

Improving the Reliability of the Pavlovian Go/No-Go Task

Samuel Zorowitz^{1,*}, Gili Karni¹, Natalie Paredes², Nathaniel Daw^{1,3}, Yael Niv^{1,3}

¹Princeton Neuroscience Institute, Princeton University, USA

²Department of Psychology, University of California, San Diego, USA

³Department of Psychology, Princeton University, USA

*Corresponding author (zorowitz@princeton.edu)

Keywords: tk

Abstract

test test test

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 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 these Pavlovian response biases (i.e., initiate action to receive reward, inhibit action to avoid punishment) and slower to learn 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 innumerable studies to probe individual-differences in reward and punishment learning. Many studies have observed altered Pavlovian biases in individuals with psychiatric conditions. For example, an increased tendency towards passive avoidance has been observed in individuals with general 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 exhibit a U-shape across the lifespan, decreasing from childhood to young adulthood and increasing again in older age [11, 12]. Pavlovian biases have also been associated with individual differences in personality (e.g., impulsivity; [13]) and genetics [14, 15]. At a finer scale, Pavlovian biases is also reportedly modulated by differences in current state, such as mood [16], anger [17], stress [18], and fear [19].

Despite its extensive use, the Pavlovian go/no-go task has received considerably less psychometric investigation. What studies exist suggest the task, in its canonical form, is suboptimal for use in individual-differences research. Specifically, two independent studies found that descriptive and computational model-based measures of task performance exhibited unacceptable test-retest reliability over short (two-week) and long (6-, 18-month) retest intervals [20, 21]. 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., symptom scores; [22]). Therefore, as reliability decreases, so too does statistical power; this in turn increases the possibility of false-negatives [23] and bias in correlations that manage to reach statistical significance [24]. One of these studies also reported evidence of practice effects (i.e., improved performance and diminished Pavlovian biases with repeated testing; [20]). Practice effects are of additional concern for individual-differences research because they can imperil reliability (e.g., by minimizing between-participant variability) and obscure effects of interest (e.g., changes in task performance with changes in mood over time).

Previous studies of the psychometrics of the Pavlovian go/no-go task, however, did not make use of numerous strategies for improving the reliability of cognitive task measures [25]. For example, multiple studies have reported that gamification, or the incorporation

of (video) game design elements into cognitive tasks, can promote participant engagement [26] and thereby improve the reliability of task measures [27, 28]. Moreover, the use of hierarchical Bayesian models — which exert a pooling effect on person-level variables, in effect correcting for measurement error [29, 30] — have been multiply shown to improve the reliability of task measures [31–33]. Neither of the previous two studies made use of the strategies, raising the possibility that the psychometrics of the task is better than previously reported.

The objectives of the current study were to investigate the reliability and repeatability of a novel version of the Pavlovian go/no-go task. We hypothesized that, by gamifying the task to enhance participant engagement and by utilizing hierarchical Bayesian models, we would observe improvements in the test-retest reliability of model-based measures of task performance.

Experiment 1

Methods

Participants

A total of $N=148$ participants were recruited in May, 2020 from Amazon Mechanical Turk via CloudResearch [34] to participate in an online behavioral experiment. Participants were eligible to participate if they were at least 18 years old and resided in the United States. Following best practice recommendations [35], no other inclusion criteria were applied. This study was approved by the Institutional Review Board of Princeton University (#5291) 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” in the Supplementary Materials), leaving a final sample of $N=103$ participants. These participants were re-invited to complete a follow-up experiment 3, 14, and 28 days later. Once invited, participants were permitted 48 hours to complete the 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 were 35.5 years old on average (range: 20–69 years).

Experimental protocol

In each session, after providing consent, participants started by completing several self-report questionnaires (see Supplementary Materials for details). 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 conveyer 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 the correct response (Go, No-Go) for each robot type based on feedback (points won/lost) following their actions.

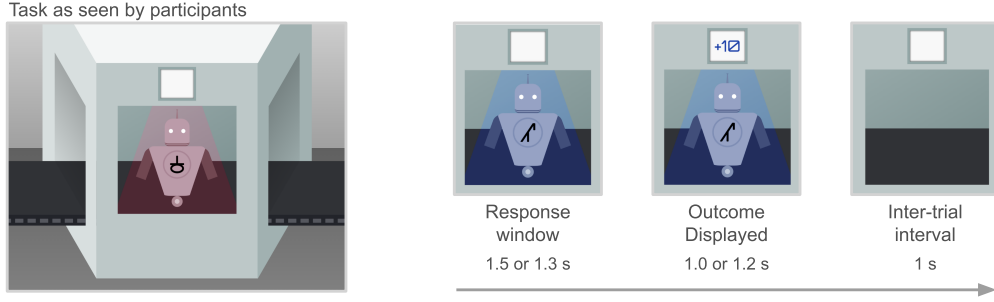
The task involved four trial types that differed in their correct action (Go, No-Go) and outcome domain (reward, punishment; Figure 1B). Specifically, the four trial types were: go to win points (GW); no-go to win points (NGW); go to avoid losing points (GAL); and no-go to avoid losing points (NGAL). 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 for each, whereas NGW and GAL trials are Pavlovian-instrumental ‘incongruent’. For the rewarding trials (GW, NGW), the possible outcomes were +10 or +1 points; for the punishing trials (GAL, NGAL), they were -1 or -10 points. The outcome domain of each robot was explicitly signaled to participants by a blue or orange ‘scanner light’ (randomized within participants across sessions). Outcomes were probabilistic such that participants received the better of the two possible outcomes with 80% chance if they made the correct action, and received the worse of the possible outcomes with 80% chance if they made the incorrect action (Figure 1C).

Participants saw eight unique robots across the task. Each individual robot was presented for 30 trials (240 trials total; Figure 1D). Trials were divided into two blocks with four robots (one of each trial type) per block. Prior to task start, participants were required to review instructions, correctly answer five comprehension questions, and complete several practice trials. Failing to correctly answer all comprehension items forced the participant to reread sections of the instructions. 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 [36] and distributed using custom web-application software (see Code Availability).

Descriptive analyses

We first evaluated participants’ choice behavior using five performance indices: overall percent correct responses; go bias, or the difference in correct responses between go and no-go trials; outcome bias, or the difference in correct responses between rewarding and punishing trials; Pavlovian bias, or the difference in correct responses between Pavlovian-instrumental congruent and incongruent trials; and feedback sensitivity, or the difference

A. Task schematic



B. Trial types

	Outcome domain	
	Gain	Loss
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	
	Gain	Loss
Correct Response	+10 80% +1 20%	-1 80% -10 20%
Incorrect Response	+1 80% +10 20%	-10 80% -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 (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 denotes outcome domain (e.g., blue denotes reward and red denotes 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 between for 8, 10, or 12 trials (240 total trials).

in correct responses between trials following veridical or sham feedback. For each session and index, we tested if the median value across participants was significantly different than zero (or 50% for overall percent correct responses). We use the median due to skew in the performance indices. We also tested if the median value of each index was significantly different between each pair of sessions. P-values were derived via permutation testing, where a null distribution of values were obtained by permuting the condition labels (for within-session tests) or session labels (for between-session tests) 5000 times. Within-session tests were not corrected for multiple comparisons as each test constitutes an individual hypothesis test; between-session tests were corrected for multiple comparisons using the familywise error rate correction [37].

Reinforcement learning models

We fit a set of seven nested reinforcement learning (RL) models to participants' choice data. All models are variants of the Rescorla-Wagner model and have previously been used to predict choice behavior on this task [2, 4, 20, 38]. In the base model (M1), the probability a participant chooses an action in response to stimulus k is determined by the difference in stimulus-action values ($Q_k(\text{Go}) - Q_k(\text{No-Go})$), scaled an outcome sensitivity parameter (β), and passed through the logistic function. In turn, stimulus-action values are learned via the Rescorla-Wagner update rule as governed by a learning rate (η). Model 2 is augmented with a single static approach/avoidance parameter (τ) that biases action towards Go responses if $\tau_j > 0$ and No-Go responses if $\tau < 0$. The succeeding model (M3) has independent approach/avoidance parameters per outcome domain (τ_+ , τ_-). Models 4 and 5 have either independent outcome sensitivity (β_+ , β_- ; M4) or learning rate parameters (η_+ , η_- ; M5) per outcome domain. In Model 6, all three parameter types are independent across outcome domains. Finally, the most complex model (M7) includes a lapse rate (ξ). Complete model equations and descriptions are provided in the Supplementary Materials.

All models were estimated within a hierarchical Bayesian modeling framework using Hamiltonian Monte Carlo as implemented in Stan (v2.30; [39]). See the Supplementary Materials for details about model estimation and our choices of prior. Briefly, we specified diffuse and uninformative priors to avoid biasing parameter estimation.

