonical Pavlovian go/no-go task, for every Go stimulus (e.g., GW) there is a corresponding No-Go stimulus (e.g., NGW). As such, learning the correct action for one stimulus provides information about the correct action for its complement. Recognizing this, savvy participants may forego reinforcement learning in favor of a process-of-elimination strategy to deduce which is the Go and which is the No-Go stimulus in each pair. Indeed, feedback from several participants in this study suggested that they may have utilized this form of top-down strategy.

This suggests that a version of the task with a less predictable trial structure might reduce practice effects. By eliminating the dependence between stimuli, motivated participants aiming to maximize their performance should have no strategy better than learning from the feedback for each of their actions. By minimizing practice effects and increasing between-participants variability, it is plausible that parameter reliability would also improve. In the next experiment, we investigated precisely this.

Experiment 2

Methods

Participants

A total of N=156 participants were recruited in December, 2020, from Amazon Mechanical Turk via CloudResearch [37]. Inclusion criteria were the same as in Experiment 1. The study was approved by the Institutional Review Board of Princeton University, and all participants provided informed consent. Total study duration was again 15-20 minutes. Monetary compensation, including the performance bonus, was the same as in Experiment 1.

Data from N=46 participants who completed the first session were excluded prior to analysis (see "Exclusion criteria" below), leaving a final sample of N=110 participants. These participants were re-invited to complete follow-up experiments 3 and 14 days later. (There was no follow-up session at 28 days due to overlap with the Christmas holiday.) Once invited, participants were permitted 48 hours to complete the follow-up experiment. Participant retention was again high for each follow-up session (Day 3: N=97 [88.2%]; Day 14: N=99 [90.0%]). Participants again received a retention bonus of $1.00 for each completed follow-up session. Detailed demographic information is presented in Table S1. The majority of participants identified as men (65 men; 53 women; 1 non-binary individual; 1 rather not say) and were 39.6 years old on average (range: 23–69 years).

Experimental protocol

The overall experimental protocol for Experiment 2 was almost identical to Experiment 1. In each session, participants started by completing the same self-report questionnaires with the exception that the 7-up/7-down was replaced with the 7-item depression subscale from the depression, anxiety, and stress scale (DASS; [53]). Participants completed the BIS/BAS scale once (on Day 0), but completed the GAD-7, DASS, and mood slider scales at the start of every session. These measures were included for exploratory analyses not reported here.

Next, participants completed a modified version of the gamified Pavlovian go/no-go task with a trial structure similar to [54]. In particular, instead of 8 unique robots each presented for 30 trials, participants saw a total of 24 unique robots presented for 8, 10, or 12 trials each. Each robot was presented for fewer trials as we were interested in measuring the learning process, where the expression of Pavlovian biases is typically largest, rather than asymptotic performance. Robots were presented to participants in mini-batches, each involving four robots and totalling approximately 40 trials. Crucially batches were not required to represent all four trial types (see Figure S2 for an example). That is, in any section of the task, participants were not guaranteed to observe one of each type of robot.

As such, learning about one robot did not imply information about another robot and participants could not rely on a top-down process-of-elimination strategy. Participants completed six mini-batches, which were divided into two blocks of 120 trials each (12 unique robots per block; three robots of each trial type; Figure 1D).

The task was visually similar to Experiment 1 except in two respects. First, the scanner colors were now blue and red (instead of blue and orange), and fixed such that these colors always indicated rewarding and punishing trials, respectively. This was intended to more naturally map on to reward (blue) and punishment (red) domains and potentially enhance Pavlovian biases. Second, the symbols on the robots' chestplates were drawn from one of two Brussels Artificial Character Sets [55] or the English alphabet (randomized within participants across sessions). These new symbols were used in order to accommodate the need for three times the number of distinctly recognizable robots. Pairwise comparisons revealed no significant differences in percent correct responses by character set (all $p > 0.90$, corrected for multiple comparisons). The timing of the task was also unchanged except the response window was shortened (from 1.5 to 1.3 seconds) and the feedback window was lengthened (from 1.0 to 1.2 seconds).

Exclusion criteria

Data from N=46 participants who completed the experiment on Day 0 were excluded prior to analysis for one or more of the following reasons: failing one or more attention checks embedded in the self-report measures (N=30; [44]); making either all Go or all No-Go responses on more than 90% of trials (N=5); exhibiting chance-level performance across all trials (<55% correct responses; N=22). These exclusion criteria left a final sample of N=110 participants. No exclusions were applied to subsequent session data.

Analyses

Analyses for Experiment 2 were identical to those for Experiment 1. In addition, we performed Wald tests to compare the magnitude of choice and practice effects between Experiments 1 and 2. P-values were derived from permutation testing, where a null distribution of values was obtained by permuting the experiment (1 or 2) and session labels (1, 2, or 3), across and within participants, respectively, 5,000 times.

Results

Descriptive analyses

Figure 4A shows trial-by-trial choice behavior for each session of the experiment. In contrast to Experiment 1 (c.f. Figure 2A), performance in all sessions conformed to the expected pattern of results. Group-averaged performance measures per session (Figure

4B) show that while performance improved after Day 0, improvement was only marginal. In particular, pairwise comparisons showed performance was significantly better on Day 3 compared to Day 0 ($p = 0.009$); however, no other pairwise comparisons were significant (complete descriptive statistics are reported in Table S6). In comparison to Experiment 1, performance accuracy on the modified task was lower (mean difference = 21.2%; $F(1,589)$ = 518.618, $p < 0.001$). This is to be expected given that the modified task was designed in part to prevent participants from reaching asymptotic performance. Crucially, practice effects (defined as the average difference in performance between the first and all follow-up sessions) were significantly reduced for the modified task in comparison to Experiment 1 (mean difference = -5.6%; $F(1,589) = 8.373$, $p < 0.001$).

In all sessions, participants performed better on Go trials than on No-Go trials ("Go bias"). The Go bias on Day 0 was significantly greater than that for all other sessions (all $p < 0.005$); no other between-session comparisons were significant. And although the practice effect for the Go bias was numerically smaller for the modified task, it was not significantly different than that for Experiment 1 (mean difference = -1.7%; $F(1,589)$ = 0.760, $p = 0.388$). Nevertheless, Go biases across sessions were significantly greater than those observed in Experiment 1 (mean difference = 8.6%; $F(1,589) = 88.026$, $p < 0.001$).

Participants also performed better on Pavlovian-instrumental congruent compared to incongruent trials in all sessions, manifesting a Pavlovian bias. The Pavlovian bias on Day 0 was significantly greater than that for all other sessions (both $p = 0.027$); no other between-session comparisons were significant. Unlike the Go bias, the practice effect for the Pavlovian bias was significantly reduced for the modified task in comparison to Experiment 1 (mean difference = -2.9%; $F(1,589) = 4.173$, $p = 0.037$). And like the Go bias, Pavlovian biases were significantly greater than those observed in Experiment 1 (mean difference = 5.3%; $F(1,589) = 59.284$, $p < 0.001$).

Regarding feedback sensitivity, across sessions participants made more correct responses following veridical compared to sham feedback. No pairwise comparison between sessions was significant (all $p > 0.10$), suggesting that feedback sensitivity was largely conserved across sessions. As a result, the practice effect for feedback sensitivity was significantly smaller in comparison to Experiment 1 (mean difference = -5.4%; $F(1,589) = 8.591$, $p < 0.001$). Moreover, feedback sensitivity was significantly greater across sessions than that observed in Experiment 1 (mean difference = 22.4%; $F(1,589) = 643.245$, $p < 0.001$). In sum, group-averaged behavior on the modified task showed evidence of residual practice effects. However, despite this, the expected choice biases were significantly larger than those observed in Experiment 1 and practice effects on the modified task were, with one exception, significantly reduced.

