Pavlovian Control of Escape and Avoidance

Alexander J. Millner[1], Samuel J. Gershman[1], Matthew K. Nock[1],
and Hanneke E. M. den Ouden[2]

Abstract

■ To survive in complex environments, animals need to have mechanisms to select effective actions quickly, with minimal computational costs. As perhaps the computationally most parsimonious of these systems, Pavlovian control accomplishes this by hardwiring specific stereotyped responses to certain classes of stimuli. It is well documented that appetitive cues initiate a Pavlovian bias toward vigorous approach; however, Pavlovian responses to aversive stimuli are less well understood. Gaining a deeper understanding of aversive Pavlovian responses, such as active avoidance, is important given the critical role these behaviors play in several psychiatric conditions. The goal of the current study was to establish a behavioral and computational framework to examine aversive Pavlovian responses (activation vs. inhibition) depending on the proximity of an aversive state (escape vs. avoidance). We introduce a novel task in which participants are exposed to primary aversive (noise) stimuli and characterized behavior using a novel generative computational model. This model combines reinforcement learning and drift-diffusion models so as to capture effects of invigoration/inhibition in both explicit choice behavior as well as changes in RT. Choice and RT results both suggest that escape is associated with a bias for vigorous action, whereas avoidance is associated with behavioral inhibition. These results lay a foundation for future work that promise to provide insights into typical and atypical aversive Pavlovian responses involved in psychiatric disorders, allowing us to quantify both implicit and explicit indices of vigorous choice behavior in the context of aversion. ■

INTRODUCTION

To survive in complex environments, animals must select actions that result in beneficial outcomes, such as obtaining food or escaping a predator. Chances of survival are vastly improved when a decision mechanism is available that selects actions quickly, with low computational costs and minimal errors that might result in disadvantageous outcomes, like loss of food or death. Fortunately, animals are endowed with such a decision mechanism, known as Pavlovian control. Pavlovian control provides rapid, computationally efficient action selection by hard-wiring certain actions to sets of stimuli. Although Pavlovian responses are generally advantageous, they are relatively inflexible because they are automatically emitted in the presence of associated stimuli, regardless of whether the response is appropriate to the current situation (Hershberger, 1986; Breland & Breland, 1961). In contrast, a second decision controller, known as instrumental control, learns advantageous responses based on outcomes that followed prior responses. By using trial-and-error learning, instrumental control permits flexible adaptation to specific environments and thus maximizes expected outcomes over the long run. However, it is slower and requires more computational resources than Pavlovian control.

Pavlovian responses differ depending on the valence of the stimulus or environmental context, promoting approach behavior toward reward-predictive stimuli and avoidance behavior toward punishment-predictive stimuli (Boureau & Dayan, 2011; Huys et al., 2011; Glickman & Schiff, 1967). Pavlovian control often results in advantageous responses because these responses typically are aligned to the statistical probabilities of outcomes in a given environment (Lloyd & Dayan, 2016; Dayan & Huys, 2009). For example, in general, an approach response to rewards and an avoid response to punishments are more likely to result in beneficial outcomes than vice versa. Advantageous Pavlovian responses can also end up assisting instrumental control. For example, Pavlovian control initiates an active approach response in the presence of food, which helps instrumental control learn the precise response to obtain the food (Dayan, Niv, Seymour, & Daw, 2006). In such circumstances, the two controllers work together to make learning fast and efficient. However, the Pavlovian controller rigidly specifies behaviors regardless of their outcomes, whereas instrumental control adapts behaviors to maximize positive outcomes (Dickinson & Balleine, 1994).