The fits of the models to the 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 Bayesian leave-one out cross-validation (LOO; [40]). A credible improvement in model fit was defined as a difference in LOO values four times larger than its corresponding standard error [41].

We investigated the reliability of the model parameters for the best-fitting model also using a Bayesian hierarchical modeling framework [30]. Briefly, the best-fitting model was fit to participants' choice data with partial-pooling over blocks within a session or over whole datasets between sessions. Split-half and test-retest reliability was respectively calculated as the Pearson correlation between model parameters estimated from each block and each pair of sessions. Complete details of the estimation procedure are reported in the Supplementary Materials.

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

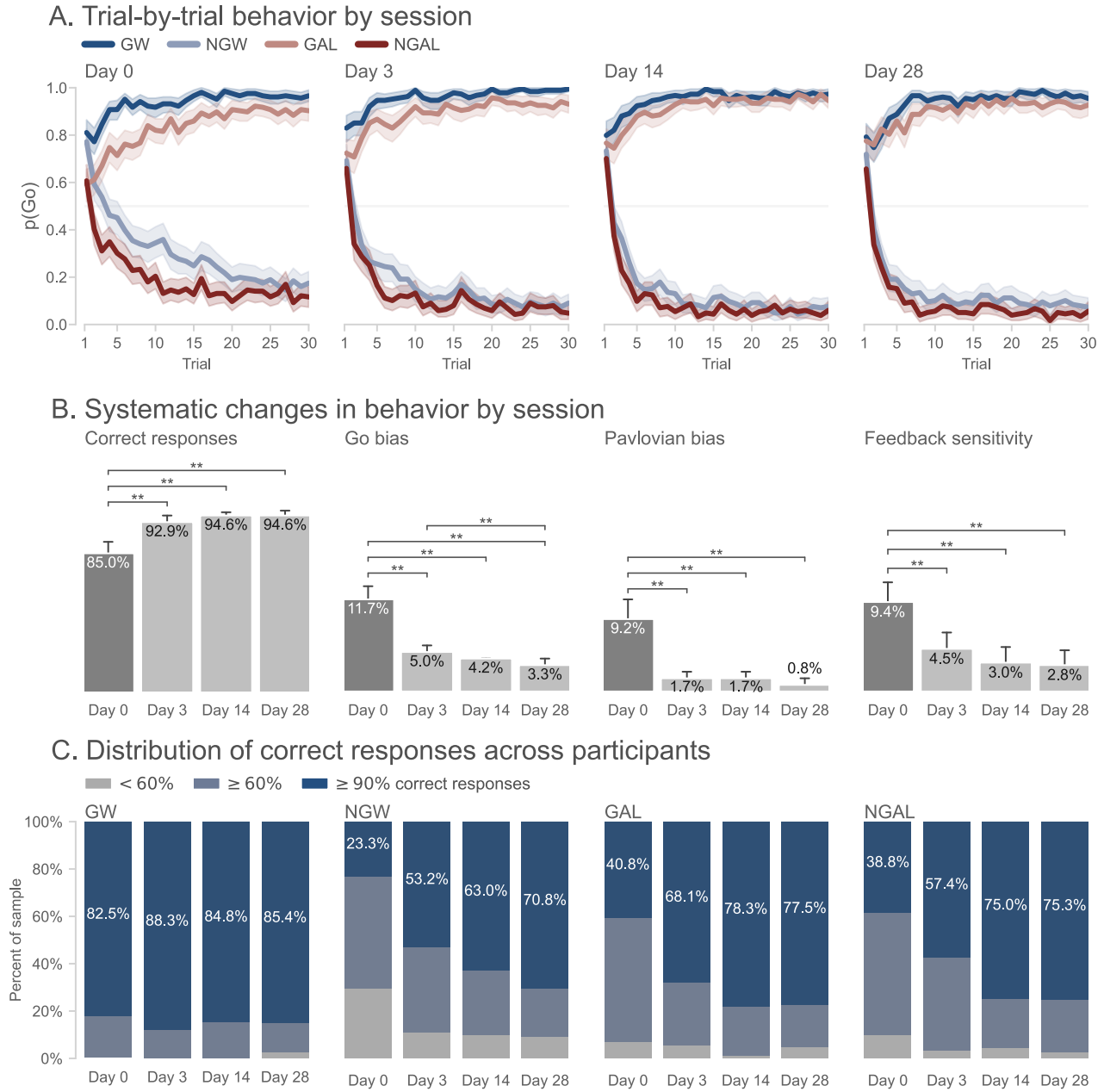


Figure 2: Summary of behavior in Experiment 1. (A) Group-averaged learning curves for each trial type and session. Shaded regions indicate 95% bootstrapped confidence intervals. (B) Group-averaged performance for each session. Performance indices from left-to-right: Correct responses, or overall accuracy; 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. Blue shading indicates significant differences in comparison to all other sessions. (C) The percentage of participants, for each session and trial type, exhibiting at- or below-chance performance (< 60% response accuracy; grey), moderate performance (≥ 60% response accuracy; light blue), or near-perfect performance (≥ 90% response accuracy; dark blue).

performance on Pavlovian-instrumental incongruent trials. However, this effect was ostensibly diminished in all follow-up sessions. We turn to the descriptive analyses to formally test these trends.

The group-averaged performance indices per session are summarized in Figure 2B. On Day 0, the overall percent correct responses was 85.0%. Performance improved 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 no-go trials. The go bias was significantly reduced in all follow-up sessions compared to Day 0 (all $p < 0.001$); so too was the bias on Day 28 compared to Day 3 ($p < 0.001$). Similarly, participants made more correct responses on Pavlovian-congruent than incongruent trials. Again, the Pavlovian bias was significantly reduced in all follow-up sessions compared to Day 0 (all $p < 0.001$); no other comparisons were significant. Participants made marginally more correct responses on rewarding trials on Day 3 only.

A similar pattern of results was observed for feedback sensitivity. Across sessions, participants made more correct responses following veridical compared to sham feedback (Day 0: $\Delta M = 0.094$, $d = 1.250$, $p < 0.001$, 95% CI = [0.072, 0.115]; Day 3: $\Delta M = 0.045$, $d = 0.770$, $p < 0.001$, 95% CI = [0.034, 0.062]; Day 14: $\Delta M = 0.030$, $d = 0.581$, $p < 0.001$, 95% CI = [0.018, 0.047]; Day 28: $\Delta M = 0.028$, $d = 0.704$, $p < 0.001$, 95% CI = [0.022, 0.043]). Feedback sensitivity was significantly less in all follow-up sessions compared to Day 0 (all $p < 0.001$); no other comparisons were significant.

The results so far summarize group-averaged performance but do not provide much insight into individual variability. As such, the proportion of participants who exhibited chance-level (<60% correct responses), intermediate ($\geq 60\%$), or near-ceiling performance ($\geq 90\%$) by session and trial type is presented in Figure 2C. Excepting GW trials, the percentage of participants nearing ceiling-level performance increases from the minority on Day 0 to the majority in all follow-up sessions. Two-way chi-squared tests confirm 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 increases in task performance (and accompanying reductions in choice biases) with repeated testing observed at the group-level extends 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., independent parameters per outcome domain plus 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).