Turning next to individual variation in performance, the proportion of participants who exhibited chance-level, intermediate, or near-ceiling performance by session and trial type is presented in Figure 4C. In contrast to Experiment 1, ceiling performance was relatively rare and the majority of participants exhibited intermediate levels of performance across all trial types and sessions (the only exception was for NGW trials on Day 0, where the majority of participants showed chance-level performance). Two-way chi-squared tests of
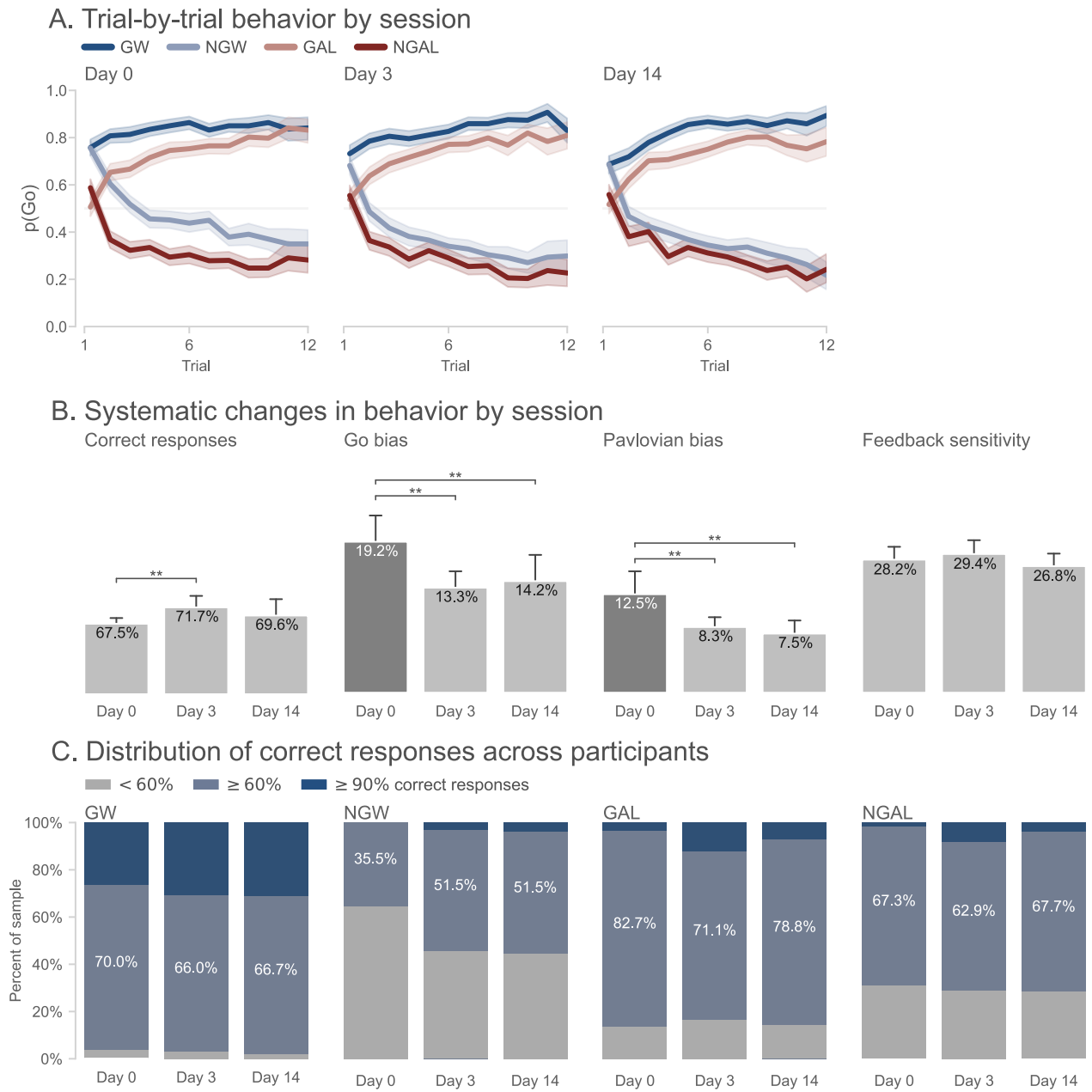
14

Figure 4: **Smaller or no practice effects on the modified Pavlovian go/no-go task in Experiment 2**. (A) Group-averaged learning curves for each trial type and session. Shaded regions indicate 95% bootstrapped confidence intervals. (B) Group-averaged performance for each session. Performance indices from left-to-right: Correct responses, or overall accuracy; Go bias, or difference in accuracy between Go and No-Go trials; Congruency effect, or difference in accuracy between Pavlovian congruent (GW, NGAL) and incongruent (NGW, GAL) trials; and Feedback sensitivity, or the difference in accuracy on trials following veridical and sham feedback. ** Denotes significant pairwise difference ($p < 0.05$, corrected for multiple comparisons). (C) The percentage of participants, for each session and trial type, exhibiting at- or below-chance performance ($< 60\%$ response accuracy; grey), intermediate performance ($\geq 60\%$ and $< 90\%$ response accuracy; light blue), or near-perfect performance ($\geq 90\%$ response accuracy; dark blue).

independence confirmed that, with an exception for NGW trials, no significant shift in participants' performance across sessions was observed (GW: $\chi^2(4) = 1.163$, $p = 0.884$; NGW: $\chi^2(4) = 13.343$, $p = 0.010$; GAL: $\chi^2(4) = 6.499$, $p = 0.165$; NGAL: $\chi^2(4) = 5.097$, $p = 0.278$). Thus, the majority of participants exhibited and maintained intermediate levels of performance on the modified Pavlovian go/no-go task.

Model comparison

Results of the model comparison are summarized in Table 2. Trial-level choice prediction for all models was worse in Experiment 2 than in Experiment 1, which is to be expected insofar as it is easier to predict asymptotic behavior, whereas the modified task primarily measures participants' performance during learning (i.e., when choice is most stochastic). As in Experiment 1, collapsing across sessions, the best-fitting model was M7, the most complex model. This was also the best-fitting model within each session (Table S7). Posterior predictive checks indicated that this model provided excellent fits to the choice data from each session (Figure S4).

Model parameters

The estimated group-level parameters from the best-fitting model are presented in Figure 5A. In comparison to Experiment 1, we observed smaller but still significant changes in the reward and punishment sensitivity parameters across days. Specifically, reward sensitivity (β_+) was significantly larger on Day 14 compared to Days 0 and 3, whereas punishment sensitivity (β_-) was significantly larger on Days 3 and 14 compared to Day 0. Both the reward and punishment sensitivity parameters were on average smaller in Experiment 2 as compared to Experiment 1 (reward sensitivity: mean difference between experiments = -15.895, 95% CI = [11.934, 19.832]; punishment sensitivity: mean difference = -14.771, 95% CI = [10.531, 18.708]). Practice effects manifest in this task

Model	Parameters	Accuracy	PSIS-LOO	ΔPSIS-LOO (se)
M1	β, η	72.9%	-95806.3	-6205.2 (73.2)
M2	β, τ, η	76.5%	-99616.0	-2395.5 (48.9)
M3	$\beta, \tau_+, \tau_-, \eta$	77.6%	-101283.0	-728.5 (28.2)
M4	$\beta_+, \beta_-, \tau_+, \tau_-, \eta$	77.5%	-101422.4	-589.0 (21.1)
M5	$\beta, \tau_+, \tau_-, \eta_+, \eta_-$	77.7%	-101519.0	-492.4 (19.1)
M6	$\beta_+, \beta_- \tau_+, \tau_-, \eta_+, \eta_-$	77.8%	-101548.7	-462.7 (17.2)
M7	$\beta_+, \beta_- \tau_+, \tau_-, \eta_+, \eta_-, \xi$	78.1%	-102011.4	-

Table 2: **Model comparison collapsing across sessions.** Accuracy = trial-level choice prediction accuracy between observed and model-predicted Go responses. PSIS-LOO = approximate leave-one-out cross-validation presented in deviance scale (smaller numbers indicate better fit). ΔPSIS-LOO = difference in PSIS-LOO values between each model and the best-fitting model (M7).

as increases in the proportion of correct responses in follow-up sessions. In the model, this appears as a between-sessions increase in the reward and punishment sensitivity parameters. Therefore, one way to quantify practice effects is as the difference in reward and punishment sensitivity parameters between Day 0 and the average of all other days. This difference was significantly smaller in Experiment 2 compared to Experiment 1 (reward sensitivity: mean difference = -17.106, 95% CI = [-23.753, -11.385]; punishment sensitivity: mean difference = -11.700, 95% CI = [-18.486, -4.777]).

The inverse pattern was observed for reward (η_+) and punishment learning rates (η_-): reward learning rates were significantly higher on Day 14 compared to Days 0 and 3, while punishment learning rates were significantly lower on Days 3 and 14 compared to Day 0. In comparison to Experiment 1, both reward and punishment learning rates were greater on average (reward learning rate: mean difference = 0.133, 95% CI = [0.094, 0.170]; punishment learning rate: mean difference = 0.186, 95% CI = [0.079, 0.301]). Practice effects for reward learning rates were not significantly different between the two
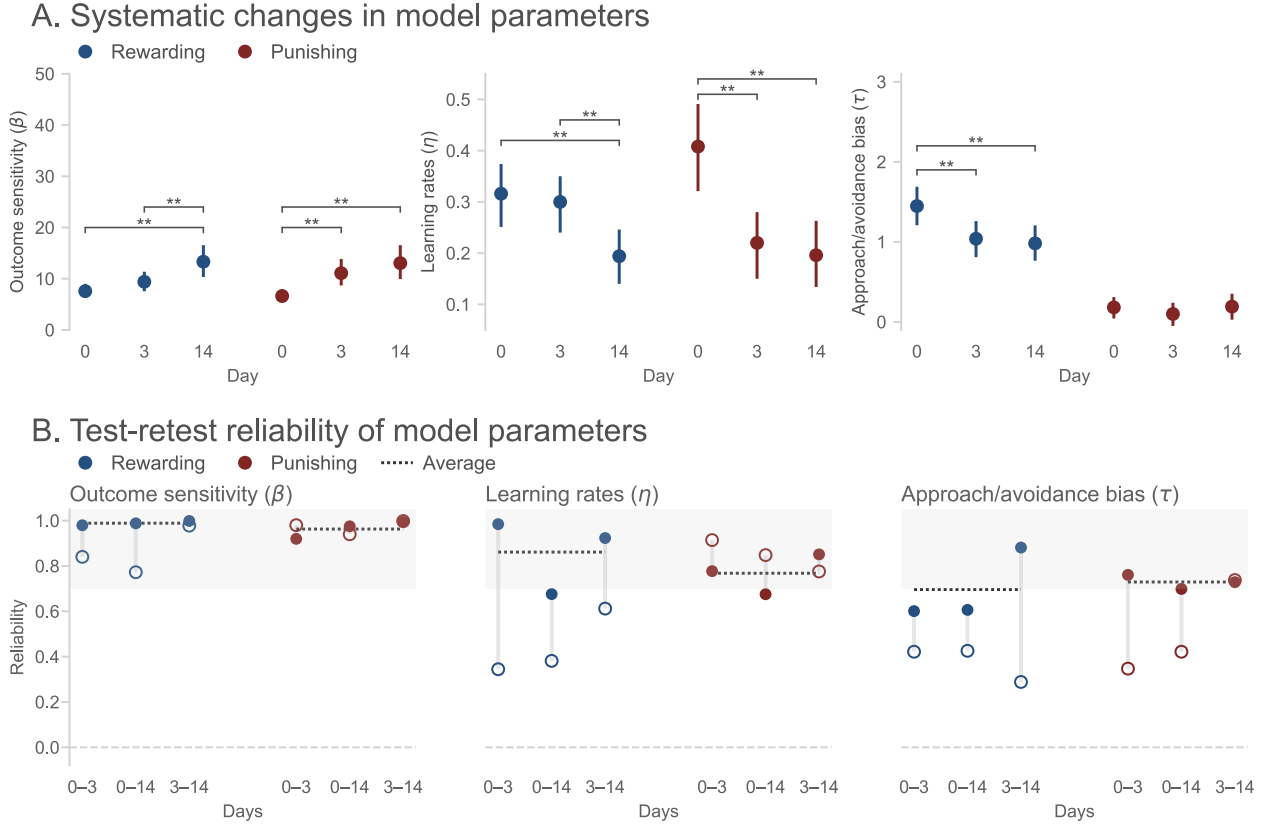

