# Journal of Experimental Psychology: Learning, Memory, and Cognition

## Cognitive Training Reduces the Strength of Pavlovian Biases

Hugo Fleming, Gloria W. Feng, Robb B. Rutledge, Jonathan P. Roiser, and Oliver J. Robinson

Online First Publication, December 11, 2025. https://dx.doi.org/10.1037/xlm0001561

CITATION

Fleming, H., Feng, G. W., Rutledge, R. B., Roiser, J. P., & Robinson, O. J. (2025). Cognitive training reduces the strength of Pavlovian biases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://dx.doi.org/10.1037/xlm0001561

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: Learning, Memory, and Cognition


ISSN: 0278-7393

https://doi.org/10.1037/xlm0001561

# Cognitive Training Reduces the Strength of Pavlovian Biases

Hugo Fleming[1, 2], Gloria W. Feng[3], Robb B. Rutledge[3, 4], Jonathan P. Roiser[2], and Oliver J. Robinson[2]

[1] Medical Research Council Cognition and Brain Sciences Unit, University of Cambridge
[2] Institute of Cognitive Neuroscience, University College London
[3] Department of Psychology, Yale University
[4] Wellcome Centre for Human Neuroimaging, University College London

Pavlovian biases are patterns of behavior that involve approaching stimuli associated with reward and avoiding those associated with punishment (regardless of whether this is actually optimal behavior). They are a ubiquitous feature of everyday decision making and are also believed to play an important role in the symptoms of anxiety and depression. Although Pavlovian biases have classically been described as fixed and automatic, some studies have indicated that their influence on behavior can actually vary over time and with task demands. While these results hint that people may have some control over their Pavlovian biases, direct behavioral evidence for this control is still lacking. In a preregistered, double-blind, sham-controlled study ($N = 800$), we tested whether a week-long cognitive training intervention could reduce Pavlovian biases on the orthogonalized go/no-go task, a well-established paradigm for isolating Pavlovian-instrumental conflict. Participants were trained on either high-conflict or no-conflict conditions of the task across 5 days. Using reinforcement learning models to dissociate components of decision making, we found that high-conflict training led to a significant reduction in Pavlovian bias—particularly avoidance bias—at follow-up. This result is incompatible with the view that Pavlovian biases are fixed and automatic, and instead implies much greater flexibility in the way that they influence cognition than has previously been understood. The training was kept deliberately simple (i.e., one stimulus per condition, with the correct responses kept constant over sessions) so as to provide a minimal proof of concept of whether Pavlovian biases can be reduced through training, but as a result, we did not observe transfer to other tasks or self-reported mood. Nonetheless, these findings demonstrate that targeted cognitive training can modulate Pavlovian biases, which may be beneficial both in everyday life and especially in the context of affective disorders like anxiety and depression.

*Keywords:* Pavlovian bias, cognitive control, reinforcement learning, anxiety, depression

*Supplemental materials:* https://doi.org/10.1037/xlm0001561.supp

Multiple systems govern how humans and other animals select actions. The instrumental system learns the associations between actions and outcomes, and can thereby select an appropriate response that will maximize reward or minimize punishment (Dickinson & Balleine, 2002). A second system is the Pavlovian system, which learns the associations among different stimuli and promotes fixed responses: The Pavlovian system invigorates action whenever rewards are expected (often described as an approach bias) and inhibits action when punishment is anticipated (avoidance bias; Dayan & Balleine, 2002; Dayan et al., 2006). While these Pavlovian biases are generally advantageous, sometimes they conflict with the responses produced by the more flexible instrumental system, such as when one needs to resist approaching an immediate reward or take action (e.g., escape) in a potentially dangerous environment (Boureau & Dayan, 2011; Guitart-Masip et al., 2014). Consider, for example, the ambush predator that starts its chase too early, allowing its prey to escape, or the proverbial rabbit in the headlights that sees a car advancing toward it at speed, but freezes instead of fleeing.

Early work on Pavlovian-instrumental interactions often characterized Pavlovian biases as automatic and evolutionarily hardwired

Ronaldo Vigo served as action editor.

Hugo Fleming https://orcid.org/0000-0001-5421-8222

This study was preregistered on the Open Science Framework (https://osf.io/m3y8u/). All study materials are publicly available at https://app.gorilla.sc/openmaterials/669092. All primary data and analysis scripts are publicly available at https://osf.io/7msvw/.

This research was supported by a Wellcome Trust 4-year PhD Studentship awarded to Hugo Fleming. Open Access was supported by the U.K. Medical Research Council.



Hugo Fleming played a lead role in formal analysis, investigation, methodology, writing–original draft, and writing–review and editing and an equal role in conceptualization. Gloria W. Feng played a supporting role in formal analysis and writing–review and editing and an equal role in resources. Robb B. Rutledge played a supporting role in formal analysis and writing–review and editing and an equal role in resources. Jonathan P. Roiser played an equal role in conceptualization, supervision, and writing–review and editing. Oliver J. Robinson played a lead role in supervision and an equal role in conceptualization and writing–review and editing.

Correspondence concerning this article should be addressed to Hugo Fleming, Medical Research Council Cognition and Brain Sciences Unit, University of Cambridge, 15 Chaucer Road, Cambridge CB2 7EF, United Kingdom. Email: hugo.fleming@mrc-cbu.cam.ac.uk

(Boureau & Dayan, 2011; Dayan et al., 2006; Guitart-Masip et al., 2011, 2014). This conclusion was initially inspired by classical conditioning results suggesting that Pavlovian responses are persistent and difficult to overcome (e.g., "punished pecking" experiments by Williams & Williams, 1969; see also Hershberger, 1986). Computational and theoretical perspectives further highlighted that dopamine plays a dual role in behavior as both a learning signal (specifically associated with the reward prediction error in temporal-difference models of learning; Schultz et al., 1997) and a potent driver of motivation and action. In this framework (e.g., Boureau & Dayan, 2011), when a cue is presented which predicts reward, dopamine is released into the dorsal striatum, which promotes approach toward and acquisition of the subsequent reward (via the "go" pathway); conversely, when a punishment cue is presented, a dip in dopamine causes inhibition of behavior (via the "no-go" pathway), facilitating avoidance. Because both action and reward predictions are implemented by the same neurochemical substrate, the link between the two—resulting in the Pavlovian biases—was thought to be fixed.

In the years since, however, emerging evidence has suggested that the expression of Pavlovian biases may be more flexible than initially thought. For example, the balance between Pavlovian and instrumental systems has been shown to vary as a function of task parameters, such as whether outcomes are stochastic versus deterministic (Dorfman & Gershman, 2019), and to correlate with neural signals associated with top-down cognitive control (e.g., frontal theta; Cavanagh et al., 2013; see also Gershman et al., 2021). These findings suggest that, in specific contexts, people can overcome their Pavlovian biases through cognitive control. Furthermore, reductions in Pavlovian bias have been observed both within a single session (Dorfman & Gershman, 2019) and across extended practice (Schurr et al., 2024), suggesting that these biases are modifiable and can diminish with experience. It remains unclear, however, whether people can learn to strengthen their control over these biases through training, rather than merely adapting incidentally to changes in task structure, stochasticity, or repeated exposure. In particular, neither study included a control condition that would allow one to isolate the specific effect of training (i.e., improved engagement of top-down control). Our study addresses this gap by including a randomized control condition, which enables a direct causal test: By comparing high-conflict training with a control group, we can attribute reductions in Pavlovian bias specifically to the targeted training intervention, rather than to general task exposure or repeated practice.