Experimental paradigms can capitalize on this differential outcome sensitivity between Pavlovian and instrumental controllers and expose Pavlovian tendencies by creating conditions in which Pavlovian and instrumental preferences conflict (Swart et al., 2017; Cavanagh,

[1]Harvard University, Cambridge, MA, [2]Radboud University Nijmegen

Eisenberg, Guitart-Masip, Huys, & Frank, 2013; Geurts, Huys, den Ouden, & Cools, 2013; Guitart-Masip et al., 2012; Huys et al., 2011; Crockett, Clark, & Robbins, 2009; Dayan et al., 2006). Experiments using such paradigms have revealed quite consistently that appetitive Pavlovian response tendencies promote active responses and degrade performance when obtaining rewards requires passive responses (Cavanagh et al., 2013; Guitart-Masip et al., 2012). Further support for the pairing of appetitive contexts and active responses is that dopamine released in mesolimbic and nigrostriatal pathways when rewards are larger than expected also invigorates motor actions through the dopamine-modulated direct "go" pathways while inhibiting striatal indirect "no-go" pathways (Lloyd & Dayan, 2016; Beierholm et al., 2013; Kravitz, Tye, & Kreitzer, 2012; Frank, 2005). Thus, dopamine responses during appetitive contexts facilitate motor activity.

Although prior research has elucidated mechanisms of appetitive responses, the corresponding responses to aversive stimuli are less understood (Dayan & Huys, 2015). Aversive stimuli have been associated with inhibitory responses (McNaughton & Corr, 2004; Graeff, Netto, & Zangrossi, 1998), and overall, aversive contexts are associated with fewer active responses and more passive responses, compared with rewarding (or neutral) contexts (Cavanagh et al., 2013; Geurts et al., 2013; Guitart-Masip et al., 2012, 2013). Further support comes from studies showing that serotonin is involved in both processing aversive events and behavioral inhibition (Graeff & Silveira Filho, 1978). For example, reducing serotonin levels is associated with reduced inhibitory responses to aversive outcomes (Crockett, Clark, Apergis-Schoute, Morein-Zamir, & Robbins, 2012; Crockett et al., 2009). Thus, Pavlovian responses to appetitive versus aversive contexts may reflect two dissociable systems subserved by different, potentially opponent, neurotransmitters (Boureau & Dayan, 2011; Cools, Nakamura, & Daw, 2011; Daw, Kakade, & Dayan, 2002).

However, the exact nature of aversive Pavlovian responses is complicated by several factors. First, the valence (appetitive vs. aversive)-by-response (active vs. passive) interaction reflecting a Pavlovian bias found in prior studies is driven by a large bias for active responses to reward and not necessarily by an inhibitory bias for aversive stimuli. Within only aversive trials, participants do not show differences in the ability to learn active versus passive responses (Cavanagh et al., 2013; Geurts et al., 2013; Guitart-Masip et al., 2012, 2013). Thus, aversive contexts are not as clearly associated with passivity as appetitive contexts are associated with active responses.

Second, there are situations in which animals show active and vigorous responses to aversive contexts. For example, responses to natural aversive stimuli, such as predator-related stimuli, depend on proximity: Distal threats are associated with the inhibition of action, but proximal threats elicit defensive responses to escape,

such as fighting or fleeing (Deakin & Graeff, 1991; Blanchard & Blanchard, 1988). Animals also exhibit vigorous active responses to achieve safety from experimentally induced ongoing aversive states, such as a floor with a continuous electric shock (Mellgren, Nation, & Wrather, 1975). Furthermore, recent research showed that participants will respond faster to larger punishments when they are informed of the extent of the punishment in advance (Griffiths & Beierholm, 2017). Thus, aversive Pavlovian responses involve both active and passive responses depending on circumstances.

Third, the neuromodulatory systems supporting active responses to aversive stimuli are not well understood. Deakin and Graeff (1991) hypothesized that two serotonin pathways were associated with active versus passive defensive responses, and subsequent research has largely supported their theory (Deakin, 2013; Paul & Lowry, 2013). Thus, this line of research implicates serotonin systems as playing a pivotal role in execution of both active and inhibitory behaviors to aversive stimuli. It also suggests that neuromodulatory systems involved in active responses to aversive contexts are unique from those involved with active responses to appetitive stimuli.

Another, although not mutually exclusive, theory posits that active responses to both appetitive and aversive systems partially rely on overlapping mechanisms. Here, dopamine release signals safety from an aversive context and motivates an active response by modulating direct and indirect pathways, similar to appetitive contexts (Lloyd & Dayan, 2016). This is supported by research showing mesolimbic dopamine release during aversive situations, such as physical pain (Navratilova & Porreca, 2014). In general, understanding the neuromodulatory systems involved in aversive and appetitive responses will help inform how these different contexts influence behavior.