Figure 5: **Reinforcement learning model parameters in Experiment 2 show improved stability and reliability**. (A) Group-level model parameters for each session. Error bars indicate 95% Bayesian confidence intervals (CIs). ** Denotes pairwise comparison where 95% CI of the difference excludes zero. (B) Test-retest reliability estimates for each model parameter. Filled circles denote estimates for Experiment 2; open circles denote estimates from Experiment 1, for comparison. Grey vertical lines show the change in reliability across experiments. Dotted lines indicates average reliability for Experiment 2. Shaded region indicates conventional range of acceptable reliability ($\rho \geq 0.7$).

experiments (mean difference = -0.015, 95% CI = [-0.098, 0.068]), but were in fact larger for the punishment learning rate in Experiment 2 (mean difference = 0.186, 95% CI =[0.079, 0.301]).

Finally, the approach bias (τ_+) was slightly but significantly larger on Day 0 compared to Days 3 and 14. No significant differences across sessions were observed in the avoidance bias (τ_-). Therefore, although Pavlovian biases were somewhat diminished through repeated testing, in both absolute and relative terms (i.e., compared to the outcome sensitivity parameters), they remained largely intact in later sessions. Relative to the magnitude of the reward sensitivity parameter, the approach bias was significantly larger on average in the modified task than in Experiment 1 (mean difference = 0.049, 95% CI = [0.027, 0.071]), although practice effects between the two experiments were not significantly different (mean difference = -0.015, 95% CI = [-0.070, 0.041]). Also in relative terms, the avoidance bias was not significantly different between the two experiments (mean difference = 0.004, 95% CI = [-0.006, 0.015]), nor was the difference in practice effects (mean difference = 0.020, 95% = [-0.005, 0.045]). The Pavlovian bias (defined here as the difference between the approach and avoidance parameters) was significantly greater in the modified task compared to Experiment 1 (mean difference = 0.045, 95% CI = [0.022, 0.071]). Thus, in line with the descriptive results, Pavlovian biases were larger in the modified task despite the residual practice effects.

The estimated test-retest reliability of the model parameters is presented in Figure 5B. In contrast to Experiment 1, averaging across session pairs, acceptable test-retest reliability was observed for essentially all parameters (reward sensitivity: average ρ = 0.989, 95% CI = [0.984, 0.992]; punishment sensitivity: average ρ = 0.963, 95% CI = [0.949, 0.973]; approach bias: average ρ = 0.696, 95% CI = [0.622, 0.765]; avoidance bias: average ρ = 0.729, 95% CI = [0.669, 0.778]; reward learning rate: average ρ = 0.861, 95% CI = [0.817, 0.898]; punishment learning rate: average ρ = 0.768, 95% CI = [0.710, 0.815]). Compared to Experiment 1, test-retest reliability was significantly improved for reward sensitivity (change in average ρ = 0.125, 95% CI = [0.101, 0.152]), approach bias (change in average ρ = 0.318, 95% CI = [0.185, 0.446]), avoidance bias (change in average ρ = 0.227, 95% CI = [0.126, 0.315]), and reward learning rate (change in average ρ = 0.415, 95% CI = [0.294, 0.531]); no parameters showed significantly worsened reliability. A similar pattern of results was observed for the split-half reliability estimates (Figure S5B).

Discussion

The goal of the second experiment was to evaluate the stability and reliability of individual differences in performance on a modified version of the Pavlovian go/no-go task that was designed to keep participants learning and to lessen practice effects. At the group level, participants showed the desired behavioral effects (e.g., go bias, Pavlovian bias, and feedback sensitivity) at significantly greater levels than observed in Experiment 1 across all sessions. Although participants continued to exhibit practice effects on the modi-

fied task, these were significantly reduced for the majority of task performance indices. Moreover, the fraction of participants maintaining an intermediate level of performance was largely conserved across sessions. These findings were reflected in the parameters of a reinforcement learning model fit to participants' choice data, where parameters were largely stable and consequently exhibited acceptable test-retest reliability.

General Discussion

Despite considerable use in individual-differences and computational psychiatry research, previous studies of the psychometric properties of the Pavlovian go/no-go task found that both descriptive and model-based measures of task performance showed poor reliability [20–22]. Here, we investigated the psychometric properties of two variants of the task in an attempt to develop a more reliable version – one that would be usable in clinical practice where patients may perform a task multiple times (e.g., before, during, and after treatment). In the first experiment, we used a gamified version of the standard task. Here, we observed considerable practice effects whereby the majority of participants exhibited near-ceiling levels of performance with repeat testing. Consequently, the test-retest reliability of multiple reinforcement-learning model parameters estimated from participants' behavior was unacceptable. To address these issues, in Experiment 2 we designed a version of the task that measures choice behavior primarily during learning and prevents undesirable process-of-elimination strategies. Participants exhibited reduced practice effects on this version of the task and, as a consequence, the test-retest reliability of reinforcement-learning model parameters was significantly improved.

The estimates of model-parameter reliability observed in both our experiments were larger than previously reported for the Pavlovian Go/No-Go task [20–22]. This likely reflects a confluence of factors. First, both versions of the task studied here were gamified. Gamification has previously been shown to promote participant engagement and minimize confusion [28] and benefit the reliability of cognitive task measures [29, 30]. Second, we used a hierarchical Bayesian modeling framework to estimate model parameters for the reliability analyses. Hierarchical models exert a pooling or regularization effect on model parameters, which decreases measurement error and improves estimates of reliability [31, 32]. Indeed, our results are consistent with previous empirical studies that have demonstrated the benefits of hierarchical Bayesian models for estimating parameter reliability [34, 35]. Finally, in Experiment 2, we redesigned the trial structure of the Pavlovian go/no-go task such as to prevent practice effects. Practice effects can harm reliability when they induce ceiling performance (as in Experiment 1) or when they are not uniformly expressed by participants (e.g., as a function of age [56]). It is possible that such effects worsened reliability estimates in a prior study where practice effects were observed in an adolescent sample [20].

The occurrence of practice effects with repeated administrations is common for cognitive tasks [57, 58]. Practice effects may reflect a number of factors, such as reductions in performance anxiety or the acquisition of task-specific knowledge or strategies. In Ex-

periment 1, practice effects were ostensibly attributable to participants adopting a qualitatively different strategy after their initial completion of the Pavlovian go/no-go task. Specifically, participants were able to exploit acquired knowledge of implicit dependencies between stimuli in the task to develop a process-of-elimination strategy that resulted in rapid learning and the attenuation of the desired choice biases. To address this issue, in Experiment 2 we redesigned the task to eliminate these dependencies and the formation of such a top-down strategy. This approach is consistent with previous research, whereby preventing participants from becoming aware of critical elements of a task design resulted in improved consistency and reliability of behavior, even with practice [36].

It is important to note that although practice effects were reduced in our modified version of the Pavlovian go/no-go task, they were not eliminated altogether. Indeed, we observed smaller but still significant reductions in participants' go and Pavlovian biases (with corresponding decreases in the approach bias model parameter) following the initial test session. For the purposes of individual-differences correlational research, these residual practice effects are tolerable because the reliabilities of the model parameters are still in an acceptable range. However, they may be worrisome for longitudinal studies where systematic changes in task performance are of interest (e.g., reduction in Pavlovian biases following psychotherapy [59]). One possible solution might be increasing the length of the practice block, which was relatively brief in this study, and could be extended to help participants reach "steady state" performance prior to starting the actual task. Indeed, our results showed stability of performance on days 3 and 14, suggesting that task administrations after a longer practice may be usable for measuring changes in performance over the course of a mental health condition or treatment.