Previous attempts to train Pavlovian biases have so far been unsuccessful. For example, Ereira et al. (2021) conducted four experiments with variants of the orthogonal go/no-go task (Guitart-Masip et al., 2011), looking at the effects of variables like gamification and timing, but found no evidence of training effects (a fifth study showed improvements on a "semantic" Pavlovian bias, which is conceptually quite different from the motor biases on which the literature is based). These findings suggest that direct training effects are difficult to achieve, if indeed this is possible at all. Thus, the question remains as to whether cognitive control over Pavlovian biases, specifically in the motoric domain, can be learned.

It is important to note that the modulation of Pavlovian biases being discussed here is not equivalent to extinction learning. In extinction, the Pavlovian association is weakened or eliminated by withholding the outcome. In contrast, in Pavlovian-instrumental conflict tasks, the Pavlovian contingencies continue to be reinforced, and the challenge lies in overcoming the prepotent behavioral bias they induce. Thus, the question is not whether the association itself can be unlearned, but whether its influence on behavior can be suppressed.

In addition to its theoretical relevance, the ability to train control over Pavlovian biases has important clinical implications. These biases—particularly the Pavlovian avoidance bias—are enhanced in patients with depression or anxiety and are thought to contribute to the development and maintenance of symptoms (Mkrtchian et al., 2017; Nord et al., 2018). For instance, a heightened avoidance bias may lead individuals with social anxiety to withdraw in social situations, which can result in more awkward or strained interactions and reinforce negative expectations. In such cases, the expression of a prepotent avoidance tendency prevents effective instrumental engagement—contributing to a self-reinforcing cycle of anxiety and social difficulty. This mirrors the structure of our task, where failure to act in the face of conflict increases the likelihood of punishment. Demonstrating that such biases can be overcome through training suggests a potential route for intervention: By strengthening top-down control over these prepotent tendencies, it may be possible to mitigate symptoms rooted in Pavlovian-instrumental conflict. While our study focused on a healthy population, it provides a foundational proof of concept for targeted cognitive training in clinical contexts.

In the present study, a large-scale ($N = 800$), preregistered, double-blind trial, we assessed Pavlovian biases using the orthogonal go/no-go task (Guitart-Masip et al., 2011; see also Guitart-Masip et al., 2014). In this task, the required response and valence of each trial are varied, such that on half the trials there is Pavlovian-instrumental conflict, requiring cognitive control over Pavlovian biases to respond correctly (see Table 1). A consistent finding from this task is that, although participants do gradually learn the contingencies up to a point, accuracy for the two high-conflict trial types reaches a plateau, which is below that of the no-conflict trial types and well below 100% (see, e.g., Figure 2 of Guitart-Masip et al., 2012). After a baseline testing session with the full task, participants doing the high-conflict training practiced only the "hard," control-demanding conflict trials once a day for 5 days, while those in the control intervention practiced the "easy," no-control trials. Finally, both groups then repeated the full task at a follow-up assessment session (using the same stimuli as at baseline and that they had trained on). Our primary interest here was in testing whether Pavlovian biases can in principle be controlled, so we simplified the task so that there was just one stimulus to learn per trial type, and we recruited a large sample of 800 participants, enabling us to

**Table 1**
*The Four Trial Types of the Orthogonal Go/No-Go Task (Guitart-Masip et al., 2011)*

| Action | Reward | Punishment |
|---|---|---|
| Go | Go to win reward | Go to avoid punishment |
| No-go | No-go to win reward | No-go to avoid punishment |

*Note.* Across four different trial types, participants have to make either a go or no-go response, for which they would either receive a reward for a correct response (and a neutral outcome otherwise) or a punishment for an incorrect response (and a neutral outcome otherwise). This produces two "easy" trial types for which the Pavlovian and instrumental systems are aligned (light gray) and two "hard" trial types for which they are in conflict (dark gray).

draw strong conclusions about the efficacy of the training. We hypothesized that the group that practiced the control-demanding, Pavlovian-instrumental conflict trials would show a greater improvement in accuracy at the follow-up session than the control group. We also included depression and anxiety self-report scales and secondary tasks to assess transfer effects to other domains.

# Method

## Preregistration

This study was preregistered on the Open Science Framework (https://osf.io/m3y8u/overview).

## Participants

In total, 800 adults participated in this study, recruited through the online platform Prolific. All were fluent in English and reported no history of psychiatric or neurological disorders. Participants were informed that they would be doing a week of practice on the task, but they were not told anything else about the nature of the training or indeed the existence of active and sham versions of the intervention.

After examining the data, we excluded 110 participants (a schedule is provided in Table 2). The exclusion criteria were all preregistered except one, which excluded participants who, during the go/no-go task, responded with keys outside the response set ("S" or "L" keys) on more than 15% of trials. This criterion led to the exclusion of three participants; no other participants made nearly so many wrong-key responses (the maximum among the other participants was just 3%).

This left us with 690 participants whose data were included in the final analysis, exceeding our preregistered minimum sample size of 676 (which was determined by *a priori* power analysis, using $d = 0.25$, $\alpha = 5\%$, and power = 90%).

The study was approved by the University College London Research Ethics Committee (6198/001).

## Procedure

The study comprised three phases, which took place over 8 days, as shown in Figure 1.

On the first day, participants completed a baseline testing session in which they completed the orthogonal go/no-go task (Guitart-Masip et al., 2011), the affective bias task (Daniel-Watanabe et al., 2022), the risk taking task (Rutledge et al., 2016), and two mental health questionnaires (the Beck Depression Inventory, Beck et al., 1996, and the State–Trait Anxiety Inventory, Spielberger et al., 1983). Then, they were randomly allocated to receive either the high-conflict or no-conflict training—those in the high-conflict group were given solely the control-demanding, Pavlovian conflict trials of the go/no-go task to practice, while those in the no-conflict group practiced just the no-conflict trials. Participants had to complete five training sessions over 6 days (with a maximum of one training session per day allowed); participants who had not completed all five sessions by the end of the training period were excluded from the study. Finally, on the 8th day of the study, participants completed a follow-up session containing the same battery of assessments as at baseline.

These sessions were conducted entirely online, using the experiment platform Gorilla (https://www.gorilla.sc), which also performed the randomization to the high-conflict or no-conflict training groups automatically.

## Measures and Tasks

### *Orthogonal Go/No-Go Task (Guitart-Masip et al., 2011)*

The full procedure for this task is set out in Figure 2A. A trial consisted of three events, each displayed for 1,000 ms with a 250 ms interstimulus interval: First, an initial fractal cue was shown in the center of the screen; then, a circle target was displayed on one side of the screen, to which participants chose whether or not to respond; finally, the outcome of their response was displayed.