Furthering the understanding of aversive Pavlovian control has the potential to provide insights into mechanisms of several psychiatric disorders. Many psychiatric conditions are characterized by both ongoing aversive psychological states from which people seek relief (i.e., escape; Yager, 2015) and problematic avoidance behaviors. For example, behaviors such as nonsuicidal and suicidal self-harm behaviors can be viewed as maladaptive coping behaviors meant to gain relief from ongoing aversive psychological states (Nock, 2009), whereas phobias as well as social and generalized anxiety can be characterized as exaggerated avoidance responses (Shin & Liberzon, 2010). Increasing our understanding of aversive Pavlovian responses could provide important insight into these forms of psychopathology.

To better understand Pavlovian responses to aversive stimuli, we adapted a previously used reinforcement learning (RL) paradigm (Guitart-Masip et al., 2012) that consisted of conditions in which instrumental and Pavlovian processes either promote congruent (e.g., active responses to obtain a reward) or incongruent (e.g.,

passive responses to obtain rewards) behaviors. The current task has only aversive stimuli and uses active (go) and passive (no-go) responses to test whether there are different Pavlovian effects on action selection within two aversive conditions: an escape condition, in which participants learn a response to escape an ongoing aversive state, and an avoid condition, in which participants learn a response to avoid an impending aversive state. We hypothesized that the escape condition would be associated with a Pavlovian bias for an active response, whereas the avoid condition would be associated with a passive (inhibitory) Pavlovian bias. Furthermore, we predicted that active responses to escape would lead to more vigorous responses (as demonstrated by faster RTs) compared with avoiding an impending punishment.

METHODS

Participants

Fifty-three participants completed the study. Fifty-two participants were analyzed, as one participant was excluded for selecting go on every escape trial. Among the remaining participants, ages ranged from 18 to 65 years ($M = 28.7$ years, $SD = 11.8$ years), with 26 women. Nearly half (44%) of the participants were of European ancestry ($n = 23$), whereas 10% were of African ancestry

($n = 5$), 29% were of Asian ancestry ($n = 15$), and the remaining 17% of the participants were of mixed races ($n = 9$). Participants were recruited from the Harvard University Psychology Study Pool and were either compensated with course credit or paid $12. The Harvard University Institutional Review Board approved the study.

Experimental Paradigm

The paradigm (Figure 1) is adapted from a similar paradigm by Guitart-Masip and colleagues (2012). In the current paradigm, on every trial, participants were presented with one of four cues (fractal images), followed by either an aversive sound ("escape" condition) or silence ("avoid" condition). The participants' goal was to learn which response (press a button: "go," withhold a button press: "no-go") that more frequently resulted in silence during feedback. As noted above, in the escape condition, the onset of the cue coincided with the onset of the aversive sound. Here, participants had to learn the response (go or no-go) that turned off the aversive sound. In the avoid condition, there was no sound during the cue presentation, and participants had to learn the response that avoided the aversive sound turning on during feedback. The two required responses (go, no-go) and two conditions (escape, avoid) that affected whether
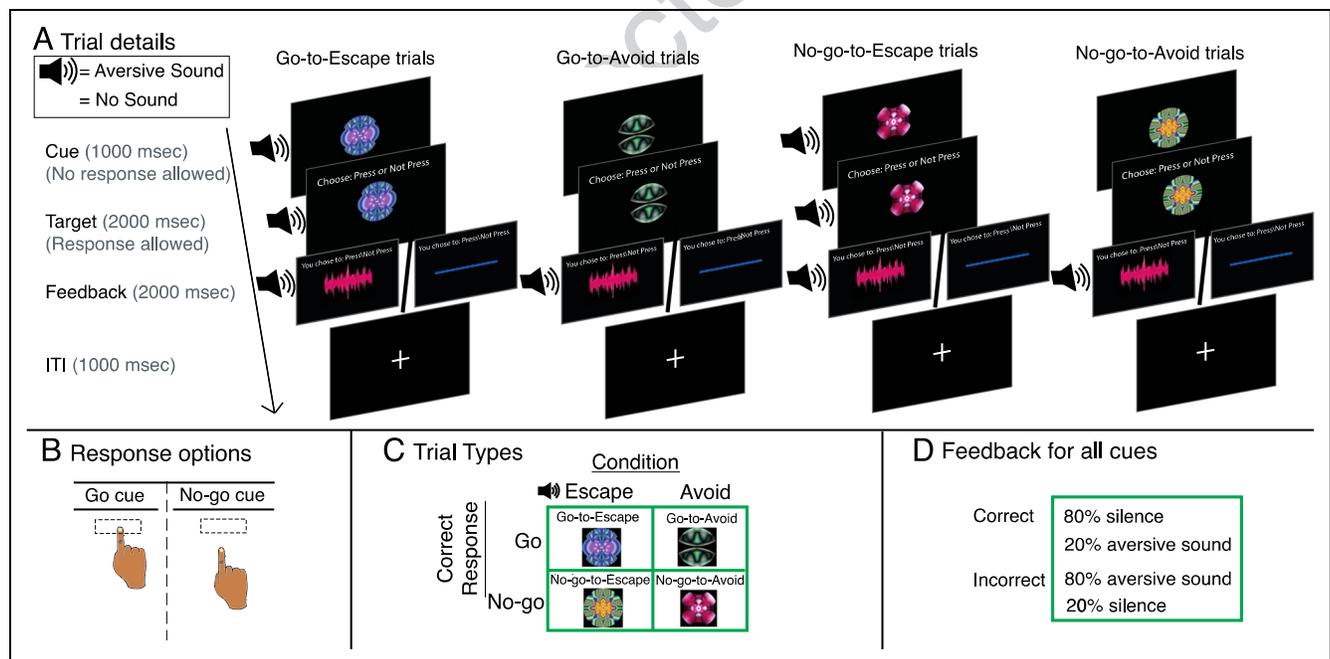