The current study has several notable limitations. We investigated the psychometric properties of two versions of the Pavlovian go/no-go task in a sample of online adult participants. The reliability of task measures, however, can vary as a function of the sample and the test setting. For example, previous research has shown that the reliability of a task completed by healthy adults can differ from that for adults with psychopathology [60] or healthy children [61]. Importantly, our general sample of adult participants rated the modified Pavlovian go/no-go task as more mentally demanding than the original task (see Table S2). As such, our task may prove to be too challenging for other groups (e.g., children; patients) which may affect reliability. Future research is therefore necessary to validate the modified version of the task in other populations, or develop simplified variants of it.

A second limitation is that we only studied participants' choice behavior. Previous studies have found that Pavlovian biases also manifest in response times [62, 63], and these may be a meaningful index of individual differences [6, 15, 64]. Previous work also introduced a computational framework for jointly modeling participants' choice and response time behavior on the task [6, 62]. This is notable because joint modeling of choice and response time had been found to improve the precision and reliability of parameter estimates from reinforcement learning models [65, 66]. As such, more research is warranted to investigate how the reliability of model-derived measures of behavior on the Pavlovian go/no-go task

could be further improved by incorporating response times.

Limitations notwithstanding, our study demonstrates that it is possible to derive performance measures from the Pavlovian go/no-go task that are sufficiently reliable for use in individual-differences research. We encourage researchers to use and further adapt the modified version of the task presented here. In support of this goal, we have made all of our data and code publicly available (see Data and Code Availability statements).

Data availability

The data that support the findings of this study are openly available on Github at `https://github.com/nivlab/RobotFactory`.

Code availability

All code for data cleaning and analysis associated with this study is available at `https://github.com/nivlab/RobotFactory`. The experiment code is available at the same link. The custom web-software for serving online experiments is available at `https://github.com/nivlab/nivturk`. A playable demo of the task is available at `https://nivlab.github.io/jspsych-demos/tasks/pgng/experiment.html`.

Citation diversity statement

Recent work in several fields of science has identified a bias in citation practices such that papers from women and other minority scholars are under-cited relative to the number of such papers in the field [67, 68]. Here we sought to proactively consider choosing references that reflect the diversity of the field in thought, form of contribution, gender, race, ethnicity, and other factors. First, we obtained the predicted gender of the first and last author of each reference by using databases that store the probability of a first name being carried by a woman [67]. By this measure (and excluding self-citations to the first and last authors of our current paper), our references contain 9.52% woman(first)/woman(last), 15.87% man/woman, 23.81% woman/man, and 50.79% man/man. This method is limited in that a) names, pronouns, and social media profiles used to construct the databases may not, in every case, be indicative of gender identity and b) it cannot account for intersex, non-binary, or transgender people. Second, we obtained predicted racial/ethnic category of the first and last author of each reference by databases that store the probability of a first and last name being carried by an author of color [69, 70]. By this measure (and excluding self-citations), our references contain 3.82% author of color (first)/author of color(last), 12.98% white author/author of color, 16.75% author of color/white author, and 66.46% white author/white author. This method is limited in that a) names and Florida Voter Data to make the predictions may not be indicative of racial/ethnic identity, and b) it cannot account for Indigenous and mixed-race authors, or those who may face differential biases due to the ambiguous racialization or ethnicization of their names. We look forward to future work that could help us to better understand how to support equitable practices in science.

Acknowledgements

Competing Interests Statement

The authors declare no competing interests.

References

1. Carver, C. S. & White, T. L. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of personality and social psychology* **67,** 319 (1994).

2. Guitart-Masip, M. *et al.* Go and no-go learning in reward and punishment: interactions between affect and effect. *Neuroimage* **62,** 154–166 (2012).

3. Guitart-Masip, M., Duzel, E., Dolan, R. & Dayan, P. Action versus valence in decision making. *Trends in cognitive sciences* **18,** 194–202 (2014).

4. Mkrtchian, A., Aylward, J., Dayan, P., Roiser, J. P. & Robinson, O. J. Modeling avoidance in mood and anxiety disorders using reinforcement learning. *Biological psychiatry* **82,** 532–539 (2017).

5. Peterburs, J., Albrecht, C. & Bellebaum, C. The impact of social anxiety on feedback-based go and nogo learning. *Psychological Research,* 1–15 (2021).

6. Millner, A. J. *et al.* Suicidal thoughts and behaviors are associated with an increased decision-making bias for active responses to escape aversive states. *Journal of abnormal psychology* **128,** 106 (2019).

7. Ousdal, O. T. *et al.* The impact of traumatic stress on Pavlovian biases. *Psychological medicine* **48,** 327–336 (2018).

8. Montagnese, M. *et al.* Reinforcement learning as an intermediate phenotype in psychosis? Deficits sensitive to illness stage but not associated with polygenic risk of schizophrenia in the general population. *Schizophrenia research* **222,** 389–396 (2020).

9. Huys, Q. J. *et al.* The specificity of Pavlovian regulation is associated with recovery from depression. *Psychological medicine* **46,** 1027–1035 (2016).

10. Albrecht, M. A., Waltz, J. A., Cavanagh, J. F., Frank, M. J. & Gold, J. M. Reduction of Pavlovian bias in schizophrenia: Enhanced effects in clozapine-administered patients. *PLoS One* **11,** e0152781 (2016).

11. Eisinger, R. S. *et al.* Pavlovian bias in Parkinson's disease: an objective marker of impulsivity that modulates with deep brain stimulation. *Scientific Reports* **10,** 13448 (2020).

12. Richter, A. *et al.* Valenced action/inhibition learning in humans is modulated by a genetic variant linked to dopamine D2 receptor expression. *Frontiers in Systems Neuroscience* **8,** 140 (2014).

13. Richter, A. *et al.* Motivational learning biases are differentially modulated by genetic determinants of striatal and prefrontal dopamine function. *Journal of Neural Transmission* **128,** 1705–1720 (2021).

14. Raab, H. A. & Hartley, C. A. Adolescents exhibit reduced Pavlovian biases on instrumental learning. *Scientific reports* **10,** 1–11 (2020).

15. Betts, M. J. *et al.* Learning in anticipation of reward and punishment: perspectives across the human lifespan. *Neurobiology of aging* **96,** 49–57 (2020).

16. Weber, I., Zorowitz, S., Niv, Y. & Bennett, D. The effects of induced positive and negative affect on Pavlovian-instrumental interactions. *Cognition and Emotion,* 1–18 (2022).

17. Wonderlich, J. A. *Anger and Pavlovian Bias: Integrating Laboratory Task Performance and Ecological Momentary Assessment* PhD thesis (George Mason University, 2020).

18. De Berker, A. O. *et al.* Acute stress selectively impairs learning to act. *Scientific reports* **6,** 1–12 (2016).

19. Mkrtchian, A., Roiser, J. P. & Robinson, O. J. Threat of shock and aversive inhibition: Induced anxiety modulates Pavlovian-instrumental interactions. *Journal of Experimental Psychology: General* **146,** 1694 (2017).

20. Moutoussis, M. *et al.* Change, stability, and instability in the Pavlovian guidance of behaviour from adolescence to young adulthood. *PLoS computational biology* **14,** e1006679 (2018).

21. Pike, A. C., Tan, K., Ansari, H. J., Wing, M. & Robinson, O. J. Test-retest reliability of affective bias tasks. *PsyArXiv* (2022).

22. Saeedpour, S., Hossein, M. M., Deroy, O. & Bahrami, B. Interindividual differences in Pavlovian influence on learning are consistent. *Royal Society Open Science* **10,** 230447 (2023).

23. Spearman, C. "General intelligence", objectively determined and measured. *Am. J. Psychol.* **15,** 201–293 (1904).

24. Parsons, S., Kruijt, A.-W. & Fox, E. Psychological Science Needs a Standard Practice of Reporting the Reliability of Cognitive-Behavioral Measurements. *Advances in Methods and Practices in Psychological Science* **2,** 378–395 (Dec. 2019).

25. Gelman, A. & Carlin, J. Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science* **9,** 641–651 (2014).

26. Schurr, R., Reznik, D., Hillman, H., Bhui, R. & Gershman, S. J. Dynamic computational phenotyping of human cognition. *Nature Human Behaviour,* 1–15 (2024).

27. Zorowitz, S. & Niv, Y. Improving the reliability of cognitive task measures: A narrative review. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* (2023).

28. Sailer, M., Hense, J. U., Mayr, S. K. & Mandl, H. How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in human behavior* **69,** 371–380 (2017).

29. Kucina, T. *et al.* Calibration of cognitive tests to address the reliability paradox for decision-conflict tasks. *Nature Communications* **14,** 2234 (2023).

30. Verdejo-Garcia, A. *et al.* A unified online test battery for cognitive impulsivity reveals relationships with real-world impulsive behaviours. *Nature Human Behaviour* **5,** 1562–1577 (2021).

31. Haines, N., Sullivan-Toole, H. & Olino, T. From Classical Methods to Generative Models: Tackling the Unreliability of Neuroscientific Measures in Mental Health Research. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* (2023).

32. Rouder, J. N. & Haaf, J. M. A psychometrics of individual differences in experimental tasks. *Psychonomic bulletin & review* **26,** 452–467 (2019).