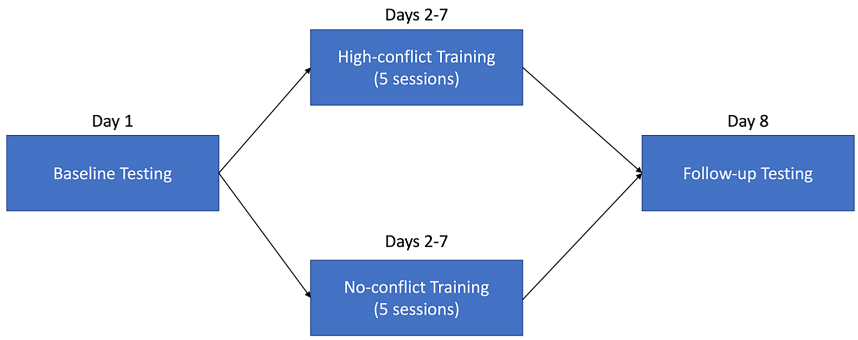
Each fractal was associated with both a required response ("go" or "no-go") and a valence (correct responses allowed participants either to win points or avoid losing them). Combining these

**Table 2**
*Schedule of Exclusions*

| Time point | Reason | *N* excluded | *N* remaining |
|---|---|---|---|
| Baseline testing | | | 800 |
| | Did not complete baseline session | 11 | 789 |
| | GNG: Go to win reward accuracy <65% | 9 | 780 |
| | GNG: Left/right accuracy <65% | 1 | 779 |
| | Aff. bias: Accuracy on unambiguous trials <60% | 29 | 750 |
| | Aff. bias: No response on >15% of trials | 2 | 748 |
| | STAI: Failed attention check | 2 | 746 |
| Training | Did not complete five training sessions | 46 | 700 |
| Follow-up testing | Did not complete follow-up session | 1 | 699 |
| | GNG: Go to win reward accuracy <65% | 2 | 697 |
| | GNG: Wrong-key responses >15%[a] | 3 | 694 |
| | Aff. bias: Accuracy on unambiguous trials <60% | 1 | 693 |
| | Aff. bias: No response on >15% of trials | 2 | 691 |
| | STAI: Failed attention check | 1 | 690 |

*Note*. GNG = go/no-go task; Aff. bias = affective bias task; STAI = State–Trait Anxiety Inventory.
[a] This was not a preregistered criterion—see the Participants section for details.

**Figure 1**
*Timeline of the Study*

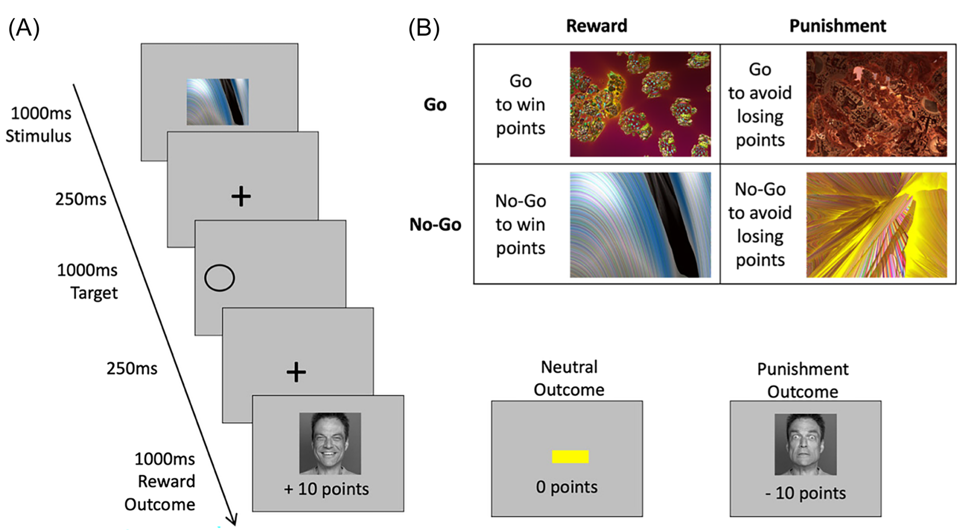



*Note*. Participants completed the full battery of tasks at baseline and then were randomized to receive either the high-conflict or no-conflict training. They completed five practice sessions over 6 days and then repeated the full battery of tasks at follow-up. See the online article for the color version of this figure.

responses and valences orthogonally produced four unique trial types (see Figure 2B), each represented by a different fractal: go to win reward, go to avoid punishment, no-go to win reward, and no-go to avoid punishment. Participants had to learn the correct responses (and maximize their points) through trial and error. Note that the outcomes were probabilistic, such that a correct response only led to reward (or avoided punishment) 80% of the time. In addition, to hold participants' attention, the exact keypress for a "go" response ("S" or "L") had to match whichever side of the screen the target was presented on.

After a number of practice rounds, during which the complexity of the task was built up gradually, the main phase of the task comprised 80 trials (20 per condition) presented in a random order. The fractal allocation was randomized for each participant at the start of the baseline session and then retained for the training and follow-up sessions.

**Figure 2**
*Procedure for the Orthogonal Go/No-Go Task*




*Note*. (A) Participants were shown an initial fractal cue, associated with a required response (go or no-go) and a valence (reward or punishment). Participants made or omitted a response when the target appeared and then received an outcome depending on their response and the fractal valence. (B) The four possible trial types (combining the possible responses and valences orthogonally). This figure builds on Table 1, additionally showing a possible allocation of the fractals to the four trial types. In the study, the fractal allocation was randomized for each participant at baseline. Example fractal images are reproduced from Mathôt et al. (2015) under a Creative Commons Attribution (Unported) License. Example face images are taken from the publicly available subset of the FACES database (Ebner et al., 2010). See the online article for the color version of this figure.

### *Pavlovian Bias Training*

Participants were trained on a variant of the same orthogonal go/no-go task described above, but with just a subset of conditions: The high-conflict training group practiced just the trial types that involved Pavlovian-instrumental conflict ("go to avoid punishment" and "no-go to win reward"), while those in the no-conflict group were trained on just the no-conflict trial types ("go to win reward" and "no-go to avoid punishment"). The fractal allocation was the same as at baseline and follow-up. Each training session comprised 48 trials (24 trials per condition, in a random order), and participants were required to complete at least five training sessions in 6 days (one session per day). We did not collect information on the time of day participants completed the training sessions, but as participants were randomized to the two conditions, this would not be expected to vary systematically between the groups.

### *Secondary Outcomes*

We included two further tasks to investigate whether the training, if successful on the main task, would also transfer to other domains. The visual affective bias task (Daniel-Watanabe et al., 2022) assessed how participants' perceptual judgments about the size of an ambiguous stimulus (a medium-sized black disc) were affected by receiving asymmetric rewards for choosing large versus small. The risk taking task (Rutledge et al., 2016) measured participants' tendency to gamble versus take a safe, certain outcome, as the expected values of these options were varied. See the Supplemental Material for further details.

We also administered two well-validated mental health questionnaires: the Beck Depression Inventory (Beck et al., 1996) and the State–Trait Anxiety Inventory (Spielberger et al., 1983). Both instruments have robust psychometric properties, including high internal consistency and test–retest reliability, as reported in the original validation studies. We removed one question from the Beck Depression Inventory, which asks about thoughts of suicide, due to the safeguarding risk. We also added a catch question ("Press the very much so button") at the end of the State–Trait Anxiety Inventory to detect inattentive participants.

## Preregistered Analyses

To test our primary hypothesis that the training would enhance control over Pavlovian biases in the orthogonal go/no-go task, we planned two related analyses, one model-agnostic and another that involved computational modeling.

### *Model-Agnostic Analysis of the Orthogonal Go/No-Go Task*

For the model-agnostic analysis, we first calculated a measure of Pavlovian bias for each participant in each session. This was defined as the sum of the accuracies for the two Pavlovian-instrumental conflict trial types (go to avoid punishment and no-go to win reward) minus the sum of the two no-conflict trial types (go to win reward and no-go to avoid punishment). We then computed a training effect, which was the change in this measure between the baseline and follow-up sessions. Finally, we tested (using an independent-samples *t* test) our primary hypothesis that there would be a difference in the change in this Pavlovian bias metric between the high-conflict and no-conflict training groups.