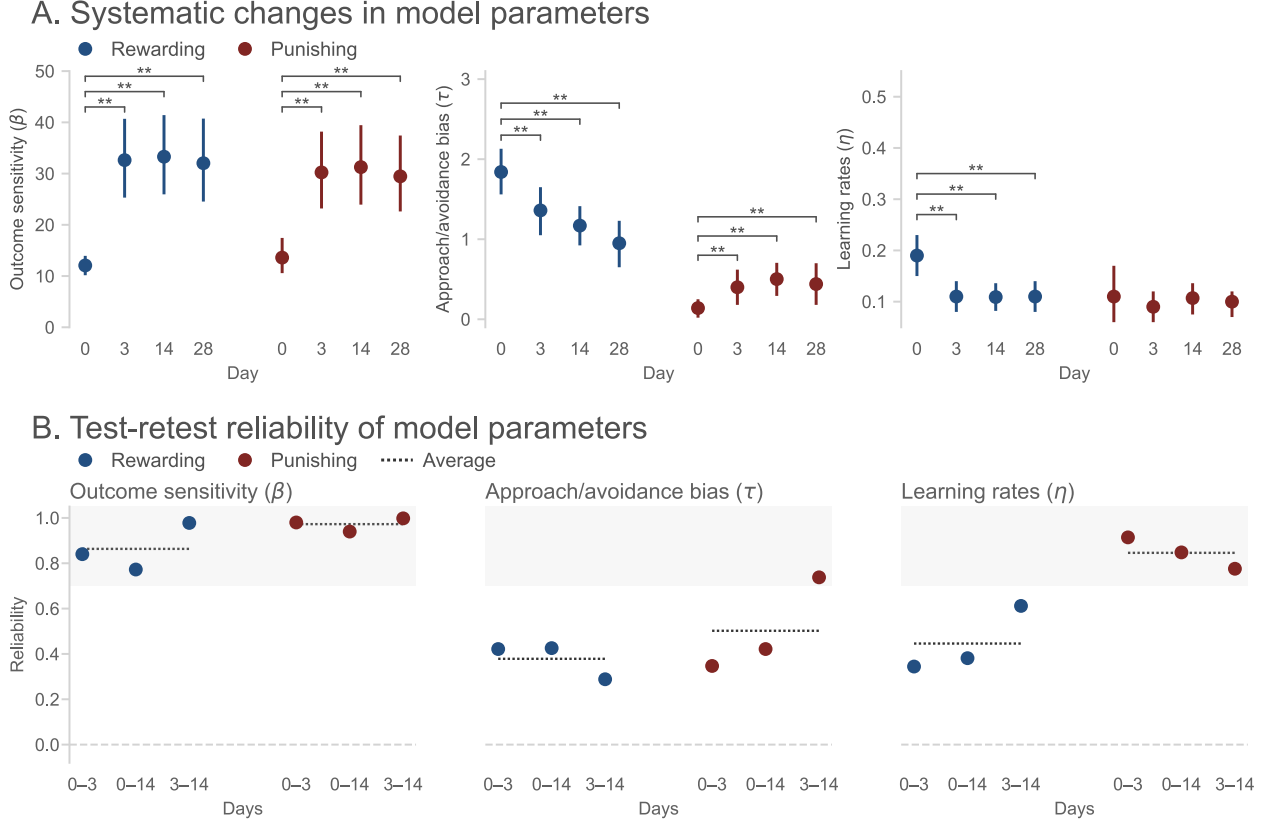
Model parameters

The estimated group-level parameters from the best-fitting model are presented in Figure 3A. Consistent with the descriptive analyses, large shifts in the parameters were observed following Day 0. The reward and punishment sensitivity parameters exhibited an almost threefold increase between Days 0 and 3, and stabilized thereafter. The inverse pattern was observed for the positive learning rate. Crucially, the approach and avoidance bias parameters followed a similar pattern. The approach bias credibly decreased between Days 0 and 3, and qualitatively declined thereafter. In turn, the avoidance bias credibly increased between Days 0 and 3, but stabilized thereafter. That is, Pavlovian biases diminished in absolute and relative (i.e., compared to the outcome sensitivity parameters) terms with repeat testing.

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

Model	Parameters	Acc. (%)	LOO	Δ LOO (se)
1	β, η	87.5	-151457.9	-5602.6 (68.3)
2	β, τ, η	89.0	-154011.9	-3048.6 (51.2)
3	$\beta, \tau_+, \tau_-, \eta$	89.8	-155817.8	-1242.7 (31.3)
4	$\beta_+, \beta_-, \tau_+, \tau_-, \eta$	89.8	-156261.6	-798.8 (22.6)
5	$\beta, \tau_+, \tau_-, \eta_+, \eta_-$	89.9	-156265.9	-794.6 (20.7)
6	$\beta_+, \beta_-, \tau_+, \tau_-, \eta_+, \eta_-$	89.9	-156401.8	-658.6 (18.8)
7	$\beta_+, \beta_-, \tau_+, \tau_-, \eta_+, \eta_-, \xi$	90.1	-157060.5	-

Table 1: Model comparison collapsing across sessions. Notes: Acc. = trial-level predictive accuracy between observed and model-predicted choice. LOO = Bayesian leave-one-out cross-validation scores presented in deviance scale (i.e., smaller numbers indicate better fit). Δ LOO = difference in LOO values between each model and the best-fitting model (M7).



*Figure 3: Summary of reinforcement learning model parameters. (A) Group-level model parameters for each session. Error bars indicate 95% HDIs. **Denotes pairwise comparison where 95% HDI of the difference excludes zero. (B) Test-retest reliability estimates for each model parameter. Dotted lines indicates overall average. Shaded region indicates conventional range of acceptable reliability ($\rho \geq 0.7$).*

Discussion

The objectives of this first experiment was to evaluate the stability and reliability of individual differences in performance on a gamified version of the Pavlovian go/no-go task. At both the group- and participant-levels, we observed significant practice effects on performance following the initial 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 halved or more after the first session. Moreover, this pattern of results was reflected in the group-level parameters of an RL model fit to participants' choice data, which indicated that Pavlovian biases were credibly weakened in follow-up sessions. Relatedly, we found that the Pavlovian bias parameters exhibited poor-to-moderate test-retest reliability. This last result is perhaps unsurprising given that, by definition, low between-participants variability decreases reliability.

The results of Experiment 1 beg two questions: what engenders these practice effects and what can be done to prevent or mitigate them? With respect to the first question, one possibility is that, after the initial session, participants employ an alternative cognitive

strategy to complete the task. Consider that, under the canonical design of the Pavlovian go/no-go task, there is an implicit dependence between stimuli: within an outcome domain, for every go stimulus (e.g., GW) there is always a corresponding no-go stimulus (e.g., NGW). That is, learning the correct action for one stimulus provides information about the correct action for its complement. Recognizing this, savvy participants may simply forego reinforcement learning in favor of a process-of-elimination strategy and thereby enhance their performance on the task. Indeed, feedback from some of our online participants in this study indicated that this sort of top-down strategy was at play.

This suggests that, in order to mitigate or prevent practice effects, there is a need for an alternative version of the Pavlovian go/no-go task without such an easily exploited structure. By eliminating this dependence between stimuli, motivated participants would have no other other option than to pay attention to each and every stimulus equally in order to maximize their performance. Furthermore, by minimizing practice effects (and thereby 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 from Amazon Mechanical Turk (via CloudResearch) in December, 2020 to participate in an online behavioral experiment. The inclusion criteria were the same as in Experiment 1. This study was approved by the Institutional Review Board of Princeton University (#11968), and all participants provided informed consent. Total study duration was again 15-20 minutes per participant. 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” in the Supplementary Materials), leaving a final sample of $N=110$ participants. These participants were re-invited to complete a follow-up experiment 3 and 14 days later. (There was no follow-up session at 28 days due to 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, on average, 39.6 years old (range: 23–69 years).

Experimental protocol

The experimental protocol for Experiment 2 was almost identical to Experiment 1. In each session, participants started by completing several self-report questionnaires (see Supplementary Materials for details). Next, participants completed an alternative version of the gamified Pavlovian go/no-go task where the trial structure had been fundamentally altered (taking inspiration from [42]). Instead of 8 unique robots each presented for 30 trials, participants now saw a total of 24 unique robots presented for 8, 10, or 12 trials each. The motivation for shortening the number of exposures to each stimulus was to measure participants' task performance primarily during learning, as opposed to asymptotic performance, where the expression of Pavlovian biases are typically largest (for similar discussion, see also [43]). The robots were presented to participants in batches, each composed of four robots totalling approximately 40 trials. Crucially, within a batch, all four trial types were not necessarily represented (see Figure S2 for an example). That is, at any point in the task, participants were not guaranteed to observe one of each type of robot. As such, participants could not utilize a process-of-elimination strategy in order to improve their task performance. The six batches were evenly divided into two blocks of 120 trials each (12 unique robots per block; three robots of each type; Figure 1D).

The task was unchanged visually except in two respects. First, the scanner colors were now blue and red (instead of blue and orange), and fixed such that they always indicated rewarding and punishing trials, respectively. Second, the symbols on the robots' chestplates were drawn from one of two Brussels Artificial Character Sets [44] 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).

Analyses

The analyses for Experiment 2 were identical to those for Experiment 1.

Results

Descriptive analyses

Trial-by-trial choice behavior for each session is presented in Figure 4A. In contrast to Experiment 1, performance in each session qualitatively conformed to the expected pattern of results. The group-averaged performance indices per session are summarized in Figure 4B. On Day 0, overall percent correct responses was 67.5%. Performance improved in all subsequent sessions, but only marginally so. Pairwise comparisons confirmed that

performance was significantly greater on Day 3 compared to Day 0 ($p = 0.009$); however, no other comparisons were significant.