33. Sullivan-Toole, H., Haines, N., Dale, K. & Olino, T. Enhancing the Psychometric Properties of the Iowa Gambling Task Using Full Generative Modeling. *Faculty/Researcher Works* (2022).

34. 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. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **5,** 601–609 (2020).

35. Waltmann, M., Schlagenhauf, F. & Deserno, L. Sufficient reliability of the behavioral and computational readouts of a probabilistic reversal learning task. *Behavior research methods,* 1–22 (2022).

36. McLean, B. F., Mattiske, J. K. & Balzan, R. P. Towards a reliable repeated-measures beads task for assessing the jumping to conclusions bias. *Psychiatry research* **265,** 200–207 (2018).

37. Litman, L., Robinson, J. & Abberbock, T. TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior research methods* **49,** 433–442 (2017).

38. Robinson, J., Rosenzweig, C., Moss, A. J. & Litman, L. Tapped out or barely tapped? Recommendations for how to harness the vast and largely unused potential of the Mechanical Turk participant pool. *PloS one* **14,** e0226394 (2019).

39. Spitzer, R. L., Kroenke, K., Williams, J. B. & Löwe, B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine* **166,** 1092–1097 (2006).

40. Youngstrom, E. A., Murray, G., Johnson, S. L. & Findling, R. L. The 7 up 7 down inventory: a 14-item measure of manic and depressive tendencies carved from the General Behavior Inventory. *Psychological assessment* **25,** 1377 (2013).

41. Pagliaccio, D. *et al.* Revising the BIS/BAS Scale to study development: Measurement invariance and normative effects of age and sex from childhood through adulthood. *Psychological assessment* **28,** 429 (2016).

42. Betella, A. & Verschure, P. F. The affective slider: A digital self-assessment scale for the measurement of human emotions. *PloS one* **11,** e0148037 (2016).

43. De Leeuw, J. R. jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior research methods* **47,** 1–12 (2015).

44. Zorowitz, S., Solis, J., Niv, Y. & Bennett, D. Inattentive responding can induce spurious associations between task behaviour and symptom measures. *Nature Human Behaviour,* 1–15 (2023).

45. Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M. & Nichols, T. E. Permutation inference for the general linear model. *Neuroimage* **92,** 381–397 (2014).

46. Rubin, M. When to adjust alpha during multiple testing: A consideration of disjunction, conjunction, and individual testing. *Synthese* **199,** 10969–11000 (2021).

47. Swart, J. C. *et al.* Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in) action. *Elife* **6,** e22169 (2017).

48. Carpenter, B. *et al.* Stan: A probabilistic programming language. *Journal of statistical software* **76,** 1–32 (2017).

49. Vehtari, A., Gelman, A. & Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and computing* **27,** 1413–1432 (2017).

50. Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J. & Friston, K. J. Bayesian model selection for group studies. *Neuroimage* **46,** 1004–1017 (2009).

51. Vehtari, A. *Cross-validation FAQ* en. https://mc-stan.org/loo/articles/online-only/faq.html. Accessed: 2023-2-2.

52. Cicchetti, D. V. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment* **6,** 284 (1994).

53. Henry, J. D. & Crawford, J. R. The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British journal of clinical psychology* **44,** 227–239 (2005).

54. Wittmann, B. C., Daw, N. D., Seymour, B. & Dolan, R. J. Striatal activity underlies novelty-based choice in humans. *Neuron* **58,** 967–973 (2008).

55. Vidal, C., Content, A. & Chetail, F. BACS: The Brussels Artificial Character Sets for studies in cognitive psychology and neuroscience. *Behavior Research Methods* **49,** 2093–2112 (2017).

56. Anokhin, A. P. *et al.* Age-related changes and longitudinal stability of individual differences in ABCD Neurocognition measures. *Developmental Cognitive Neuroscience* **54,** 101078 (2022).

57. Hausknecht, J. P., Halpert, J. A., Di Paolo, N. T. & Moriarty Gerrard, M. O. Retesting in selection: a meta-analysis of coaching and practice effects for tests of cognitive ability. *Journal of Applied Psychology* **92,** 373 (2007).

58. Scharfen, J., Peters, J. M. & Holling, H. Retest effects in cognitive ability tests: A meta-analysis. *Intelligence* **67,** 44–66 (2018).

59. Geurts, D. E. *et al.* Aversive Pavlovian inhibition in adult attention-deficit/hyperactivity disorder and its restoration by mindfulness-based cognitive therapy. *Frontiers in Behavioral Neuroscience,* 289 (2022).

60. Cooper, S. R., Gonthier, C., Barch, D. M. & Braver, T. S. The role of psychometrics in individual differences research in cognition: A case study of the AX-CPT. *Frontiers in psychology* **8,** 1482 (2017).

61. Arnon, I. Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behavior research methods* **52,** 68–81 (2020).

62. Millner, A. J., Gershman, S. J., Nock, M. K. & den Ouden, H. E. Pavlovian control of escape and avoidance. *Journal of Cognitive Neuroscience* **30,** 1379–1390 (2018).

63. Algermissen, J., Swart, J. C., Scheeringa, R., Cools, R. & den Ouden, H. E. Striatal BOLD and midfrontal theta power express motivation for action. *Cerebral Cortex* **32,** 2924–2942 (2022).

64. Scholz, V., Kandroodi, M. R., Algermissen, J. & den Ouden, H. Dissociable effects of mood-anxiety and compulsive symptom dimensions on motivational biases in decision-making. *Biological Psychiatry* **87,** S382–S383 (2020).

65. Ballard, I. C. & McClure, S. M. Joint modeling of reaction times and choice improves parameter identifiability in reinforcement learning models. *Journal of Neuroscience Methods* **317,** 37–44 (2019).

66. Shahar, N. *et al.* Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. *PLoS computational biology* **15,** e1006803 (2019).

67. Dworkin, J. D. *et al.* The extent and drivers of gender imbalance in neuroscience reference lists. *Nature neuroscience* **23,** 918–926 (2020).

68. Bertolero, M. A. *et al.* Racial and ethnic imbalance in neuroscience reference lists and intersections with gender. *BioRxiv* (2020).

69. Ambekar, A., Ward, C., Mohammed, J., Male, S. & Skiena, S. *Name-ethnicity classification from open sources* in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge Discovery and Data Mining* (2009), 49–58.

70. Sood, G. & Laohaprapanon, S. Predicting race and ethnicity from the sequence of characters in a name. *arXiv preprint arXiv:1805.02109* (2018).

Supplementary materials

Comparative stability of self-report measures

A. Systematic changes in self-report measures

B. Test-retest reliability of self-report measures
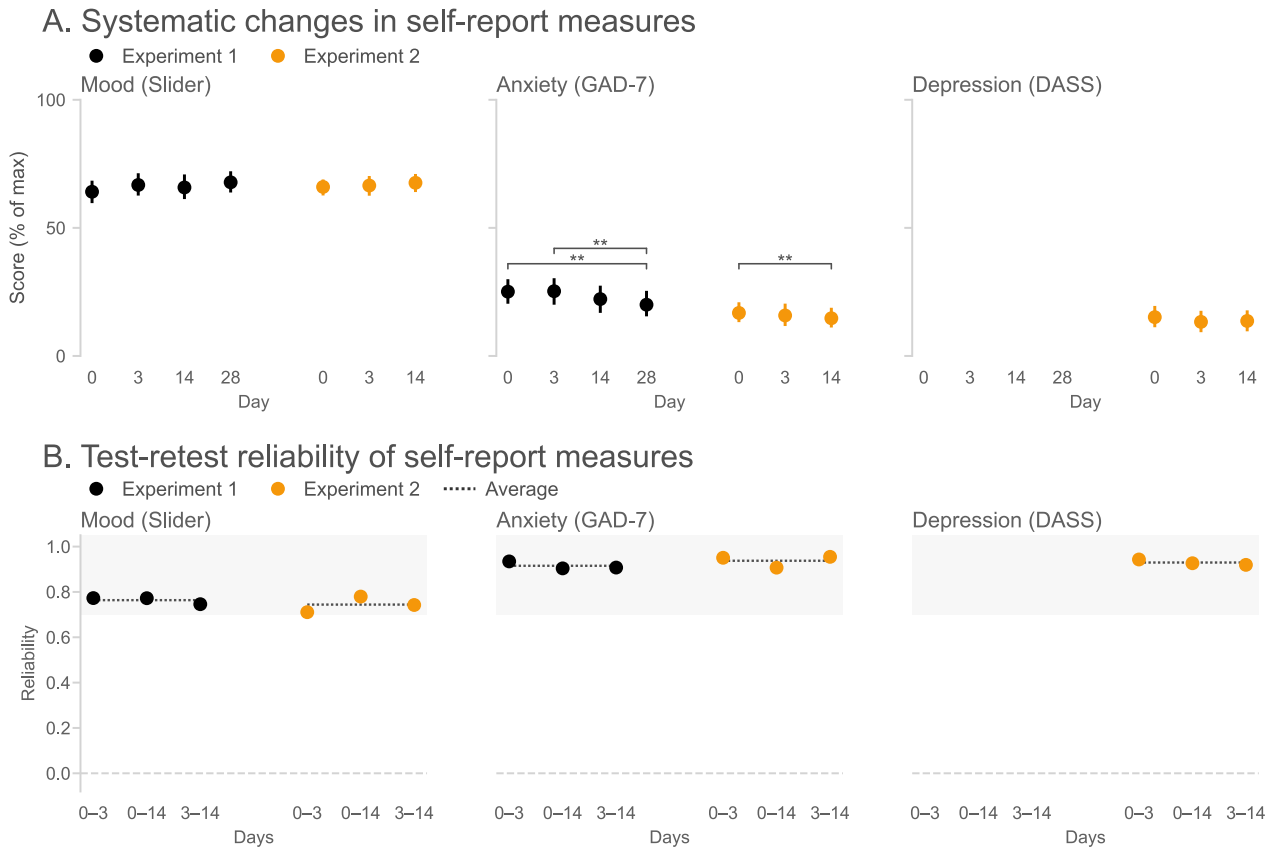
Figure S1: **Self-report measures suggest that mood and anxiety were relatively stable over the study period across participants.** (A) Group-level mood and symptom scores for each session. Error bars indicate 95% bootstrapped confidence intervals. ** Denotes significant pairwise difference ($p < 0.05$, corrected for multiple comparisons). (B) Test-retest reliability estimates for each self-report measure. Dotted lines indicate overall average. Shaded regions indicate conventional range of acceptable reliability ($\rho \geq 0.7$).