We also preregistered a secondary hypothesis that participants in both groups would exhibit Pavlovian biases at baseline. To test this, we ran a 2 × 2 (Required Response × Valence) repeated measures analysis of variance on the accuracy data from the baseline session, followed by four planned paired-samples *t* tests comparing the go to avoid punishment and no-go to win reward conditions with each of the go to win reward and no-go to avoid punishment conditions.

### *Computational Modeling of the Orthogonal Go/No-Go Task*

In parallel, we also tested our primary hypothesis using computational modeling. Our models build on established reinforcement learning frameworks that combine instrumental Q-learning (Rescorla & Wagner, 1972; Sutton, 1988) with Pavlovian value influences (Dayan et al., 2006). In particular, we use as a starting point the winning model from Guitart-Masip et al. (2012) and call this the base model. In this model, each trial's action value is computed from the expected instrumental reward (the difference between the $q$ values for go and no-go), modulated by fixed biases reflecting approach and avoidance tendencies. These biases multiply the Pavlovian value of the present stimulus and so determine to what extent the Pavlovian estimates can influence the action weights; they promote go responses when reward is anticipated (i.e., value is positive) and no-go responses when punishment is expected (i.e., value is negative). In addition, we also assume that participants have a general go bias, which invigorates action regardless of the instrumental or Pavlovian values on that trial. These factors are summarized in Equation 1 below, which shows how the action weight $w(s)$, the log odds of making a go response, is calculated on each trial (where $s$ indexes the stimulus shown on that trial):

$$\begin{aligned} w(s_t) = q_{\text{go}}(s_t) - q_{\text{nogo}}(s_t) + \text{GoBias}_{\text{subject}} \\ + \text{Pavbias}_{\text{subject}} \times \text{value}(s_t). \end{aligned} \tag{1}$$

Here, $q_{go}(s_t)$ and $q_{nogo}(s_t)$ are the instrumental $Q$ values for the go and no-go actions respectively; $\text{GoBias}_{\text{subject}}$ reflects a general tendency to make go responses; $\text{Pavbias}_{\text{subject}}$ is the Pavlovian bias parameter; and $\text{value}(s_t)$ is the Pavlovian value of stimulus $s_t$.

The other key subject-level parameters in the base model are the outcome sensitivity, which acts as a multiplier on the outcome received on each trial (following Guitart-Masip et al., 2012, we assume that participants may value rewards and punishments differently), and the learning rate, which acts as a multiplier on the prediction error observed on each trial.

We subsequently considered three extensions to this model: one with separate Pavlovian approach and avoidance biases ("Base + 2PavBias"), which applied when the Pavlovian value of the stimulus was rewarding or punishing, respectively; another with distinct learning rates for reward and punishment outcomes ("Base + 2LR"); and a third with both approach/avoidance Pavlovian biases and reward/punishment learning rates ("Base + 2PavBias + 2LR").

The full model specifications for all four models, including parameter definitions, formal equations, parameter recovery, and model comparison results, are provided in the Supplemental Material (p. 2), where there is also further discussion of the relationship between these models and previous work. All model code and data are also publicly available on the Open Science Framework (https://osf.io/7msvw/overview).

For the baseline session, we fitted the models to the entire sample of participants all together, and then for the follow-up session, we fitted the models to the high-conflict and no-conflict groups separately (thus assuming that the two groups were identical before the intervention but might differ afterward).

All models were fitted using Markov Chain Monte Carlo in Stan (Stan Development Team, 2023). Sampling was run for four chains, each with 2,000 iterations. Subsequent to fitting, we carried out diagnostics (visual inspection of the chains; divergences or treedepth warnings; Estimated Bayesian Fraction of Missing Information < 0.3, effective sample size > 400, split-Rhat < 1.01; Betancourt, 2018). Besides a negligible number of divergences (<1% for all models), there were no issues. We also inspected the posterior predictions for the winning model and observed a good overall fit to the data (Supplemental Figure S12).

We compared these models using the Widely Applicable Information Criterion (Watanabe, 2010), which provides an estimate of out-of-sample predictive accuracy, then examined the posterior parameter values from the winning model. As with the model-agnostic approach described above, our preregistered analysis used an independent-samples *t* test to assess whether the mean change in the Pavlovian bias term differed between the high-conflict and no-conflict training groups. Further exploratory analyses are reported in the Supplemental Material.

### *Affective Bias Task*

Our preregistered analysis for this task involved assessing (using an independent-samples *t* test) whether the two training groups differed in their change in affective bias between baseline and follow-up.

### *Risk Taking Task*

Our preregistered analysis for this task involved assessing, using a 3 × 2 × 2 analysis of variance (Framing × Time Point × Training Group), whether there was an interaction between time point and group on gambling choices.

## Materials, Data, and Code Availability

All study materials are publicly available at https://app.gorilla.sc/openmaterials/669092. All primary data and analysis scripts are publicly available at https://osf.io/7msvw/overview.

# Results

As expected, both groups exhibited significant Pavlovian biases at baseline, action-by-valence interaction: $F(1, 689) = 709$, $p < .001$, $\eta_p^2 = 0.51$ (see Figure 3A, left panel). Accuracy was worse when participants had to go to avoid punishment compared with go to win reward, $t(689) = 18.4$, $p < .001$, $d = 0.70$, and likewise when participants had to no-go to win reward compared with no-go to avoid punishment, $t(689) = 23.9$, $p < .001$, $d = 0.91$ (both tests remained significant using a Bonferroni correction of $\alpha = .025$). We also found significant action and valence biases: Participants found it easier to learn go than no-go responses ($M = 0.76$, $SD = 0.17$ vs. $M = 0.40$, $SD = 0.25$), $F(1, 689) = 1{,}680$, $p < .001$, $\eta_p^2 = 0.71$, and they were better at avoiding punishment than winning reward ($M = 0.61$, $SD = 0.15$ vs. $M = 0.55$, $SD = 0.36$), $F(1, 689) = 97.3$, $p < .001$, $\eta_p^2 = 0.12$.

Our second planned analysis looked at the model-agnostic measure of Pavlovian bias. At follow-up, we saw a substantial difference between the training groups in the change in their Pavlovian biases, $t(688) = 11.9$, $p < .001$, $d = 0.91$. Specifically, the high-conflict training led to a large reduction in Pavlovian bias, $t(344) = 9.90$, $p < .001$, $d = 0.53$, while this bias in fact became stronger following the no-conflict training, $t(344) = 6.86$, $p < .001$, $d = 0.37$ (both *post hoc* tests again remained significant after Bonferroni correction). These results are plotted in Figure 3B, and descriptive statistics are given in Table 3.

In an exploratory analysis, we also used analysis of covariance to control for baseline differences in Pavlovian bias. In agreement with our preregistered analysis, we found there was a significant effect of training condition, $F(1, 687) = 210$, $p < .001$, $\eta_p^2 = 0.23$. Separately, we also verified that the high-conflict group showed a significantly greater improvement in overall accuracy, $t(688) = 2.29$, $p = .022$, $d = 0.17$.