In all sessions, participants performed better on go trials than no-go trials. The go bias on Day 0 was significantly greater than that for all other sessions (all $p < 0.005$); no other comparisons were significant. Participants also performed better on Pavlovian congruent compared to incongruent trials in all sessions. The Pavlovian bias on Day 0 was significantly greater than that for all other sessions (both $p = 0.027$); no other comparisons were significant. In sum, group-averaged behavior showed evidence of practice effects but these were minor by comparison to those observed in Experiment 1.

A similar pattern of results was observed for feedback sensitivity. Across sessions, participants made more correct responses following veridical compared to sham feedback. No pairwise comparison was significant (all $p > 0.10$), suggesting that feedback sensitivity was largely conserved across sessions.

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, the majority of participants now exhibited intermediate performance across all trial types and sessions with the only exception being for NGW trials on Day 0. Two-way chi-squared tests of independence confirmed that, with the exception of 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 exhibit and sustain performance in the dynamic range on the modified Pavlovian go/no-go task.

Model comparison

The results of the model comparison are summarized in Table 2. It should be noted that trial-level predictive performance for all models were worse in Experiment 2 than in Experiment 1, which is to be expected insofar that the modified task samples more choice during learning (i.e., when choice behavior is most stochastic). As in Experiment 1, collapsing across sessions, the best-fitting model was the most complex one (i.e., independent parameters per outcome domain plus lapse rate; M7). 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. Small but credible shifts were observed for the reward and punishment sensitivity parameters. Reward sensitivity was credibly larger on Day 14 compared to Days 0 and 3, whereas punishment sensitivity was credibly larger on Days 3 and 14 compared to Day

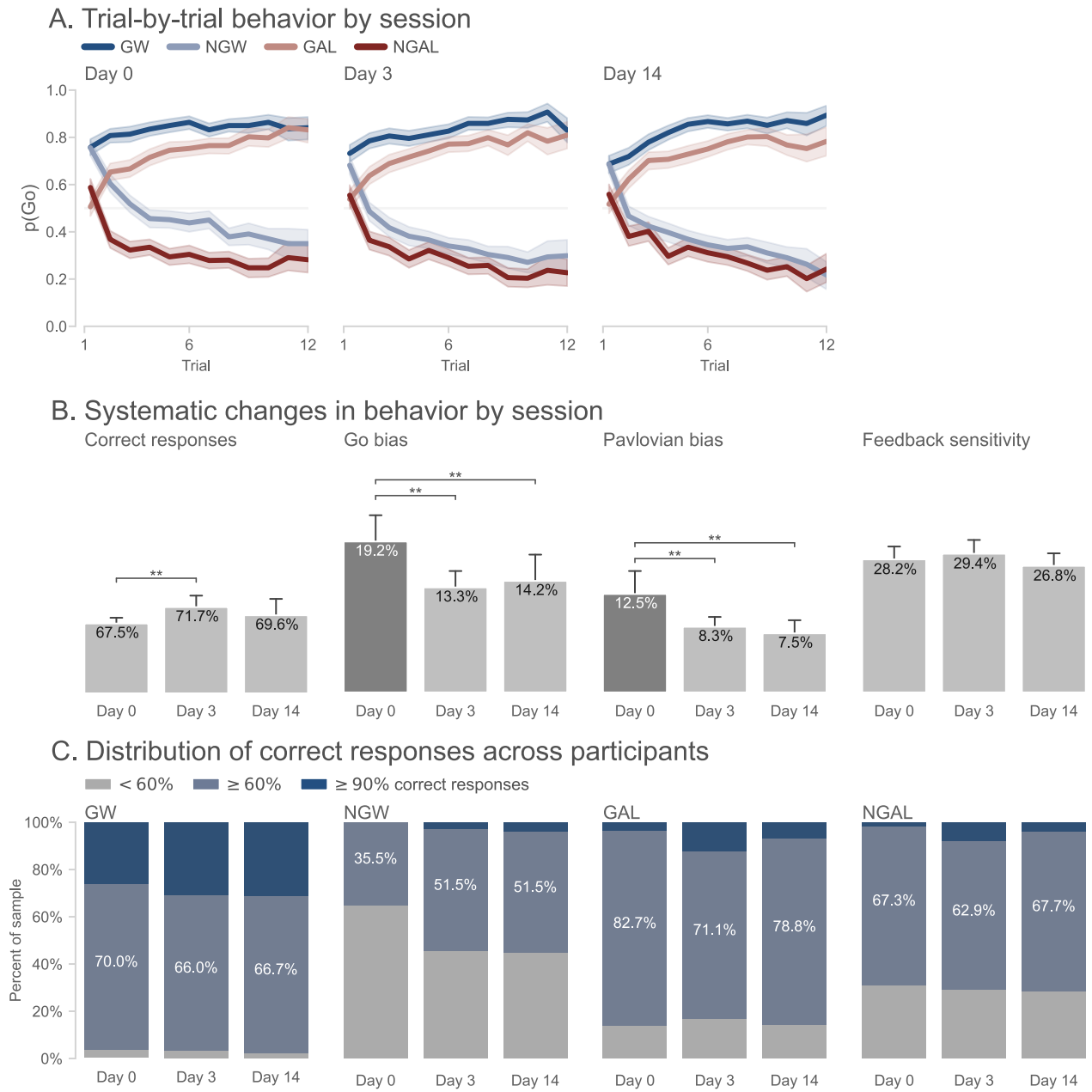


Figure 4: Summary of behavior in Experiment 2. (A) Group-averaged learning curves for each trial type and session. Shaded regions indicate 95% bootstrapped confidence intervals. (B) Group-averaged performance for each session. Performance indices from left-to-right: Correct responses, or overall accuracy; 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. Blue shading indicates significant differences in comparison to all other sessions. (C) The percentage of participants, for each session and trial type, exhibiting at- or below-chance performance (< 60% response accuracy; grey), moderate performance (≥ 60% response accuracy; light blue), or near-perfect performance (≥ 90% response accuracy; dark blue).

0. The inverse pattern was observed for the reward and punishment learning rates. The approach bias was marginally but credibly larger on Day 0 compared to Days 3 and 14. No credible differences in the avoidance bias across sessions were observed. Therefore, though Pavlovian biases were somewhat diminished, in both absolute and relative (i.e., compared to the outcome sensitivity parameters) terms, they remain largely intact with repeat testing.

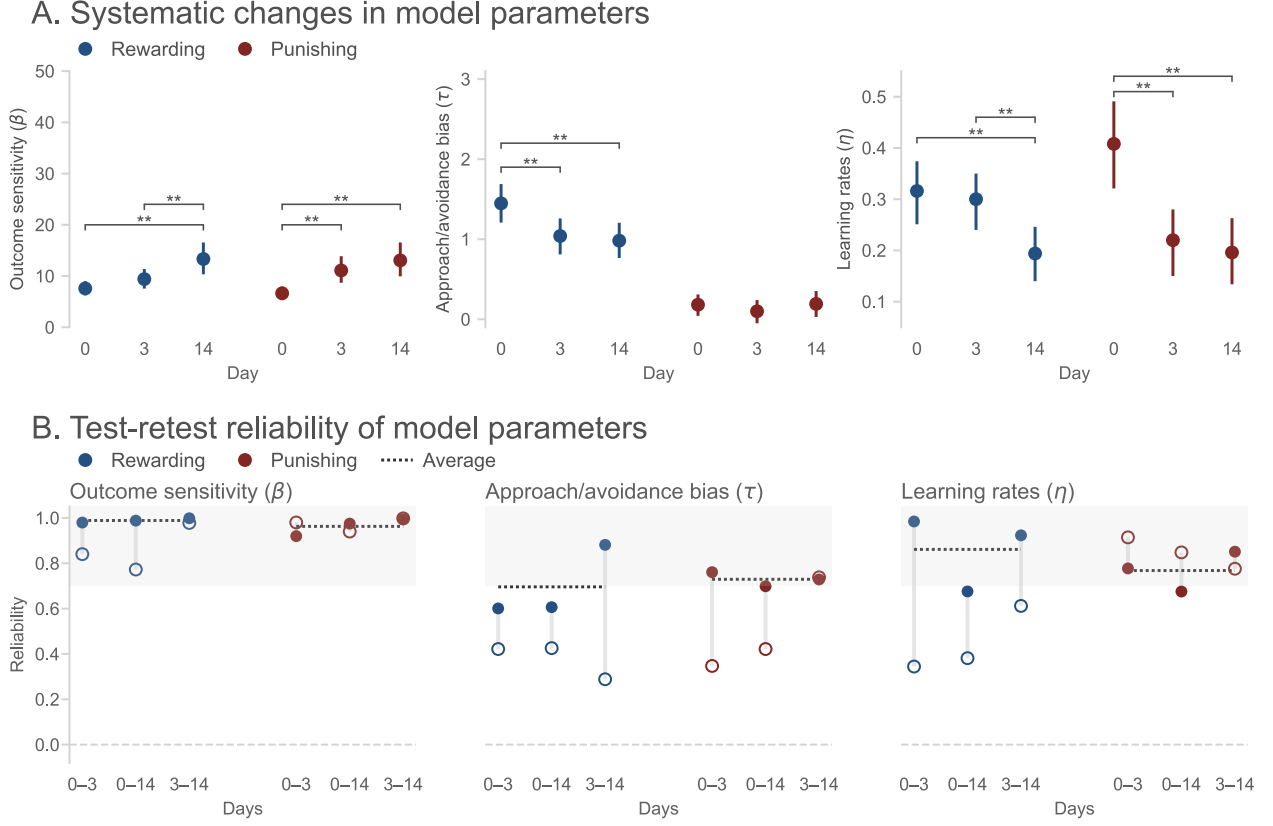
The estimated test-retest reliability of the model parameters is presented in Figure 5B. In contrast to Experiment 1, collapsing across session-pairs, acceptable test-retest reliability was observed for essentially all parameters (reward sensitivity: $\bar{\rho} = 0.989$, 95% CI = [0.984, 0.992]; punishment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