Example block of the modified Pavlovian go/no-go task
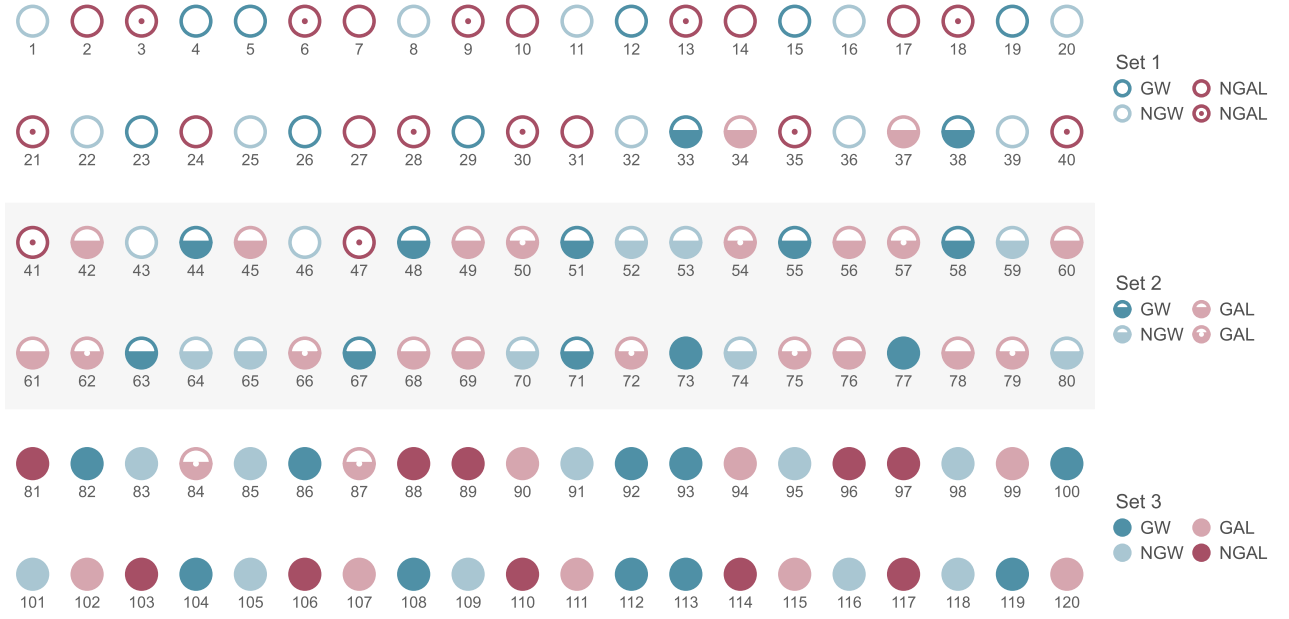
Figure S2: **Trial structure of an example block of the modified Pavlovian go/no-go task from Experiment 2**. In one task block, there were 12 unique stimuli (three of each trial type: go to win [GW; dark blue]; no-go to win [NGW; light blue]; go to avoid losing [GAL; light red]; no-go to avoid losing [NGAL; dark red]) divided into three sets (Set 1: empty circles; Set 2: half-filled circles; Set 3: filled circles). Each set was composed of approximately 40 trials (120 trials total). Each set, however, did not necessarily involve all four trial types. In this example block, Set 1 involved two NGAL stimuli and no GAL stimuli, and Set 2 involved two GAL stimuli and no NGAL stimuli.

Posterior predictive check for the best-fitting reinforcement learning model (Experiment 1)

A. Observed & model-predicted learning curves

B. Observed & model-predicted choice behavior
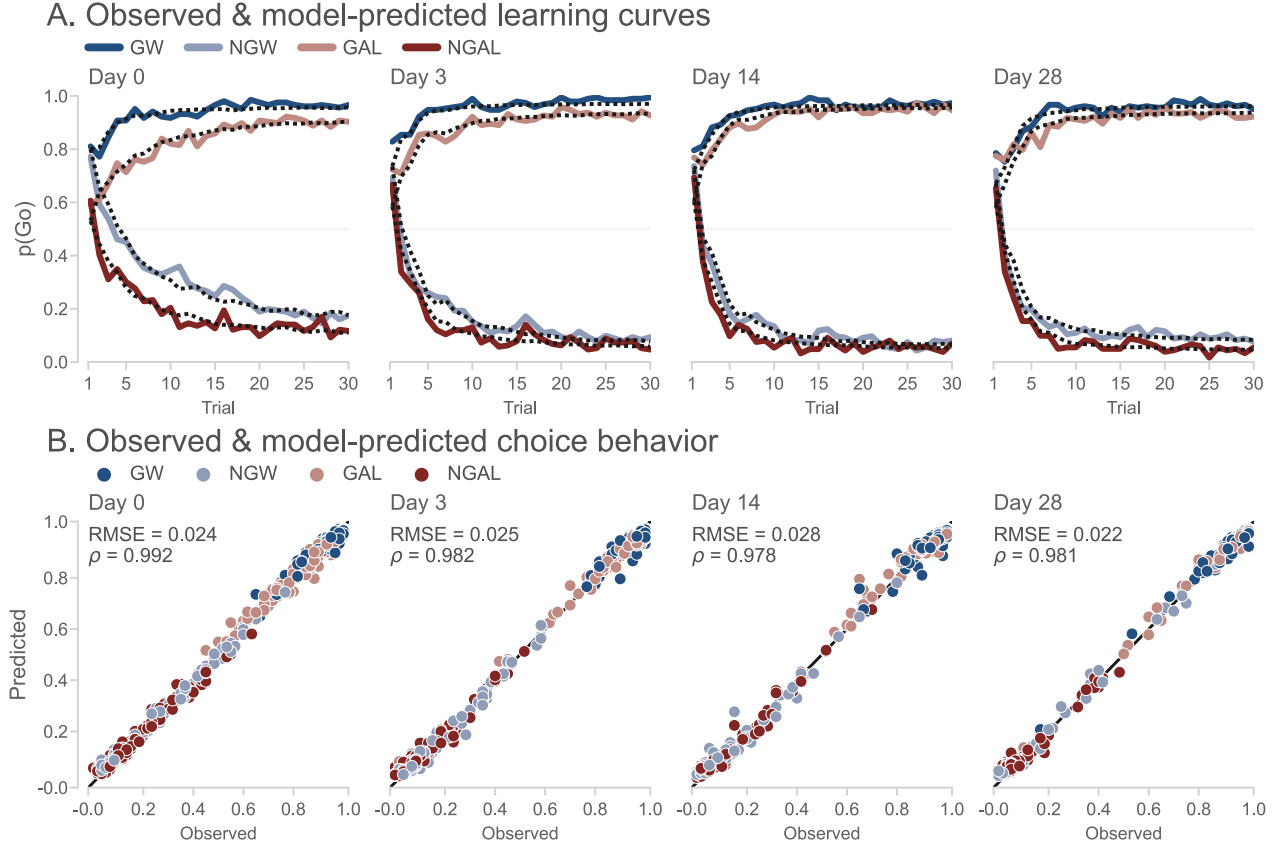
Figure S3: **Observed & model-predicted choice behavior for Experiment 1.** (A) Trial-by-trial choice behavior. Solid and dotted lines depict observed and model-predicted choice behavior, respectively. (B) Observed (x-axis) and model-predicted (y-axis) proportions of go responses for each trial type. Each point corresponds to one participant and condition. RMSE = root-mean-squared error between observed and model-predicted choice behavior. ρ = Spearman's rank correlation between observed and model-predicted choice behavior. All ρ's were exceptionally high at > 0.975.