## Performance During the Training Phase

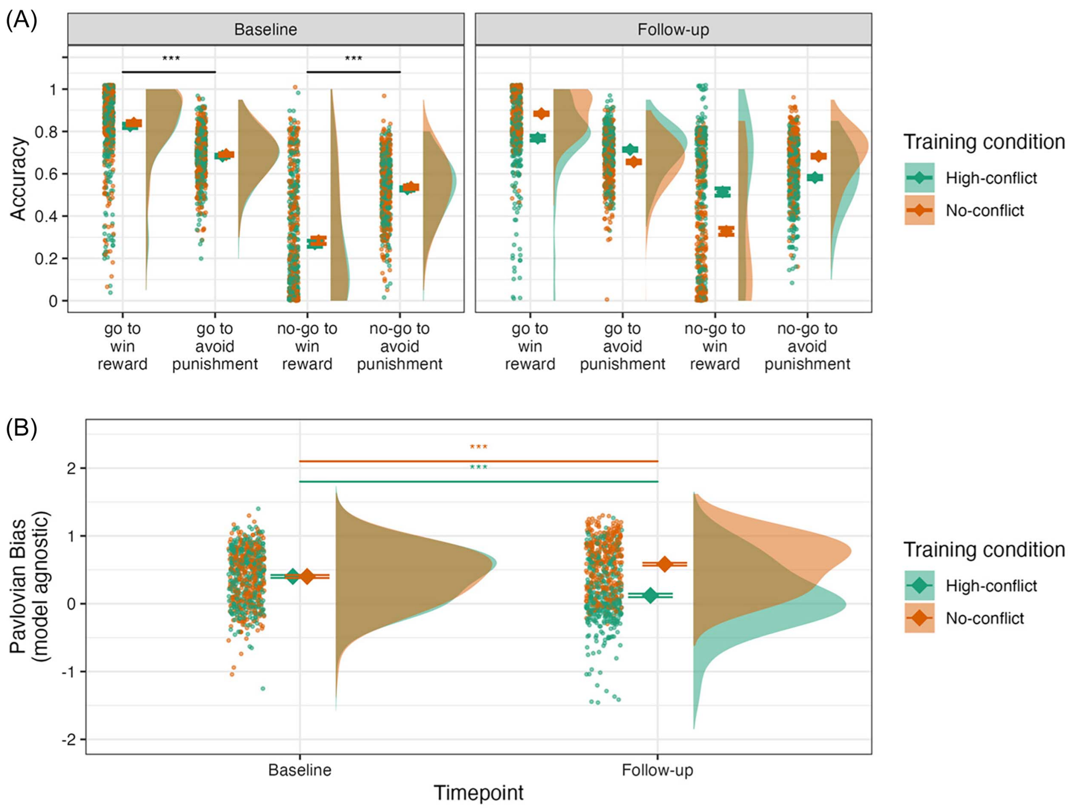
Turning to performance during the training phase itself, we observed a significant difference between groups (see Figure 4 and Supplemental Table S1). While both training groups improved their accuracy over the course of training, there was a significant interaction between training group and time point, $F(4, 2752) = 63.0$, $p < .001$, $\eta_p^2 = 0.08$, with the high-conflict group improving by a greater amount. Indeed, the average improvement in accuracy from the first to the fifth training sessions for the high-conflict training group was 0.16 ($SD = 0.18$), while for the no-conflict group, it was 0.05 ($SD = 0.09$), $t(491) = 9.91$, $p < .001$, $d = 0.76$.

Detailed analysis of the sequential differences between training sessions is provided in the Supplemental Material (p. 11).

## Computational Modeling

The best performing model had separate Pavlovian approach and avoidance biases and separate reward and punishment learning rates (see the Supplemental Material for full details of the model comparison). The Pavlovian avoidance parameter reflects the degree to which participants tended to withhold action in the presence of stimuli that had a Pavlovian association with punishment. Reductions in this parameter therefore indicate a diminished influence of the Pavlovian system on behavior. Examining the posterior estimates from this model (Figure 5), we found a significant difference between the high-conflict and no-conflict training groups in their change in avoidance bias, $t(688) = 36.1$, $p < .001$, $d = 2.75$, but not approach bias, $t(688) = 0.35$, $p = .73$. Specifically, the high-conflict training reduced the strength of participants' avoidance bias almost to zero on average (from 1.49 at baseline to 0.09 at follow-up: mean change = –1.40, $SD = 0.55$), $t(344) = 46.8$, $p < .001$, $d = 2.52$, whereas those in the no-conflict group showed a much smaller reduction (mean change = –0.24, $SD = 0.21$), $t(344) = 21.2$, $p < .001$, $d = 1.14$, again remaining significant after Bonferroni correction. It is worth noting that, by modeling the approach and avoidance biases separately, we have been able to reveal a much

**Figure 3**
*Control Over Pavlovian Biases Significantly Enhanced by High-Conflict Training*




*Note.* Both plots show the individual data points, the $M \pm SE$ (diamonds and horizontal lines), and the overall distribution. (A) Groups were closely matched at baseline and showed clear signs of Pavlovian bias (impaired accuracy at go to avoid punishment and no-go to win reward trials). Following training, the high-conflict (but not no-conflict) group showed a significant improvement in accuracy. (B) The high-conflict training, but not the no-conflict training, led to a significant decrease in the model-agnostic measure of Pavlovian bias. *SE* = standard error. See the online article for the color version of this figure.
*** $p < .001$.

larger change in (avoidance) bias than had been indicated by our model-agnostic analysis above.

As before, we also ran an exploratory analysis with analysis of covariance to control for any differences in baseline Pavlovian biases, and again this led to the same conclusions: a significant training effect on avoidance bias, $F(1, 687) = 1{,}310$, $p < .001$, $\eta_p^2 = 0.66$, but not approach bias, $F(1, 687) = 0.60$, $p = .4$.

Of note, we also observed a significant decrease between groups in their change in go bias, $t(688) = 11.4$, $p < .001$, $d = 0.87$. While both groups demonstrated significantly reduced go biases after training, the reduction was substantially greater in the active training group versus the sham group, $t(344) = 22.1$, $p < .001$, $d = 1.19$ and $t(344) = 8.51$, $p < .001$, $d = 0.46$, respectively, a reduction from $M = 1.29$, $SD = 0.43$ to $M = 0.54$, $SD = 0.59$ for the active group, compared with $M = 1.28$, $SD = 0.44$ to $M = 1.03$, $SD = 0.46$ for the sham group. This may partially account for the absence of an effect on the approach bias parameter (since both parameters promote go responses), a point we return to in the discussion (note that these *post hoc t* tests remained significant after correction for multiple comparisons).