Figure 1. Experimental paradigm. (A) On each trial, one of four fractal images was presented. Participants had to learn, for each cue, whether pressing a button (i.e., go) or withholding a button press (i.e., no-go) resulted in silence, rather than an aversive sound, during feedback. For all trials, participants were presented with a cue for 1 sec where they were unable make a response, followed by a 2-sec target where they could choose to go or no-go and followed by 2 sec of feedback that consisted of an aversive sound or silence, followed by a 1-sec intertrial stimulus. During escape trials, the aversive sound played during the cue and target, whereas during avoid trials, there was no sound during the cue and target. Participants were informed of the onset of the target when the words "Choose: Press or Not Press" appeared on the screen. (B) Response options. On every trial, participants could choose to either go or no-go. (C) Within each condition (i.e., escape and avoid), there was one cue where go was the correct response and one cue where no-go was the correct response. (D) Feedback was probabilistic such that a correct response resulted in silence 80% of the time and the aversive sound 20% of the time and vice versa for an incorrect response. ITI = intertrial interval.

the sound was played during the cue and target presentation resulted in a 2 × 2 design with the following four conditions: go-to-avoid, go-to-escape, no-go-to-avoid, and no-go-to-escape (Figure 1). For each of the four fractal cues, participants had to learn the response that most likely resulted in silence during feedback. Following the 2 × 2 design, each cue was associated with one required response (go or no-go). The feedback was probabilistic, such that a required response resulted in silence during the feedback phase 80% of the time, whereas the other response resulted in silence 20% of the time.

Aversive stimuli consisted of a fork scraping on slate altered with a high-frequency sound and presented over headphones at 80–85 dB. In pilot studies, this sound and volume induced sufficient distress (average subjective distress rating of 7.1/10) without causing lasting effects on participants, such as ringing ears. To increase aversion and decrease habituation, two different clips of the sound were played simultaneously.

Additional Measures

For transparency, we disclose all measures administered and analyses conducted in this study. Some measures in the current study were administered to correspond with a separate study that included a clinical sample. Given that these measures were collected for a purpose other than the main goal of this study, we did not analyze the data collected from them. The measures include self-reported rumination and cognitive flexibility as well as income level, educational level, and employment status.