Posterior predictive check for the best-fitting reinforcement learning model (Experiment 2)

A. Observed & model-predicted learning curves

B. Observed & model-predicted choice behavior
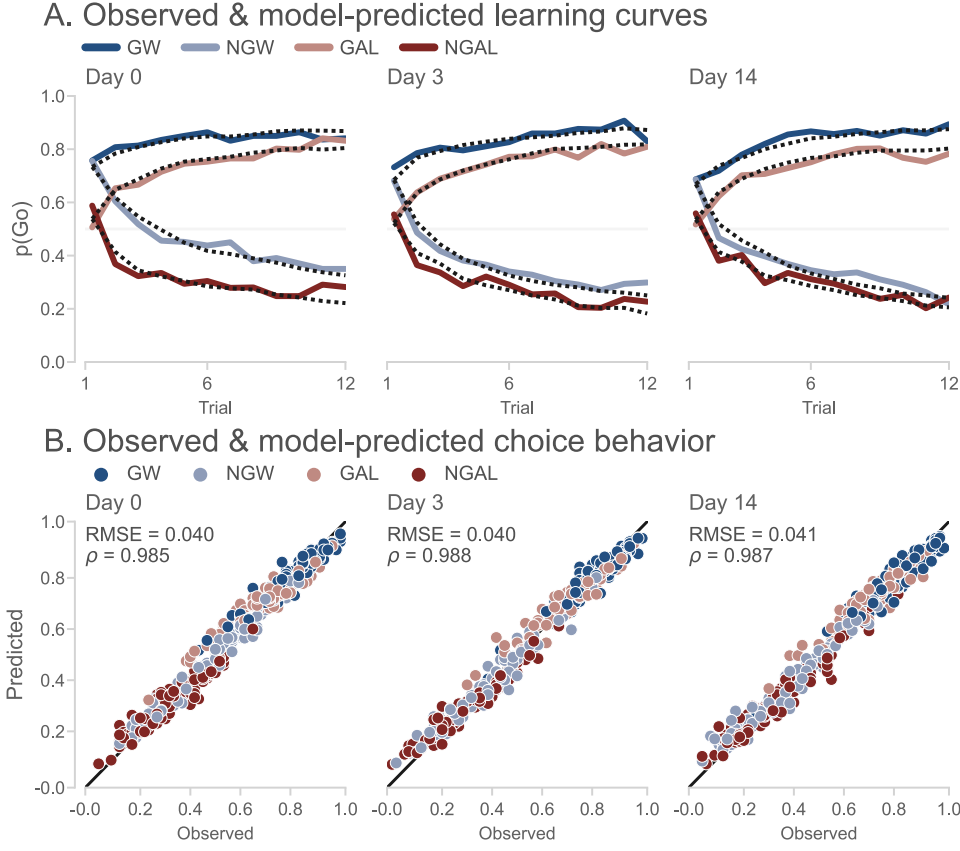
Figure S4: **Observed & model-predicted choice behavior for Experiment 2.** (A) Trial-by-trial choice behavior. Solid and dotted lines depict observed and model-predicted choice behavior, respectively. (B) Observed (x-axis) and model-predicted (y-axis) proportions of go responses for each trial type. Each point corresponds to one participant and condition. RMSE = root-mean-squared error between observed and model-predicted choice behavior. ρ = Spearman's rank correlation between observed and model-predicted choice behavior. Here, too, all ρ's were exceptionally high at > 0.985.

Split-half reliability of model parameters

A. Split-half reliability of model parameters (Experiment 1)

B. Split-half reliability of model parameters (Experiment 2)
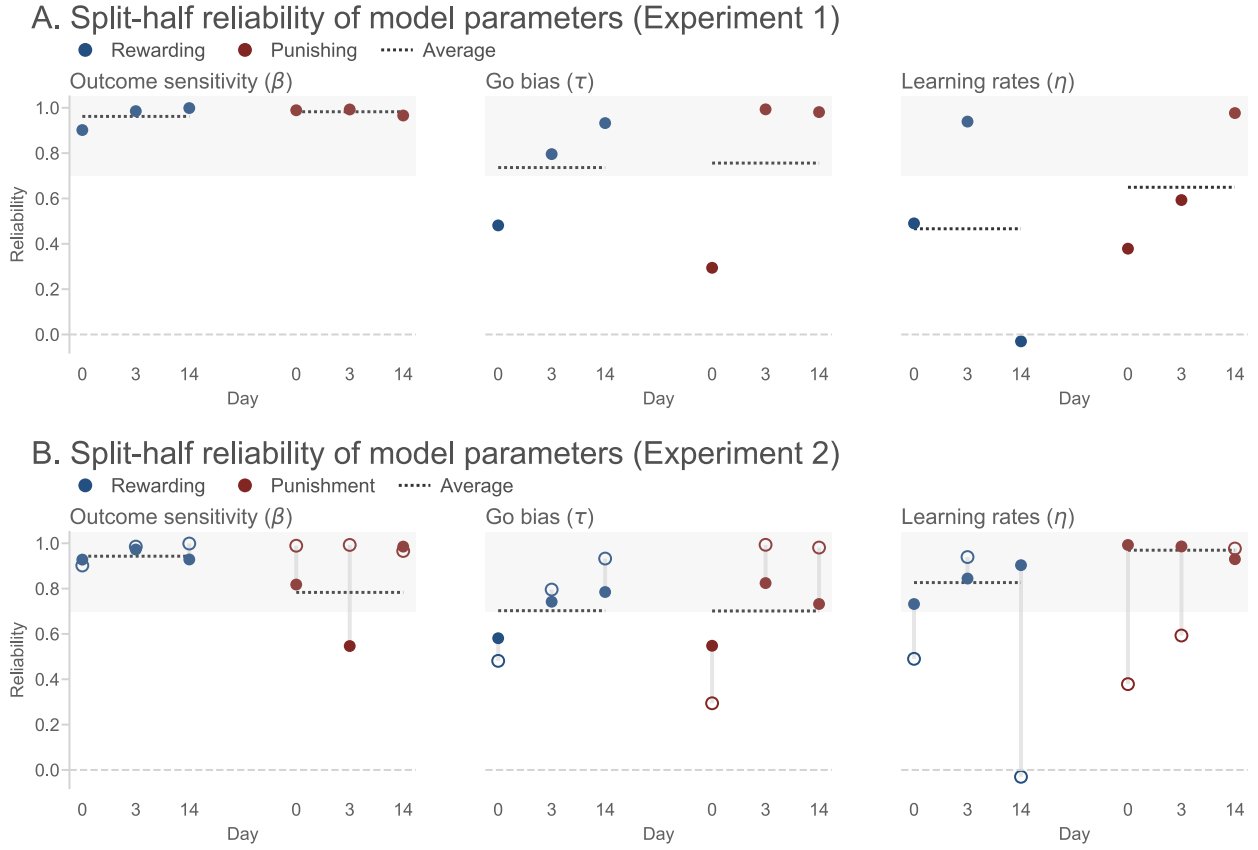
Figure S5: **Split-half reliability estimates for the best-fitting model parameters for (A) Experiment 1 and (B) Experiment 2.** Filled circles denote estimates for each experiment; open circles denote estimates from Experiment 1. Grey vertical lines show the change in reliability across experiments. Dotted lines indicate average reliability. Shaded regions indicate conventional range of acceptable reliability ($\rho \geq 0.7$).

Participant demographics

Variable	Experment 1 (N=103)	Experiment 2 (N=110)	p-value
Gender, N (%)			0.403
Men	55 (53.4%)	65 (59.1%)	
Women	47 (45.6%)	43 (39.1%)	
Transgender or nonbinary	1 (1.0%)	1 (0.9%)	
Rather not say	0 (0.0%)	1 (0.9%)	
Age, years			0.006
Mean (range)	35.5 (20–69)	39.6 (23–69)	
Race & Ethnicity, N (%)			0.264
White	80 (77.7%)	97 (83.6%)	
Black or African American	9 (8.7%)	10 (8.6%)	
Hispanic or Latino	9 (8.7%)	8 (7.3%)	
Asian	10 (9.7%)	3 (2.6%)	
American Indian/Alaska Native	0 (0.0%)	3 (2.6%)	
Rather not say	4 (3.9%)	3 (2.6%)	

Table S1: **Demographic characteristics of the participants in Experiments 1 and 2.** Participants could select more than one ethnic and racial identity. Therefore, participant counts and percentages in the corresponding section sum to more than 100%. The mean age of the two samples was compared via an independent samples t-test ($df = 211$, $\alpha = 0.05$, two-sided). The proportion of participants in each sample identifying as men or white was compared via the two sample proportions z-test ($df = 211$, $\alpha = 0.05$, two-sided). The two experiments differed significantly only in the age of participants, which was significantly older in Experiment 2.

Task appraisals

	Experiment	Day 0	Day 3	Day 14	Day 28
Difficulty	1	2.5 (1.1)	2.2 (0.9)	2.1 (1.0)	1.8 (0.8)
	2	3.3 (1.1)	3.0 (1.1)	3.1 (1.1)	-
Fun	1	3.9 (0.9)	4.0 (1.0)	3.9 (0.9)	4.1 (0.9)
	2	3.9 (0.9)	4.1 (0.8)	4.0 (0.8)	-
Clarity	1	4.8 (0.6)	4.9 (0.4)	5.0 (0.2)	4.9 (0.3)
	2	4.9 (0.5)	4.9 (0.2)	4.9 (0.2)	-

Table S2: **Mean (sd) of participants' ratings of task difficulty, fun, and clarity of the tasks.** All ratings were made on a 5-point Likert scale. Difficulty: "How difficult was the task?" (Very easy = 1, Very hard = 5); Fun: "How fun was the task?" (Very boring = 1, Very fun = 5); Clarity: "How clear were the instructions?" (Very confusing = 1, Very Clear = 5). Participants rated Experiment 2 as more difficult than Experiment 1.