General Discussion

The test-retest reliability of reinforcement learning model parameters observed here are markedly improved compared to previous studies [20, 21]. Indeed, previous studies investigating the reliability of both descriptive and model-based behavioral measures of performance on the Pavlovian go/no-task found them to be moderate-to-poor in range. The discrepancy in findings between ours and past studies likely reflect a confluence of factors. First, our task made use of gamification principles intended to promote partic-

Model	Parameters	Acc. (%)	LOO	Δ LOO (se)
1	β, η	72.9	-95806.3	-6205.2 (73.2)
2	β, τ, η	76.5	-99616.0	-2395.5 (48.9)
3	$\beta, \tau_+, \tau_-, \eta$	77.6	-101283.0	-728.5 (28.2)
4	$\beta_+, \beta_-, \tau_+, \tau_-, \eta$	77.5	-101422.4	-589.0 (21.1)
5	$\beta, \tau_+, \tau_-, \eta_+, \eta_-$	77.7	-101519.0	-492.4 (19.1)
6	$\beta_+, \beta_-, \tau_+, \tau_-, \eta_+, \eta_-$	77.8	-101548.7	-462.7 (17.2)
7	$\beta_+, \beta_-, \tau_+, \tau_-, \eta_+, \eta_-, \xi$	78.1	-102011.4	-

Table 2: Model comparison collapsing across sessions. Notes: Acc. = trial-level classification accuracy between observed and model-predicted go responses. LOO = Bayesian leave-one-out cross-validation presented in deviance scale (i.e., smaller numbers indicate better fit). Δ LOO = difference in LOO values between models.



*Figure 5: Summary of reinforcement learning model parameters. (A) Group-level model parameters for each session. Error bars indicate 95% HDIs. **Denotes pairwise comparison where 95% HDI 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. The 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 ($\rho \geq 0.7$).*

ipant engagement and minimize confusion, which has previously been found to improve the reliability of cognitive task measures [27, 28]. A second possibility is that, similar to our Experiment 1, previous studies were hampered by practice effects. Practice effects can substantially impact reliability when they are not uniformly expressed by participants, such as when they manifest as a function of age [45]. Indeed we observed the largest improvements in parameter reliability after minimizing practice effects in Experiment 2. A third possibility is our use of method. By exploiting partial-pooling, hierarchical Bayesian models have been shown to minimize parameter estimation noise and thereby improve reliability estimates [29, 30]. Previous empirical studies have demonstrated the impact of hierarchical Bayesian models for improving reliability estimates [32, 33].

Relatedly, we also observed large practice effects for the standard version of the task (Experiment 1). This is not in itself surprising, as practice effects are common to cognitive tasks in general [46, 47] and were previously observed for the Pavlovian go/no-go task (even across lengthy retest intervals; [20]). In our case, participants appeared to be using a qualitatively different behavioral strategy to complete the task after the initial administration. Specifically, participants ostensibly exploited acquired knowledge of the

task structure to utilize a “process-of-elimination” strategy that allowed them to identify the correct response to certain quickly and without the need for trial-and-error (i.e., reinforcement) learning. Our solution was to redesign the task in such a way to prevent such strategies from forming or being advantageous to use. This approach is similar to previous studies, where preventing participants from becoming aware of critical elements of the task design improved both the consistency and reliability of behavior even with practice [48].

The current study is not without limitations. Here we have only investigated the psychometric properties of our modified Pavlovian go/no-go task in a sample of online adult participants. The reliability of task measures, however, are not absolute; reliability can vary by participant demographics and the experiment setting. For example, previous studies have found that the reliability of specific task measures in healthy adults is not the same as for adults with psychopathology [49] or healthy children [50]. To the extent that our modified version of the Pavlovian go/no-go task is more challenging (see S2), it may be the case that it proves too challenging for other groups, which may harm reliability (e.g., [50]). Additional studies are necessary to validate this version of the task in other populations. It should also be noted that the current study is not necessarily informative about the reliability of Pavlovian biases over longer periods of time (i.e., beyond 2–4 weeks).

A second limitation is that we studied only choice but not response times. Previous studies have found that Pavlovian biases manifest in response times in addition to choice behavior [51, 52], and may be a meaningful index of individual-differences in behavior [6, 12, 53]. Furthermore, previous studies have introduced computational frameworks for jointly modeling participants’ choice and response time behavior on the task [6, 51]. This is notable because the joint modeling of choice and response time has previously been found to improve the precision and reliability of estimated parameters in reinforcement learning models [54, 55]. Therefore, future research should investigate how the reliability of model-based 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 behavioral 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, we have made all of our data and code publicly available (see Code Availability).

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/jpspsych-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 [56, 57]. 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 [56]. By this measure and excluding self-citations to the first and last authors of our current paper), our references contain 12.0% woman(first)/woman(last), 16.0% man/woman, 26.0% woman/man, and 46.0% 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 [58, 59]. By this measure (and excluding self-citations), our references contain 4.7% author of color (first)/author of color(last), 14.1% white author/author of color, 18.2% author of color/white author, and 63.0% 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

The authors are grateful to Daniel Bennett for helpful feedback on the task design. The research reported in this manuscript was supported in part by the National Center for Advancing Translational Sciences (NCATS), a component of the National Institute of Health (NIH), under award number UL1TR003017 (ND, YN). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. SZ was supported by an NSF Graduate Research Fellowship. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

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. Raab, H. A. & Hartley, C. A. Adolescents exhibit reduced Pavlovian biases on instrumental learning. *Scientific reports* **10**, 1–11 (2020).
12. Betts, M. J. *et al.* Learning in anticipation of reward and punishment: perspectives across the human lifespan. *Neurobiology of aging* **96**, 49–57 (2020).
13. 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).
14. 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).
15. 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).

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. Spearman, C. “General intelligence”, objectively determined and measured. *Am. J. Psychol.* **15**, 201–293 (1904).
23. 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).
24. Gelman, A. & Carlin, J. Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science* **9**, 641–651 (2014).
25. Zorowitz, S. & Niv, Y. Improving the reliability of cognitive task measures: A narrative review. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* (2023).
26. 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).
27. Kucina, T. *et al.* A solution to the reliability paradox for decision-conflict tasks (2022).
28. 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).
29. 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).
30. Rouder, J. N. & Haaf, J. M. A psychometrics of individual differences in experimental tasks. *Psychonomic bulletin & review* **26**, 452–467 (2019).
31. 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).

32. 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).
33. 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).
34. 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).
35. 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).
36. De Leeuw, J. R. jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior research methods* **47**, 1–12 (2015).
37. 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).
38. Swart, J. C. *et al.* Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in) action. *Elife* **6**, e22169 (2017).
39. Carpenter, B. *et al.* Stan: A probabilistic programming language. *Journal of statistical software* **76**, 1–32 (2017).
40. 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).
41. Vehtari, A. *Cross-validation FAQ* en. <https://mc-stan.org/loo/articles/online-only/faq.html>. Accessed: 2023-2-2.
42. Wittmann, B. C., Daw, N. D., Seymour, B. & Dolan, R. J. Striatal activity underlies novelty-based choice in humans. *Neuron* **58**, 967–973 (2008).
43. Zorowitz, S., Quoshe, L., Karni, G. & Niv, Y. risk paper. *PsyArXiv* (2023).
44. 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).
45. Anokhin, A. P. *et al.* Age-related changes and longitudinal stability of individual differences in ABCD Neurocognition measures. *Developmental Cognitive Neuroscience* **54**, 101078 (2022).
46. 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).
47. Scharfen, J., Peters, J. M. & Holling, H. Retest effects in cognitive ability tests: A meta-analysis. *Intelligence* **67**, 44–66 (2018).