Data Analysis

Behavioral data were first analyzed using generalized linear mixed-effects regression (GLMER) models with the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015). To test whether accuracy to go and no-go required responses varied as a function of Condition (escape/avoid), we ran a logistic GLMER. The correct choice was defined as the required response (go/no-go) that resulted in the higher probability of silence during feedback. All GLMERs had trial accuracy (i.e., 0 = *incorrect choice*, 1 = *correct choice*) as the dependent variable. For every model, within-participant factors were added as a random factor for the intercept and slopes for all fixed factors and interactions (i.e., maximal models; Barr, 2013). We tested four GLMERs, incrementally adding more regressors (Table 1). To assess whether the addition of a new factor resulted in an improved model (i.e., model comparison), we used a likelihood ratio test (De Boeck et al., 2011), implemented using the *anova* function in R. Additional regressors were determined to have improved the model enough to warrant their inclusion if the *p* value for likelihood ratio test was <.05.

The first (null) model included only the dependent variable and an intercept as a fixed factor in a random

Table 1. Model Comparison of Logistic Linear Mixed-effects Models for Accuracy and RT

	χ^2	df	p
Accuracy			
M1 (null)			
M2 condition (escape, avoid)	77.0	3	1.39E−16
M3 Condition + Response (go, no-go)	104.5	4	1.11E−21
M4 Condition × Response	239.6	5	9.34E−50
RT			
M1 (null)			
M2 condition (escape, avoid)	179.7	3	1.03E−38
M3 Condition + Response (go, no-go)	182.6	4	2.09E−38
M4 Condition × Response	14.0	5	.0155

intercept model. For the subsequent models, we then added incrementally Condition (escape/avoid, M2), Response (go/no-go, M3), and the Condition × Response interaction (M4). After model comparison, coefficient confidence intervals for the winning model were calculated using the lsmeans R package (Lenth, 2016). Simple effects within significant interactions were computed using the phia package in R, which uses Wald χ^2 to determine their significance (Martinez, 2015). Multiple comparisons were corrected by the Holm–Bonferroni sequential procedure.

We followed the same approach to analyze RT, except that we used a gamma GLMER with an identity link function and RT on each trial was entered as the dependent variable. Like the accuracy models, all models contained random intercepts and random slopes for each participant. For both accuracy and RT, statistical significance was set at .05 with a two-tailed test.

Computational Model

We hypothesized that Pavlovian responses for escaping an ongoing aversive stimulus would be associated with a bias for choosing active responses and increased vigor as assessed by faster RTs, whereas the opposite would be the case when avoiding an impending aversive stimulus. Although linear mixed models could provide support for the hypotheses, choice and RT data are analyzed independently and we hypothesized that they are driven by the same Pavlovian bias. Therefore, we sought to use a computational model to identify latent processes, including a Pavlovian bias parameter, that can capture the behavioral effects of both choice behavior and RT.

Many computational models of RL focus on solely modeling response choice. Thus, these models operationalize how the values of each response (i.e., choice) are

updated over the course of learning and how those updated values are translated into subsequent responses. By providing an updated value for each response through prediction errors on a trial-by-trial basis, these models can specify a probability distribution of responses on each trial (Niv & Montague, 2008). Prior studies using similar paradigms to the current study have added a Pavlovian model parameter to demonstrate value-based response biases, and these models have provided higher model evidence than models without a Pavlovian parameter (Cavanagh et al., 2013; Guitart-Masip et al., 2012). One important difference between the task used in these prior studies and the one in the current study is that, in these prior studies, participants had to learn which cues were associated with reward and punishment over the course of the task, whereas in the current task, like in Swart et al. (2017), the condition (escape/avoid) was known at the cue onset because of the presence or absence of the aversive sound. Thus, on the basis of these task differences, we followed Swart et al. (2017) and implemented a static Pavlovian bias parameter, in contrast to these prior studies, which modeled a Pavlovian parameter that dynamically updated over the task.

In a tradition separate from RL models, drift-diffusion models (DDMs) have been used to model RT data successfully across a variety of two-choice RT tasks, such as visual discrimination and memory tasks (Ratcliff, Smith, Brown, & McKoon, 2016). A standard two-alternative DDM consists of a decision variable that evolves over time according to two components: a deterministic linear component whose slope is given by a drift rate parameter and a Gaussian noise component that causes the decision variable to diffuse over time. The decision variable begins its trajectory at a starting point and evolves stochastically until it reaches one of two decision boundaries at which point a response is made (see Figure 2A).