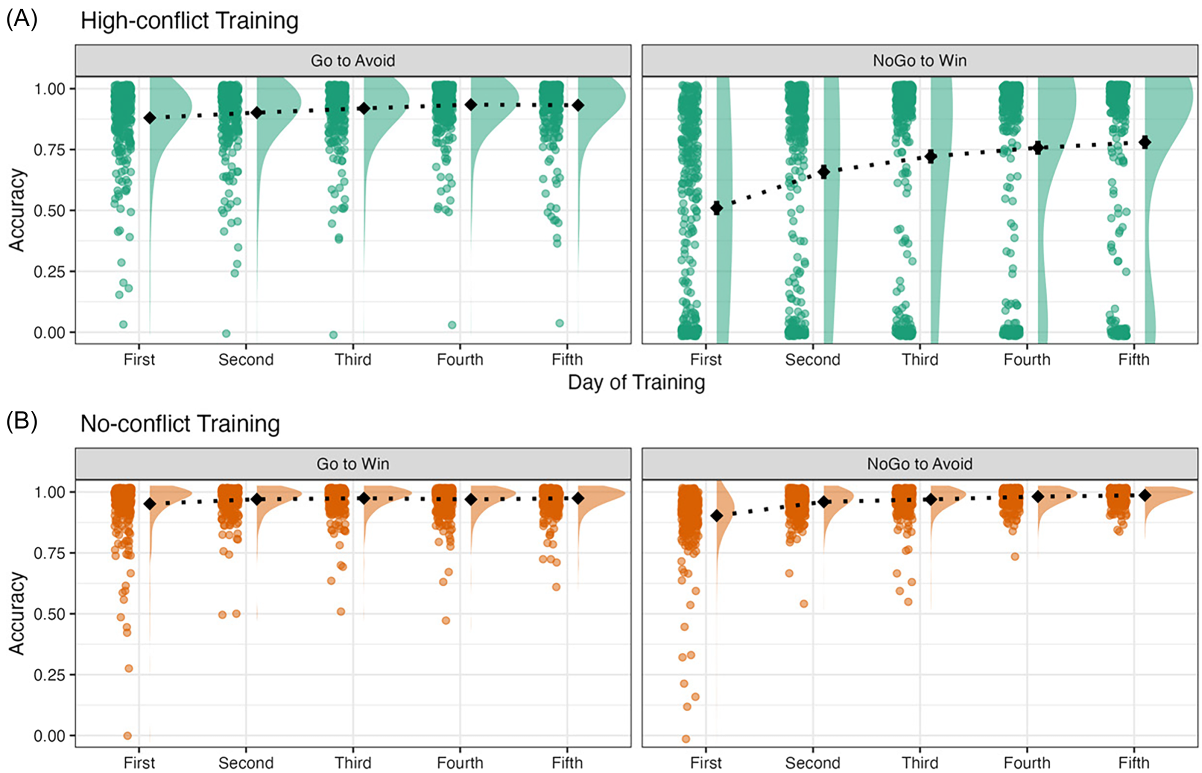
To assess the validity of these model-based inferences, we conducted a parameter recovery analysis by fitting the model to simulated data generated using the observed parameter values. This analysis tests whether the model can recover the original parameter estimates, which is essential for interpreting individual- and group-level differences in behavior. For the go bias parameter, both Pavlovian bias parameters and the punishment learning rate parameters could all be recovered reliably ($r > 0.8$). In contrast, reward learning rate, reward sensitivity, and punishment sensitivity parameters showed poorer recovery. Correlation analyses (see the Supplemental Material) suggested that this was not due to collinearity, implying that poor recovery of these

**Table 3**
*Model-Agnostic Pavlovian Bias Measure in Each Condition*

| Training condition | Time point | Pavlovian bias *M* (*SD*) |
|---|---|---|
| High-conflict | Baseline | 0.40 (0.39) |
| | Follow-up | 0.12 (0.48) |
| No-conflict | Baseline | 0.40 (0.40) |
| | Follow-up | 0.58 (0.38) |

**Figure 4**

*Performance on the Pavlovian Bias Training, Split by Trial Type*




*Note*. Average accuracy in both groups improved over the course of training, but the improvement was greater in the high-conflict compared with the no-conflict group. Plots show individual data points and distributions (color) and $M \pm SE$ (black). $SE$ = standard error. See the online article for the color version of this figure.

parameters is unlikely to have distorted estimation of the reliably recovered ones. To further validate our findings, we reran the primary analysis using the simpler Base model (which does not have the poorly recovering parameters) and obtained identical results (see Supplemental Figure S15). These checks confirm that the training effect on Pavlovian bias is robust and not dependent on model complexity or parameterization.

Finally, we investigated a possible mechanistic account of this change in Pavlovian bias, based on the model proposed in Dorfman and Gershman (2019). In that article, the authors proposed an algorithm in which the balance between Pavlovian and instrumental controllers is updated dynamically, based on their relative predictive power, through a process of Bayesian model averaging. After simulating new participants with their model (details are provided in the Supplemental Material), we found that it provides a qualitatively good match to our empirical results, reproducing an improvement following training that is specific to the conflict training condition only. This suggests that an adaptive process, which updates the weight given to the Pavlovian and instrumental controllers based on their ongoing predictive accuracy, could underlie the training effects that we observed here (Figure 6).

### Transfer Effects

We did not observe transfer to the other cognitive tasks or self-reported mood: The change in affective bias between time points did not differ between groups, $t(688) = 0.11$, $p = .91$, nor was there an interaction between time point and group for the risk taking task, $F(1, 688) = 0.03$, $p = .87$; depression, $F(1, 688) = 0.67$, $p = .41$; state anxiety, $F(1, 688) = 1.25$, $p = .26$; or trait anxiety, $F(1, 688) = 0.56$, $p = .46$. Other analyses of these tasks are reported in the Supplemental Material.

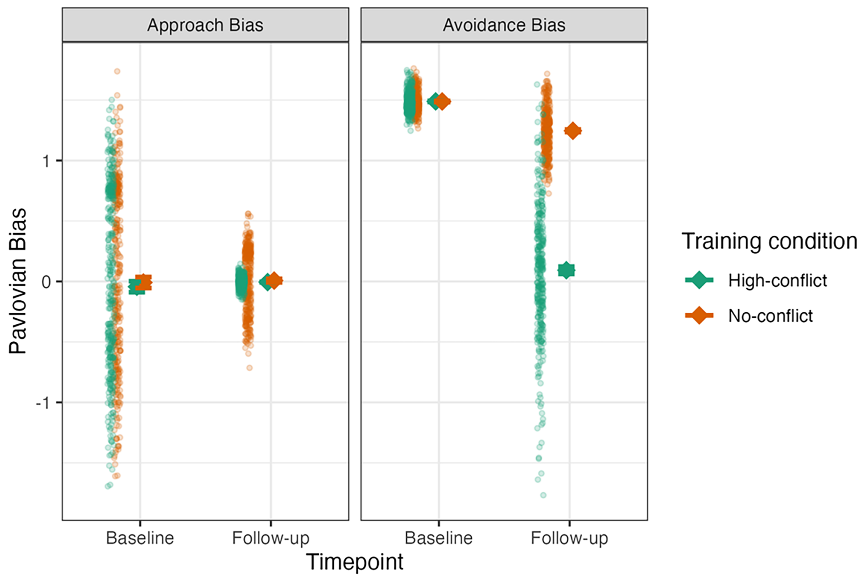
## Discussion

In this double-blind study, we examined whether control over Pavlovian biases can be learned. Over five sessions, participants in the high-conflict group practiced the control-demanding, Pavlovian-instrumental conflict trials of the orthogonal go/no-go task. Converging evidence from both model-agnostic and computational modeling analyses revealed that these participants were able to substantially reduce their Pavlovian biases (compared with participants in the no-conflict group); indeed, the avoidance bias was almost entirely eliminated, according to the results of the modeling. This clearly demonstrates that people can learn to control their Pavlovian biases through training.