48. 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).
49. 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).
50. 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).
51. 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).
52. 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).
53. 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).
54. 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).
55. 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).
56. Dworkin, J. D. *et al.* The extent and drivers of gender imbalance in neuroscience reference lists. *Nature neuroscience* **23**, 918–926 (2020).
57. Bertolero, M. A. *et al.* Racial and ethnic imbalance in neuroscience reference lists and intersections with gender. *BioRxiv* (2020).
58. 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.
59. Sood, G. & Laohaprapanon, S. Predicting race and ethnicity from the sequence of characters in a name. *arXiv preprint arXiv:1805.02109* (2018).

Supplementary materials

Self-report measures

In Experiment 1, prior to starting the Pavlovian go/no-go task, participants completed the following self-report questionnaires: the 7-item generalized anxiety disorder scale (GAD-7; [1]); the 14-item manic and depressive tendencies scale (7-up/7-down; [2]); and the abbreviated 12-item behavioral activation/inhibition scale (BIS/BAS; [3]). Participants also indicated their current mood using the affective slider [4]. Participants completed the 7-up/7-down and BIS/BAS scales twice, once on Day 0 and Day 28. Participants completed the GAD-7 and mood slider scales at the start of every session.

In Experiment 2, prior to starting the task, participants completed the same set of 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; [5]). Participants completed the BIS/BAS scale once, on Day 0 only. Participants completed the GAD-7, DASS, and mood slider scales at the start of every session.

Exclusion criteria

To ensure data quality in Experiments 1 and 2, the data from multiple participants from the initial session were excluded prior to analysis. In Experiment 1, the data from multiple participants who completed the experiment on Day 0 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) or exhibiting chance-level performance (<55% correct responses) on go to win trials (N=43). Data from N=45 participants who completed the first session were excluded under these criteria, leaving a final sample of N=103 participants. No exclusions were applied to subsequent session data.

In Experiment 2, the data from multiple participants who completed the experiment on Day 0 were excluded prior to analysis for one or both of the following reasons: failing one or more attention checks embedded in the self-report measures (N=30); making go or no-go responses on more than 90% of trials (N=5); exhibiting chance-level performance (<55% correct responses) across all trials (N=22). Data from N=46 participants who completed the first session were excluded under these criteria, leaving a final sample of N=110 participants. No exclusions were applied to subsequent session data.

Reinforcement learning models

To more precisely characterize participants' performance on the Pavlovian go/no-go task in Experiments 1 and 2, we fit a nested set of reinforcement learning models to their choice data. All models are variants of the Rescorla-Wagner model and have previously been

used to predict choice behavior on this task [6–9]. Under the most complex model (M7), the probability that a participant makes a go response following stimulus k is defined as:

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

where $Q_k(\text{go})$ and $Q_k(\text{no-go})$ are the learned stimulus-action values for the go and no-go responses, respectively; β_{v_k} is the reward sensitivity (if stimulus k is rewarding) or punishment sensitivity (if stimulus k is punishing) parameter; τ_{v_k} is an approach bias (if stimulus k is rewarding) or avoidance bias (if stimulus k is punishing) parameter; and ξ is the lapse rate. In turn, the value of a response is learned through feedback according to a learning rule:

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

where r is the observed outcome on a particular trial; and η_{v_k} is the positive learning rate (if stimulus k is rewarding) or the negative learning rate (if stimulus k is punishing). In all models, outcomes were rescaled so that the the better and worse of the two possible outcomes were equal to one and zero, respectively. The Q-values were accordingly initialized to $Q = 0.5$.

Simplifications of this model involved either fixing parameters to be equal to zero or fixing parameters to be equal their counterpart. Specifically, the base model (M1) had two free parameters: a single outcome sensitivity parameter and a single learning rate, both shared across outcome domains (i.e., $\beta_+ = \beta_-$; $\eta_+ = \eta_-$; $\tau_+ = \tau_- = 0$). Model 2 added a static bias parameter that was shared across outcome domains (i.e., $\tau_+ = \tau_-$). Model 3 relaxed this assumption and involved for independent approach (τ_+) and avoidance (τ_-) parameters. The next two models involved independent outcome sensitivity (β_+, β_- ; M4) or learning rate (η_+, η_- ; M5) parameters by outcome domain. Model 6 involved both independent outcome sensitivity and learning rate parameters. Finally, the most complex model (M7) included a lapse rate (ξ).

All models were estimated within a hierarchical Bayesian modeling framework using Hamiltonian Monte Carlo as implemented in Stan (v2.30; [10]). 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. Thus, 5,000 post-warmup samples from the joint posterior were retained. 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).

Reliability analyses

A primary objective of Study 2 was to measure the split-half reliability (within session 1) and test-retest reliability (between sessions 1 & 2) of individual differences across model parameters (especially the asymmetry index, κ). To do so, we used a nested hierarchical modeling approach where parameters were pooled both within- and across-participants (see [11]). Briefly, for a particular parameter type, separate parameters were estimated

per participant and session (in the case of test-retest reliability) or task-half (in the case of split-half reliability). Specifically, the parameters were estimated as follows:

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

where θ_{i1} and θ_{i2} are some parameter (e.g., inverse temperature) for participant i in sessions (or task-halves) 1 and 2, respectively; μ_1 and μ_2 are the group-averaged parameters for sessions (or task-halves) 1 and 2; θ_{ic} is the common effect for participant i (i.e., the parameter component that is stable across sessions or task-halves); and θ_{id} is the difference effect for participant i (i.e., the parameter component that varies across sessions or task-halves). The collection of θ_{ic} parameters across participants represents between-participants variability, whereas the collection of θ_{id} parameters represent 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-halves and sessions, respectively.

Comparative stability of self-report measures

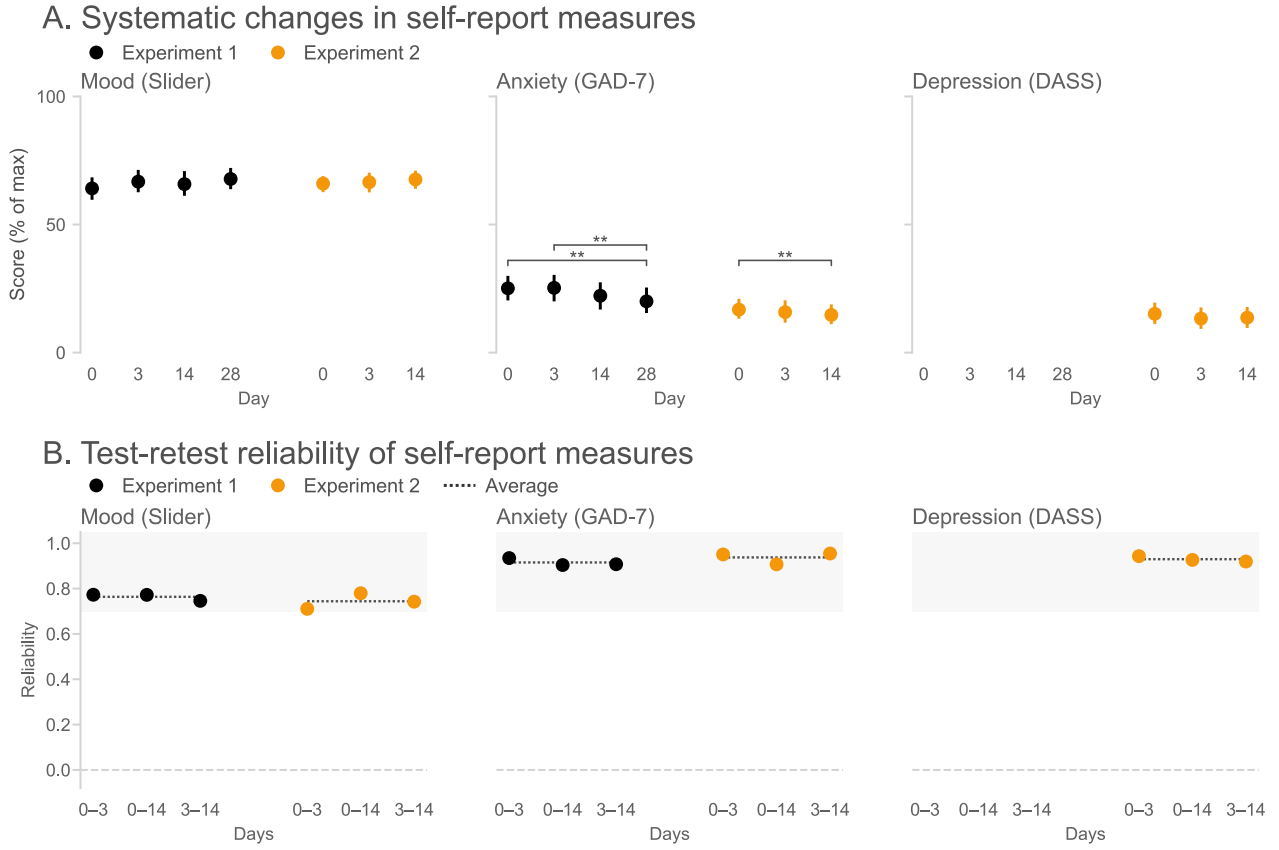


Figure S1: Summary of self-report measures indicating that mood and anxiety across participants were relatively stable over the study period. (A) Group-level mood and symptom scores for each session. Error bars indicate 95% HDIs. **Denotes significant pairwise comparison. (B) Test-retest reliability estimates for each model parameter. Dotted lines indicates overall average. Shaded region indicates conventional range of acceptable reliability ($\rho \geq 0.7$).