Recent work has integrated RL models and DDMs so that a single generative model can specify a joint distribution of responses and RT (Frank et al., 2015; Milosavljevic, Malmaud, Huth, Koch, & Rangel, 2010). These integrated models use an RL model to track the value of the response options on a trial-by-trial basis, and then these RL values are passed to a DDM, where the drift rate is parametrized as a linear function of the value (i.e., the difference in value between the two response options determines the drift rate). Thus, by using the DDM to define the mapping from values to actions (i.e., policy), we can model the dynamics of choice and RT over the course of learning (Figure 2B; Pedersen, Frank, & Biele, 2017).

Two aspects of the models in the study warrant additional comments. First, because no-go choices are, by definition, the absence of a response and there is no index for the timing of the decision, we modeled the no-go option using an implicit decision boundary, consistent with prior work supporting this assumption (Ratcliff, Huang-Pollock, & McKoon, 2016; Gomez, Ratcliff, & Perea, 2007). Thus, the model is fit to RTs and choice probabilities for go choices but only choice probabilities for no-go choices (i.e., when no-go was selected). Second, we included a go bias, which captures individual variability in the overall tendency to make a go response, thus better explaining data.

We aimed to achieve two main goals using the computational modeling. First, we verified whether including a Pavlovian response parameter captured the behavioral results better than a model without such a parameter. Second, we contrasted two mechanisms by which a Pavlovian bias affects decisions. In the first mechanism, the Pavlovian bias was modeled by allowing the starting points to vary among the escape/avoid conditions. This parameterization affects choice and RT by allowing the condition to push one response option to a starting point closer to the decision boundary, therefore requiring less "evidence" of a value signal to select that response. Alternatively, in our second mechanism, the Pavlovian bias was modeled by allowing the drift rates to vary among escape/avoid conditions. This parameterization affects choice and RT by allowing each condition to amplify the value difference between go and no-go differently. A priori, we hypothesized the first mechanism to be more likely than the second, as studies on biases in DDMs for perceptual decision-making tasks have found that changes to the starting point represent a response bias (e.g., one response is more likely to be correct), whereas changes to the drift rate represent a stimulus discrimination bias (e.g., one stimulus is easier to detect; White & Poldrack, 2014).

As noted previously, in both models, the Pavlovian parameter is modeled as a static bias that, unlike prior studies (Cavanagh et al., 2013; Guitart-Masip et al., 2012), does not dynamically update over time. This is because the condition (escape or avoid) was known at cue onset because of the presence or absence of the aversive sound.

Each model we tested used the following implementation to integrate RL models and DDMs. Instrumental Q values were updated on each trial using a simple delta rule (Rescorla & Wagner, 1972):

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha[r_t - Q_t(s_t, a_t)] \qquad (1)$$

where α is a learning rate, s_t is the stimulus, r_t is the reward, and a_t is the action (go or no-go) on trial t. Then, to translate these Q values into actions and RTs, we used the following DDM specification. Q values determined the drift rate μ_t on trial t:

$$\mu_t = \beta_0 + \beta_1[Q_t(s_t, \text{go}) - Q_t(s_t, \text{no-go})]$$

where β_0 captures a constant go bias, β_1 captures a go bias shared across responses and s_t is the cue on trial t and $Q_t(s_t, \text{go}) - Q_t(s_t, \text{no-go})$ represents the difference between Q values for go and no-go. After a nondecision time T, the drift-diffusion process starts at z, which varies between 0 and ω (the boundary separation parameter), and then proceeds until a bound (0 or ω) is reached.