Priors for Bayesian reinforcement-learning models

Parameter	Participant-level	Group-level	
		Mean	Standard deviation
Outcome sensitivity	$\beta_i \sim 10 \cdot \mathcal{N}(\mu_\beta, \sigma_\beta)$	$\mu_\beta \sim \mathcal{N}(0,1)$	$\sigma_\beta \sim \text{Half-}t(3,0,1)$
Approach/avoidance bias	$\tau_i \sim 5 \cdot \mathcal{N}(\mu_\tau, \sigma_\tau)$	$\mu_\tau \sim \mathcal{N}(0,1)$	$\sigma_\tau \sim \text{Half-}t(3,0,1)$
Learning rate	$\eta_i \sim \Phi\left(\mathcal{N}(\mu_\eta, \sigma_\eta)\right)$	$\mu_\eta \sim \mathcal{N}(0,1)$	$\sigma_\eta \sim \text{Half-}t(3,0,1)$
Lapse rate	$\xi_i \sim \Phi\left(-2 + \mathcal{N}(\mu_\xi, \sigma_\xi)\right)$	$\mu_\xi \sim \mathcal{N}(0,1)$	$\sigma_\xi \sim \text{Half-}t(3,0,1)$

Table S3: **Participant- and group-level priors specified for each parameter in the hierarchical Bayesian reinforcement-learning models.** $\mathcal{N}$ denotes the Normal distribution, Φ denotes the cumulative density function for the standard normal distribution (used to constrain learning and lapse rates to be in the range $\in [0,1]$).

Complete descriptive statistics (Experiment 1)

Variable		Within-session statistics				Between-session comparisons			
	Day	Mdn	d	p	95% CI	Day 0	Day 3	Day 14	Day 28
Correct responses (%)	0	85.0	2.982	<0.001	[80.8, 87.9]	-			
	3	92.9	6.947	<0.001	[90.4, 94.6]	<0.001	-		
	14	94.6	10.310	<0.001	[93.1, 95.2]	<0.001	0.312	-	
	28	94.6	12.029	<0.001	[93.8, 95.8]	<0.001	0.659	0.986	-
Go bias (Δ%)	0	11.7	1.049	<0.001	[10.8, 13.3]	-			
	3	5.0	1.156	<0.001	[4.2, 5.8]	<0.001	-		
	14	4.2	1.124	<0.001	[2.9, 4.6]	<0.001	0.310	-	
	28	3.3	0.899	<0.001	[2.5, 4.2]	<0.001	<0.001	0.515	-
Valence bias (Δ%)	0	0.8	0.135	0.064	[-0.8, 2.5]	-			
	3	0.8	0.169	0.017	[0.0, 2.5]	0.992	-		
	14	0.0	0.000	0.264	[-0.8, 0.8]	1.000	0.807		
	28	0.0	0.000	0.257	[-0.8, 0.8]	0.851	0.851	1.000	-
Pavlovian bias (Δ%)	0	9.2	1.237	<0.001	[6.7, 11.7]	-			
	3	1.7	0.450	<0.001	[0.8, 2.5]	<0.001	-		
	14	1.7	0.450	<0.001	[0.8, 2.5]	<0.001	0.808	-	
	28	0.8	0.225	0.001	[0.0, 1.7]	<0.001	0.270	0.998	-
Feedback sensitivity (Δ%)	0	9.4	1.250	<0.001	[7.2, 11.5]	-			
	3	4.5	0.770	<0.001	[3.4, 6.2]	<0.001	-		
	14	3.0	0.581	<0.001	[1.8, 4.7]	<0.001	0.841	-	
	28	2.8	0.704	<0.001	[2.2, 4.3]	<0.001	0.180	0.630	-

Table S4: **Within- and between-session descriptive statistics for Experiment 1.** Between-session values are p-values for the pairwise comparisons. Mdn = median.

Model comparison by session (Experiment 1)

Session	Model	Accuracy (%)	PSIS-LOO	ΔPSIS-LOO
Day 0	1	82.2	-37241.1	-2050.3 (42.9)
	2	84.1	-38327.4	-964.0 (30.7)
	3	85.2	-38992.9	-298.5 (17.7)
	4	85.2	-39145.6	-145.8 (11.1)
	5	85.3	-39129.9	-161.5 (8.9)
	6	85.4	-39190.3	-101.1 (6.7)
	7	85.6	-39291.4	-
Day 3	1	88.7	-38446.8	-1468.0 (34.0)
	2	90.0	-38979.3	-935.5 (27.5)
	3	90.8	-39492.8	-422.0 (17.7)
	4	90.8	-39644.9	-269.9 (12.9)
	5	90.8	-39629.8	-285.0 (12.3)
	6	90.9	-39666.1	-248.7 (11.6)
	7	91.0	-39914.8	-
Day 14	1	90.1	-38350.9	-928.6 (26.5)
	2	91.5	-38777.0	-502.5 (19.1)
	3	91.5	-38971.4	-308.1 (14.7)
	4	91.7	-39066.7	-212.8 (11.6)
	5	91.7	-39072.2	-207.3 (11.3)
	6	91.8	-39093.8	-185.7 (10.8)
	7	91.9	-39279.5	-
Day 28	1	89.4	-37419.1	-1155.6 (30.7)
	2	91.0	-37928.1	-646.6 (23.5)
	3	92.1	-38360.7	-214.1 (11.8)
	4	92.1	-38404.5	-170.3 (9.3)
	5	92.0	-38434.0	-140.7 (8.5)
	6	92.1	-38451.6	-123.1 (7.6)
	7	92.3	-38574.7	-

Table S5: **Model comparison broken down by session for Experiment 1.** Accuracy = trial-level choice prediction accuracy between observed and model-predicted Go responses. PSIS-LOO = approximate leave-one-out cross-validation presented in deviance scale (smaller numbers indicate better fit). ΔPSIS-LOO = difference in LOO values between each model and the best fitting model (M7).

Complete descriptive statistics (Experiment 2)

Variable	Day	Median	d	p	95% CI	Day 0	Day 3	Day 14
		Within-session statistics				**Between-session comparisons**		
Correct responses (%)	0	67.5	1.828	<0.001	[65.8, 69.0]	-		
	3	71.7	2.063	<0.001	[69.6, 74.6]	0.003	-	
	14	69.6	1.441	<0.001	[66.5, 73.8]	0.146	0.335	-
Go bias (Δ%)	0	19.2	1.108	<0.001	[15.0, 22.5]	-		
	3	13.3	0.899	<0.001	[10.0, 15.0]	0.007	-	
	14	14.2	0.882	<0.001	[7.5, 17.5]	0.003	0.382	-
Valence bias (Δ%)	0	-2.5	0.289	0.002	[-4.2, 0.8]	-		
	3	0.8	0.096	0.245	[-0.8, 3.3]	0.096	-	
	14	2.5	0.337	0.004	[0.0, 4.2]	0.002	0.851	
Pavlovian bias (Δ%)	0	12.5	1.065	<0.001	[10.0, 15.0]	-		
	3	8.3	0.843	<0.001	[6.7, 10.0]	0.027	-	
	14	7.5	0.867	<0.001	[5.0, 9.2]	0.027	0.997	-
Feedback sensitivity (Δ%)	0	28.2	2.254	<0.001	[25.7, 30.9]	-		
	3	29.4	2.142	<0.001	[24.7, 32.0]	0.632	-	
	14	26.8	2.267	<0.001	[24.1, 29.4]	0.942	0.461	-

Table S6: **Within- and between-session descriptive statistics for Experiment 2.** Between-session values are p-values for the pairwise comparisons.

Model comparison by session (Experiment 2)

Session	Model	Accuracy (%)	PSIS-LOO	ΔPSIS-LOO
Day 0	1	71.5	-33517.7	-2746.3 (48.6)
	2	75.4	-35145.3	-1118.8 (33.1)
	3	77.3	-36002.4	-261.6 (16.8)
	4	77.2	-36074.8	-189.2 (12.0)
	5	77.4	-36117.1	-147.0 (10.7)
	6	77.4	-36124.5	-139.6 (9.1)
	7	77.6	-36264.1	-
Day 3	1	74.0	-31023.5	-1656.0 (38.0)
	2	77.4	-31996.2	-683.3 (26.0)
	3	78.1	-32447.1	-232.5 (15.7)
	4	78.2	-32513.7	-165.9 (11.3)
	5	78.1	-32508.9	-170.7 (10.6)
	6	78.3	-32543.9	-135.7 (9.1)
	7	78.7	-32679.6	-
Day 14	1	73.3	-31265.0	-1802.8 (39.2)
	2	76.8	-32474.5	-593.4 (24.8)
	3	77.4	-32833.5	-234.3 (16.4)
	4	77.3	-32833.9	-233.9 (13.1)
	5	77.6	-32893.1	-174.7 (11.9)
	6	77.6	-32880.4	-187.5 (11.3)
	7	78.1	-33067.8	-

Table S7: **Model comparison broken down by session for Experiment 2.** Accuracy = trial-level choice prediction accuracy between observed and model-predicted Go responses. PSIS-LOO = approximate leave-one-out cross-validation presented in deviance scale (smaller numbers indicate better fit). ΔPSIS-LOO = difference in PSIS-LOO values between each model and the best fitting model (M7).