The existing literature has historically regarded Pavlovian biases as being highly persistent and resistant to change. Our results here suggest that, while such biases are indeed strong, they can nevertheless be overcome through training. The inclusion of a sham training condition was essential for showing that the training effect

**Figure 5**

*Subjects' Pavlovian Bias Values, According to the Winning Model (Base Plus Two Pavlovian Biases Plus Two Learning Rates)*




*Note*. The plot shows each participant's mean bias and the $M \pm SE$ in each condition. *SE* = standard error. See the online article for the color version of this figure.

resulted specifically from practicing the high-conflict trials and was not simply a continuation of the asymptotic performance improvement seen at baseline. The earlier work of Cavanagh et al. (2013) suggests a possible mechanism for this effect: They found that accuracy on the go/no-go task covaried with frontal theta measured via electroencephalography, a neural signature of top-down control; similarly, Guitart-Masip et al. (2012) found using functional magnetic resonance imaging that activity in the inferior frontal gyrus is associated with successful performance and speculated that this region may help to regulate the balance between the Pavlovian and instrumental systems. Conceivably, then, in our study the training may have taught participants how and when to engage these control signals in order to mitigate Pavlovian bias and maximize performance on the task.

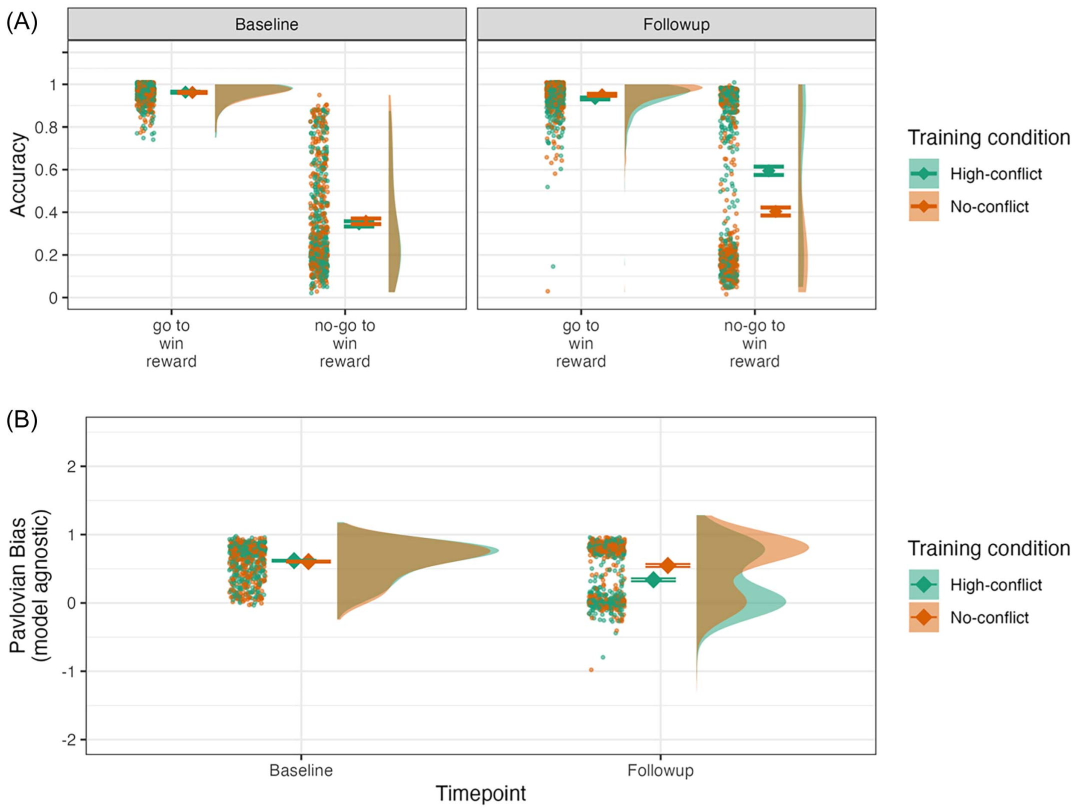
Training could have affected control in at least two ways: It could have increased participants' general ability to monitor for Pavlovian-instrumental conflict and deploy cognitive control as and when needed (termed "reactive control" by Braver, 2012), or it could have led to narrower learning, in the sense that participants may simply have learned which stimuli predicted the need to exert greater control later in the trial ("proactive control"). In the present study, we are not able to discriminate between these two possibilities, although we might speculate that reactive control requires more extensive training specifically aimed at transferring learning to novel situations and stimuli. A third possibility is that the training may not have impacted control at all and instead led to an improvement in performance through some other process, such as habit learning (Everitt & Robbins, 2005).

An important question is why the present study produced a clear training effect, while earlier attempts—such as Ereira et al. (2021)—did not. Both studies build on the original orthogonal go/no-go task (Guitart-Masip et al., 2011) but differ in several critical respects. Ereira et al. conducted four experiments with motor variants of the task and found no training effects; in a fifth experiment, they reported a reduced "semantic" bias—based on choices between words implying approach or avoidance—but this did not replicate in a nongamified version of their task. Their task also varied stimuli across sessions, which may have disrupted consistent learning. In contrast, our study followed the original task design but held all stimulus–response mappings constant, while the training phase focused solely on the conflict conditions, isolating the effect of repeated high-conflict exposure. Crucially, our findings go beyond methodological refinement: We provide the first causal demonstration that behavioral expressions of Pavlovian bias—particularly avoidance—can be deliberately suppressed through targeted training. This supports the view that cognitive control over these biases is not only possible but also trainable. Although constraining the training (e.g., one stimulus per condition, fixed mappings) likely limited the transfer to other contexts, it enabled a clean test of whether bias reduction could occur at all. Future studies can now build on this foundation to explore whether richer training regimes support generalization.

The present study also constitutes a successful proof of principle for cognitive bias modification, which may have wider applications. Patients with depression or anxiety (Mkrtchian et al., 2017; Nord et al., 2018) have been shown to have enhanced Pavlovian biases, possibly as a result of deficits in cognitive control (Robinson et al., 2013), and this is thought to contribute to symptoms through the maintenance of avoidance behavior. Our results open the possibility that cognitive bias training could eventually be used to treat some symptoms of depression and anxiety. This is a particularly exciting and significant avenue for future research because, as we have demonstrated here, such training is low-cost and can be deployed at scale through online platforms. We note that although in the present

**Figure 6**

*Simulated Accuracy and Pavlovian Bias From the Adaptive Model of Pavlovian Instrumental Control (Dorfman & Gershman, 2019)*




*Note.* (A) Simulated accuracy data across the two reward conditions. (B) Simulated model-agnostic Pavlovian bias metric. Comparing with the empirical data (Figure 3) indicates that this model provides a qualitatively good fit and so may help to explain the mechanism by which the training had its effect. Note that the model does not at present handle punishment outcomes, and so we simulated just the two reward conditions. See the online article for the color version of this figure.

study we did not see any effects on depression or anxiety, we had specifically recruited participants with no history of psychiatric illness, so it is likely there was already a floor to improvement in symptom scores; nevertheless, we will need to examine how these results extend to patient groups in future research. Additionally, the current design focused on immediate training effects, so further work is required to assess whether reductions in Pavlovian bias are sustained over time.

One further limitation of this study is that, as a control condition, we included only a sham training condition and not a passive control (i.e., one with no training at all, which would indicate the simple effect of passage of time on the go/no-go task). It is conceivable that people may improve over time even without training and that we observed a difference between the high-conflict and no-conflict groups not because the high-conflict training was effective but conversely because the no-conflict training interfered with this improvement. While we cannot definitively refute this explanation here, we suggest that the very fact that Pavlovian biases have until now been widely considered to be extremely resistant to modification indicates that improvement simply due to the passage of time is unlikely.