Figure 2. (A) A value-based DDM schematic. Choices and RTs are modeled in a DDM as the combination of (1) the starting point (w), where the decision process begins (the best fitting model in the current study had two separate starting points [w_{escape}, w_{avoid}] for the two conditions); (2) the drift rate (modeled as a linear function of value [see below] as well as a constant go bias [β_0] and a go bias shared [β_1] across the two possible responses [not shown]), which guides the trajectory of the decision process; (3) a nondecision time (T), where the stimulus is still being processed; and (4) the boundary separation (ω), which represents caution (more caution will lead to longer RTs). Thus, the stimulus is presented and processed, the decision processes start and are guided by the drift rate, and a choice is selected once the processes reach one boundary. (B) Example of six trials for one of the cues with model-based value calculations and DDMs. On the basis of feedback, values for each response (go, no-go) on a given cue are updated on a trial-by-trial basis. For example, on Trial 1, a "go" response is followed by silence, which

means that the value for "go" is increased. The "no-go" value is not updated as there was no no-go response. On Trial 2, a "no-go" response is followed by the aversive sound, so the value of the "no-go" response is decreased. Again, the value of the "go" response remains unchanged. On each trial, the DDM drift rate is modeled using the difference in value between the two responses. Early in the task, the difference in value between the two responses is small, resulting in a smaller drift rate and longer RTs to come to a decision. As the value difference increases over the course of the task, the drift rate increases, which leads to faster RTs. In the example shown, the starting point is closer to the go decision boundary, which was the case in the escape condition. In the avoid condition, the results of model fitting suggest that the starting point was closer to the no-go decision boundary, as illustrated.
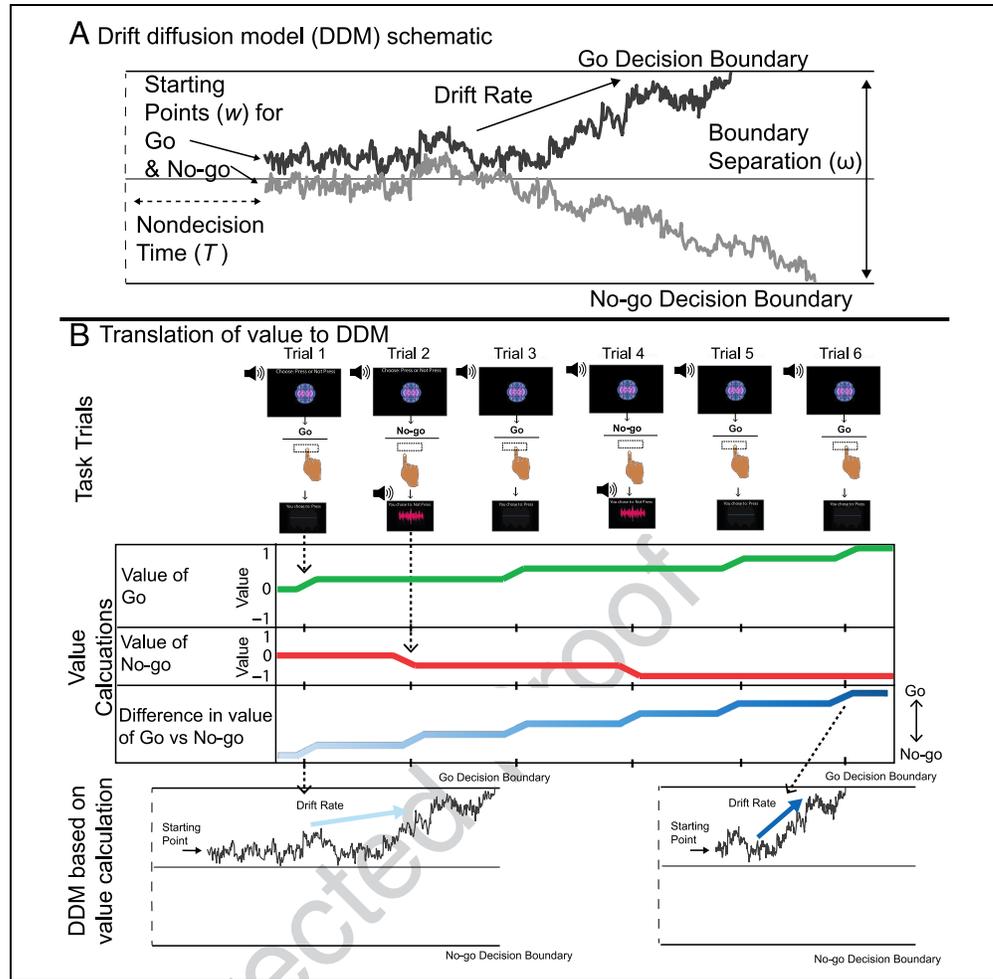
Following Navarro and Fuss (2009), we used a relative parameterization of the starting point, $w = z/\omega$, that varies between 0 and 1. The Wiener first passage time density defines the joint likelihood for the choice and RT on each trial induced by the DDM (i.e., the distribution over the time at which one of the decision boundaries is crossed).