Example block of the modified Pavlovian go/no-go task

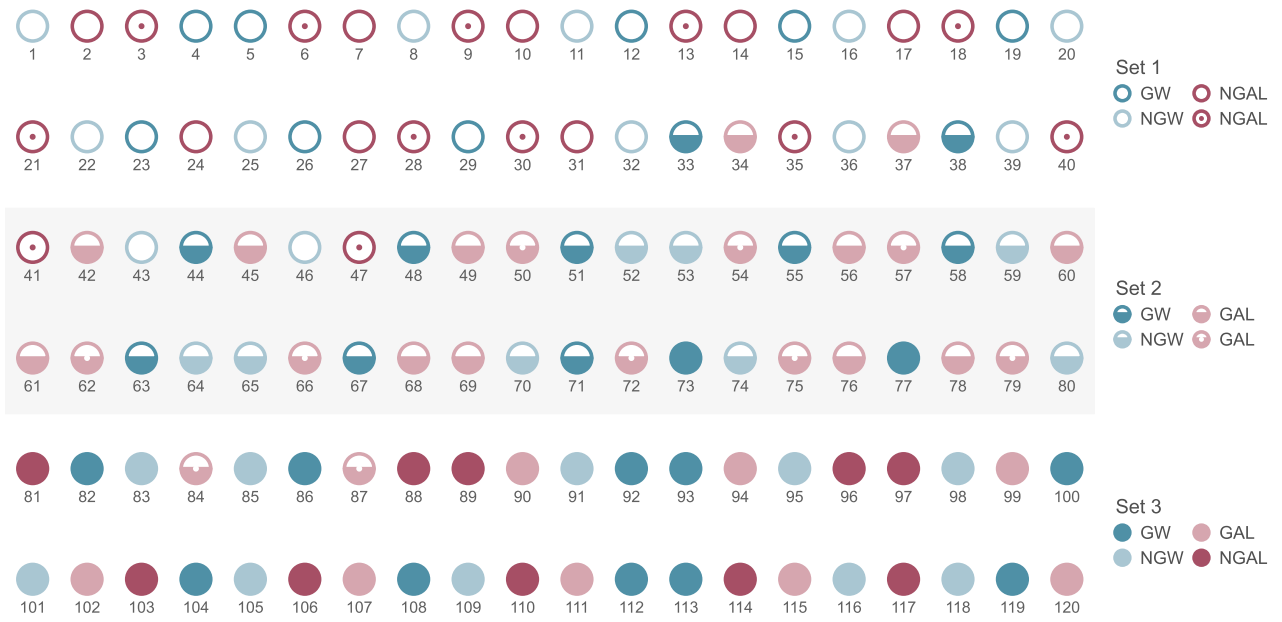


Figure S2: The trial structure of an example block of the modified Pavlovian go/no-go task. In one task block, there are 12 unique stimuli (three of each trial type: go to win [GW; dark blue]; no-go to win [NGW; light blue]; go to avoid losing [GAL; light red]; no-go to avoid losing [NGAL; dark red]) divided into three sets (Set 1: empty circles; Set 2: half-filled circles; Set 3: filled circles). Each set is composed of approximately 40 trials (120 trials total). Each set, however, may not involve all four trial types. In the example block above, Set 1 involves two NGAL stimuli and zero GAL stimuli; in turn, Set 2 involves two GAL stimuli and zero NGAL stimuli.

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

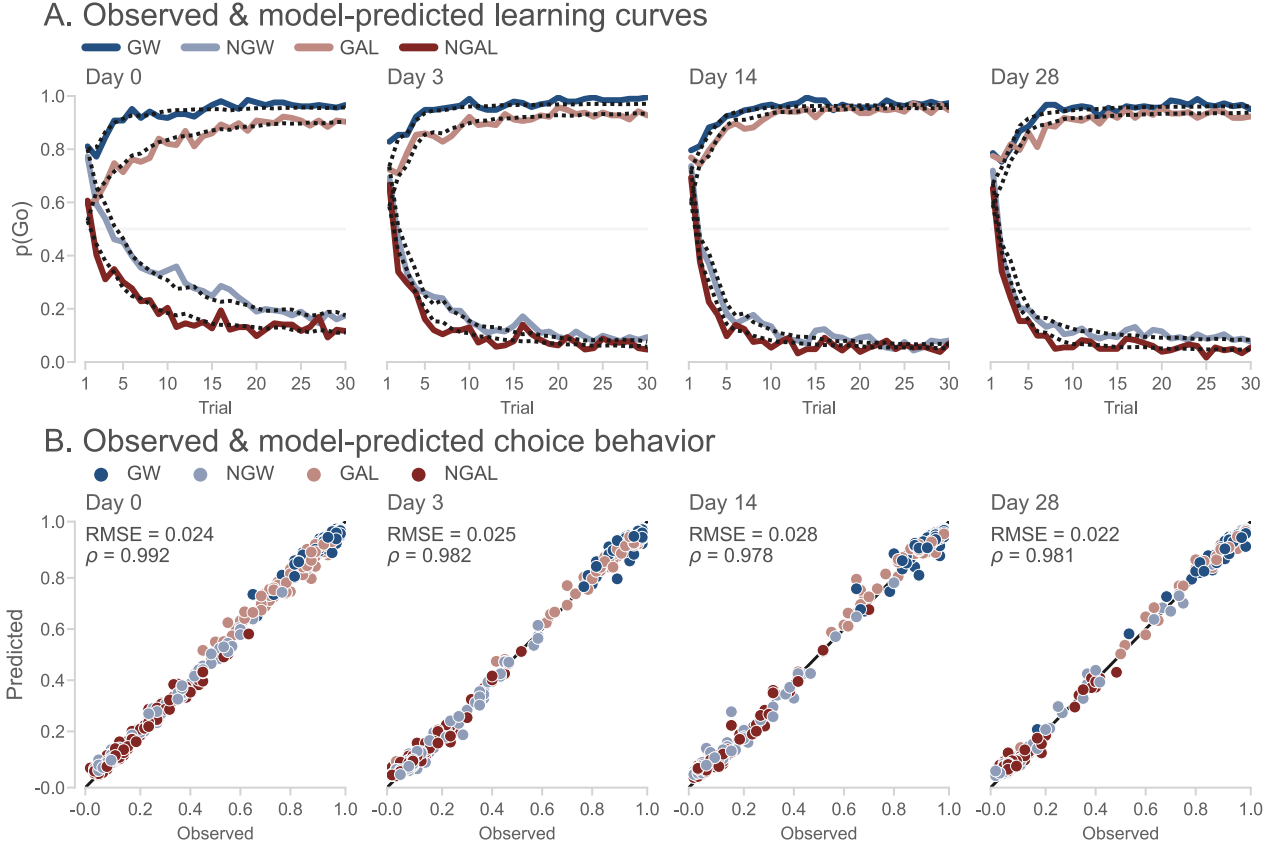


Figure S3: Observed & model-predicted choice behavior for Experiment 1. (A) Trial-by-trial choice behavior. Solid and dotted lines depict observed and model-predicted choice behavior, respectively. (B) Observed (x-axis) and model-predicted (y-axis) proportions of go responses for each trial type. Each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