Finally, it is notable that the high-conflict training led to a large reduction in avoidance bias but no change in approach bias. This may be because avoidance is more amenable to change. However, it may also be driven by floor effects, as the approach bias parameters were already around zero at baseline (the effect of training was then to shrink the variance in these estimates, rather than shift the mean). It is also possible that some of the variance in approach-related responding was absorbed by the go bias parameter, thereby obscuring potential training effects on the approach bias itself. In any case, our sensitivity analysis using the base model (which includes a single, combined Pavlovian bias term) produced qualitatively identical results. This provides reassurance that the training effect on Pavlovian bias is robust and not dependent on parametrization or model structure.

## Conclusions

In sum, this study provides causal evidence that Pavlovian bias—particularly avoidance bias—can be selectively reduced through targeted, repeated training on control-demanding trials. While previous work has shown that such biases can change with context or decline over time, we show that they can be actively suppressed through focused training. Crucially, our sham-controlled, randomized design allows us to rule out general explanations for the improvement, such as increased task familiarity, and suggests instead that it was the repeated practice of employing cognitive control that was important. These findings establish that Pavlovian biases are not fixed and instead are amenable to training, which provides a foundation for future research aimed at generalizing such training to new stimuli, wider environments, and clinical populations where Pavlovian biases contribute to maladaptive behavior.

## References

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. Psychological Corporation.

Betancourt, M. (2018). *A conceptual introduction to Hamiltonian Monte Carlo*. arXiv. https://doi.org/10.48550/arXiv.1701.02434

Boureau, Y.-L., & Dayan, P. (2011). Opponency revisited: Competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology*, *36*(1), 74–97. https://doi.org/10.1038/npp.2010.151

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106–113. https://doi.org/10.1016/j.tics.2011.12.010

Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q., & Frank, M. J. (2013). Frontal theta overrides pavlovian learning biases. *The Journal of Neuroscience*, *33*(19), 8541–8548. https://doi.org/10.1523/JNEUROSCI.5754-12.2013

Daniel-Watanabe, L., McLaughlin, M., Gormley, S., & Robinson, O. J. (2022). Association between a directly translated cognitive measure of negative bias and self-reported psychiatric symptoms. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *7*(2), 201–209. https://doi.org/10.1016/j.bpsc.2020.02.010

Dayan, P., & Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron*, *36*(2), 285–298. https://doi.org/10.1016/S0896-6273(02)00963-7

Dayan, P., Niv, Y., Seymour, B., & Daw, N. D. (2006). The misbehavior of value and the discipline of the will. *Neural Networks*, *19*(8), 1153–1160. https://doi.org/10.1016/j.neunet.2006.03.002

Dickinson, A., & Balleine, B. (2002). The role of learning in the oepration of motivational systems. In H. Pashler & R. Gallistel (Eds.), *Steven's handbook of experimental psychology: Learning, motivation, and emotion* (pp. 497–533). Wiley.

Dorfman, H. M., & Gershman, S. J. (2019). Controllability governs the balance between Pavlovian and instrumental action selection. *Nature Communications*, *10*(1), Article 5826. https://doi.org/10.1038/s41467-019-13737-7

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, *42*(1), 351–362. https://doi.org/10.3758/BRM.42.1.351

Ereira, S., Pujol, M., Guitart-Masip, M., Dolan, R. J., & Kurth-Nelson, Z. (2021). Overcoming Pavlovian bias in semantic space. *Scientific Reports*, *11*(1), Article 3416. https://doi.org/10.1038/s41598-021-82889-8

Everitt, B. J., & Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: From actions to habits to compulsion. *Nature Neuroscience*, *8*(11), 1481–1489. https://doi.org/10.1038/nn1579

Gershman, S. J., Guitart-Masip, M., & Cavanagh, J. F. (2021). Neural signatures of arbitration between Pavlovian and instrumental action selection. *PLOS Computational Biology*, *17*(2), Article e1008553. https://doi.org/10.1371/journal.pcbi.1008553

Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision making. *Trends in Cognitive Sciences*, *18*(4), 194–202. https://doi.org/10.1016/j.tics.2014.01.003

Guitart-Masip, M., Fuentemilla, L., Bach, D. R., Huys, Q. J., Dayan, P., Dolan, R. J., & Duzel, E. (2011). Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *The Journal of Neuroscience*, *31*(21), 7867–7875. https://doi.org/10.1523/JNEUROSCI.6376-10.2011

Guitart-Masip, M., Huys, Q. J., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. *NeuroImage*, *62*(1), 154–166. https://doi.org/10.1016/j.neuroimage.2012.04.024

Hershberger, W. A. (1986). An approach through the looking glass. *Animal Learning & Behavior*, *14*(4), 443–451. https://doi.org/10.3758/BF03200092

Mathôt, S., Siebold, A., Donk, M., & Vitu, F. (2015). *Large pupils predict goal-driven eye movements*. Retrieved October 20, 2025, from https://github.com/smathot/materials_for_P0010.5

Mkrtchian, A., Aylward, J., Dayan, P., Roiser, J. P., & Robinson, O. J. (2017). Modeling avoidance in mood and anxiety disorders using reinforcement learning. *Biological Psychiatry*, *82*(7), 532–539. https://doi.org/10.1016/j.biopsych.2017.01.017

Nord, C. L., Lawson, R. P., Huys, Q. J. M., Pilling, S., & Roiser, J. P. (2018). Depression is associated with enhanced aversive Pavlovian control over instrumental behaviour. *Scientific Reports*, *8*(1), Article 12582. https://doi.org/10.1038/s41598-018-30828-5

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II* (pp. 64–99). Appleton-Century-Crofts.

Robinson, O. J., Vytal, K., Cornwell, B. R., & Grillon, C. (2013). The impact of anxiety upon cognition: Perspectives from human threat of shock studies. *Frontiers in Human Neuroscience*, *7*, Article 203. https://doi.org/10.3389/fnhum.2013.00203

Rutledge, R. B., Smittenaar, P., Zeidman, P., Brown, H. R., Adams, R. A., Lindenberger, U., Dayan, P., & Dolan, R. J. (2016). Risk taking for potential reward decreases across the lifespan. *Current Biology*, *26*(12), 1634–1639. https://doi.org/10.1016/j.cub.2016.05.017

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Schurr, R., Reznik, D., Hillman, H., Bhui, R., & Gershman, S. (2024). Dynamic computational phenotyping of human cognition. *Nature Human Behaviour*, *8*, 917–931. https://doi.org/10.1038/s41562-024-01814-x

Spielberger, C. D., Gorsuch, R. L., Lushene, R., & Vagg, P. R. J. G. (1983). *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press. https://doi.org/10.1111/j.1758-0854.2009.01017.x

Stan Development Team. (2023). *Stan modelling language users guide and reference manual*. https://mc-stan.org

Sutton, R. (1988). Learning to predict by the methods of temporal differences. *Machine Learning*, *3*(1), 9–44. https://doi.org/10.1023/A:1022633531479

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely application information criterion in singular learning theory. *Journal of Machine Learning Research*, *11*, 3571–3594.

Williams, D. R., & Williams, H. (1969). Auto-maintenance in the pigeon: Sustained pecking despite contingent non-reinforcement. *Journal of the Experimental Analysis of Behavior*, *12*(4), 511–520. https://doi.org/10.1901/jeab.1969.12-511

Received February 26, 2025
Revision received October 1, 2025
Accepted October 8, 2025 ■