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

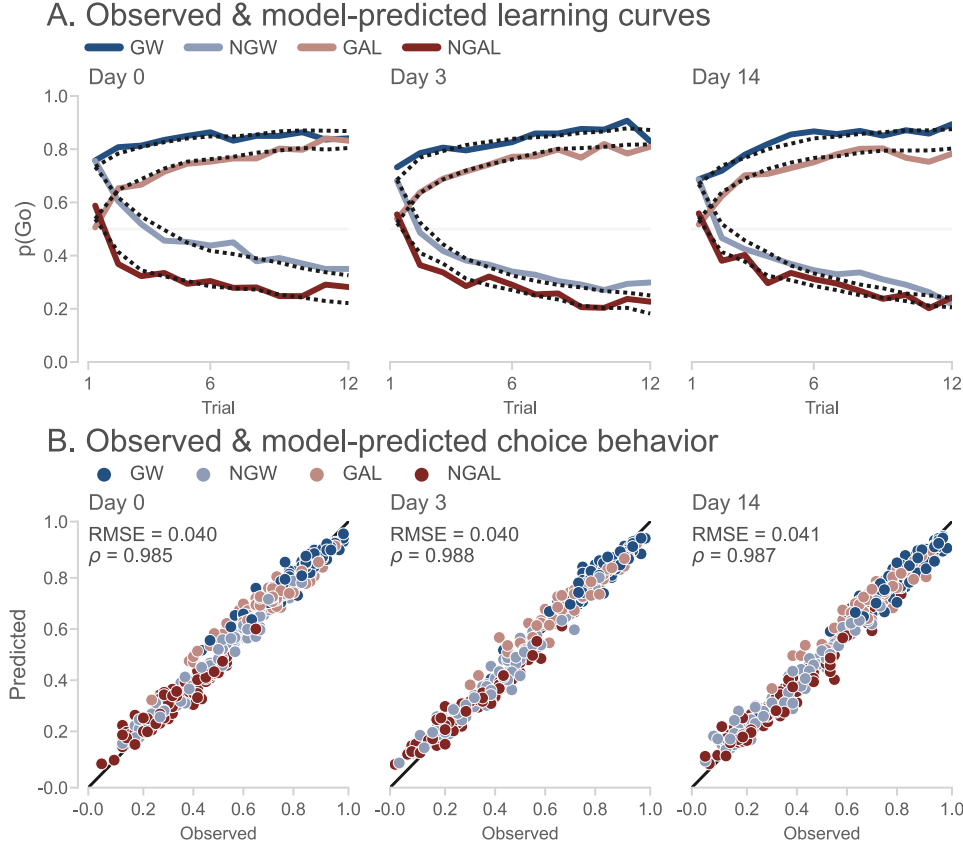


Figure S4: Observed & model-predicted choice behavior for Experiment 2. (A) Trial-by-trial choice behavior. Solid and dotted lines depict observed and model-predicted choice behavior, respectively. (B) Observed (x-axis) and model-predicted (y-axis) proportions of go responses for each trial type. Each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

Split-half reliability of model parameters

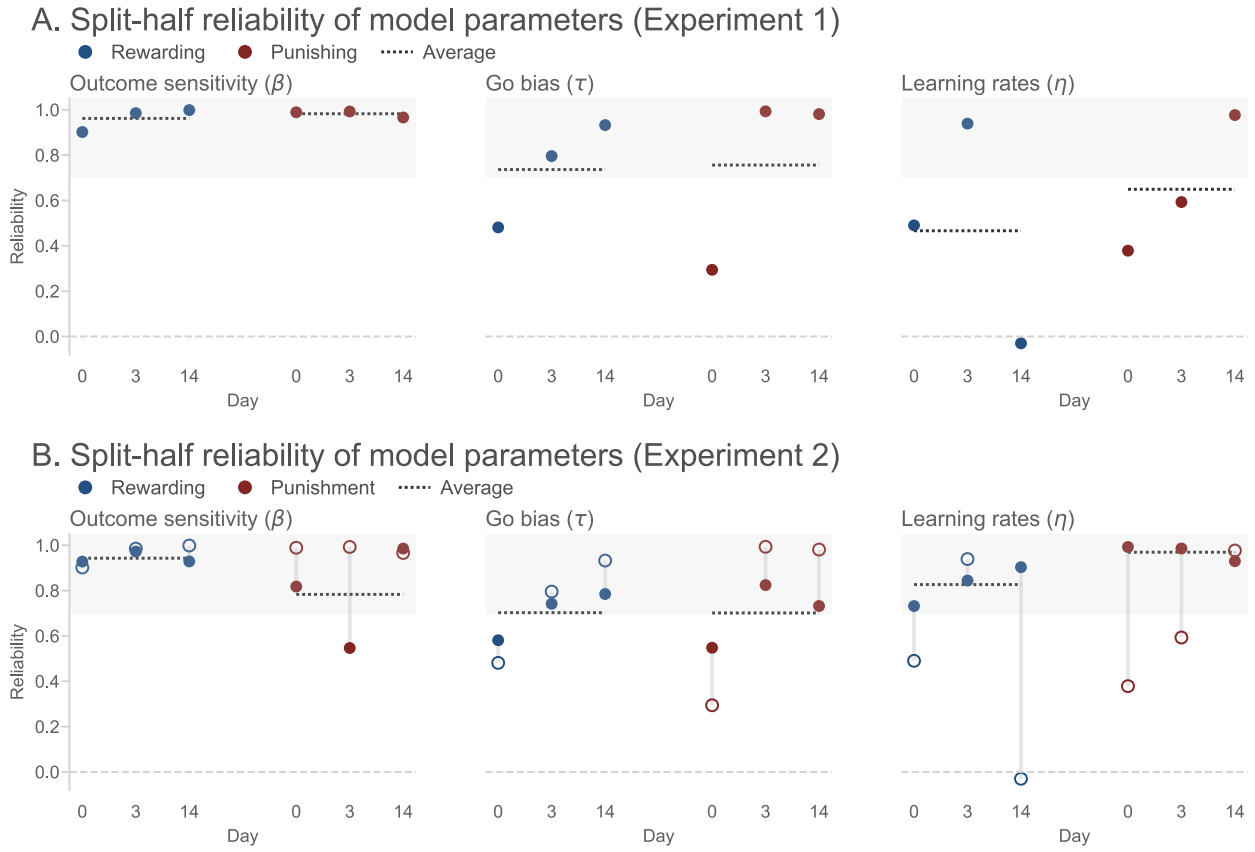


Figure S5: Split-half reliability estimates for the best-fitting model parameters for Experiment 1 (A) and Experiment 2 (B). Filled circles denote estimates for each experiment; open circles denote estimates from Experiment 1. The grey vertical lines show the change in reliability across experiments.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

Participant demographics

Variable	Experiment 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, yrs			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. Notes: Participants were able to select more than one ethnic and racial identity. Therefore, the participant counts and percentages in this section sum up to more than 100%. The mean age of each sample was compared via the 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).

Task appraisals

	Study	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. 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).

Priors for hierarchical Bayesian reinforcement learning models

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

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

ahahahhahha

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 ($\Delta\%$)	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 ($\Delta\%$)	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 ($\Delta\%$)	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 ($\Delta\%$)	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: Caption

Model comparison by session (Experiment 1)

Session	Model	Acc. (%)	LOO	Δ 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. Notes: Acc. = trial-level classification accuracy between observed and model-predicted go responses. LOO = Bayesian leave-one-out cross-validation presented in deviance scale (i.e., smaller numbers indicate better fit). Δ LOO = difference in LOO values between models.

bbbbbbbbb

Variable	Within-session statistics					Between-session comparisons		
	Day	Mdn	d	p	95% CI	Day 0	Day 3	Day 14
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 ($\Delta\%$)	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 ($\Delta\%$)	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 ($\Delta\%$)	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 ($\Delta\%$)	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: Caption

Model comparison by session (Experiment 2)

Session	Model	Acc. (%)	LOO	Δ 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. Notes: Acc. = trial-level classification accuracy between observed and model-predicted go responses. LOO = Bayesian leave-one-out cross-validation presented in deviance scale (i.e., smaller numbers indicate better fit). Δ LOO = difference in LOO values between models.

Supplementary references

1. 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).
2. 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).
3. 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).
4. Betella, A. & Verschure, P. F. The affective slider: A digital self-assessment scale for the measurement of human emotions. *PloS one* **11**, e0148037 (2016).
5. 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).
6. Guitart-Masip, M. *et al.* Go and no-go learning in reward and punishment: interactions between affect and effect. *Neuroimage* **62**, 154–166 (2012).
7. 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).
8. 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).
9. Swart, J. C. *et al.* Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in) action. *Elife* **6**, e22169 (2017).
10. Carpenter, B. *et al.* Stan: A probabilistic programming language. *Journal of statistical software* **76**, 1–32 (2017).
11. Rouder, J. N. & Haaf, J. M. A psychometrics of individual differences in experimental tasks. *Psychonomic bulletin & review* **26**, 452–467 (2019).