For the first parameterizations of the Pavlovian bias (Model 1), we fit separate starting points (w) within the DDM for the two conditions: w_{escape} on escape trials and w_{avoid} on avoid trials. In Model 2, the conditions shared the same starting point but varied in their drift rate:

M2:

$$\mu_t = \beta_{escape} + \beta_1[Q_t(s_t, \text{go}) - Q_t(s_t, \text{no-go})]$$

for escape trials and

$$\mu_t = \beta_{avoid} + \beta_1[Q_t(s_t, \text{go}) - Q_t(s_t, \text{no-go})]$$

for avoid trials (see above for how these parameterizations affect decision-making). To be clear, the starting point was set as a free parameter in both models, but in Model 1, it consisted of two separate free parameters for avoid and escape (across go and no-go) trials, and in Model 2, it was a single free parameter fit across conditions. We used the following stepwise procedure for the computational modeling approach.

Step 1: Model Fitting

We fit the model parameters to data from each participant individually using maximum likelihood estimation with the fast approximation of the Wiener first passage time density derived by Navarro and Fuss (2009). The Wiener first passage time density defines the joint likelihood for the choice and RT on each trial induced by the

DDM (i.e., the distribution over time at which one of the decision boundaries is crossed). We used the following parameter constraints: learning rate (α): [0, 1]; non-decision time (T): [−0.2, 0.5]; boundary separation (ω): [0.001, 20]; starting point (w): [0.001, 0.999]; and drift rate coefficients (β): [−20, 20].

Step 2: Capturing Qualitative Effects

Our principal approach to evaluating the models was to assess whether models captured the qualitative pattern of results across both accuracy (e.g., higher accuracy for go-to-escape and no-go-to-avoid) and RT (e.g., faster RT for go-to-escape than go-to-avoid) after we computed the expected choice probability and mean RT on each trial.

Step 3: Quantitative Model Comparison

We also used random-effects Bayesian model selection (Rigoux, Stephan, Friston, & Daunizeau, 2014) to compare models, reporting protected exceedance probabilities (the probability that a particular model is more frequent in the population than any other model under consideration, while accounting for the probability that none of the models explains the data).

Step 4: Parameter Evaluation

Finally, we also compared parameter differences (starting point or drift rate, depending on the model) between conditions using a paired t test. For summary statistics, we computed bootstrapped 95% confidence intervals across participants.

RESULTS

Behavioral Results

Accuracy

Participants displayed high overall accuracy (overall: M = 90.5%, SD = 29.3%; condition: Ms = 93.1–86.7%) on the task, showing that they learned the required responses. We tested the effect of Condition (escape/avoid) and Response (go/no-go) on task accuracy in four different logistic GLMERs, where we started with a null model with only an intercept and then incrementally added fixed factors and then an interaction term. In each case, including additional factors improved model fit (Table 1), and adding the Condition × Response interaction term significantly improved model fit (χ^2 = 239.6, p = 9.34E−50) over a model with the two main effects of Condition and Response alone.

For the winning model, Wald χ^2 tests showed a significant Condition × Response interaction (b = −1.17, 95% CI [−1.88, −0.46], p < .001). When breaking this down into simple effects by Condition, we observed that, as

hypothesized, participants were more likely to go-to-escape than no-go-to-escape, resulting in higher accuracy for go-to-escape (M = 91.9%, SD = 10.7%) than no-go-to-escape (M = 86.7%, SD = 15.7%; χ^2 = 13.9, p < .001). Conversely, there was a marginally significant effect that participants were more likely to no-go-to-avoid than go-to-avoid (χ^2 = 3.6, p = .058), resulting in higher accuracy to no-go-to-avoid (M = 93.1%, SD = 12.9%) than go-to-avoid (M = 90.2%, SD = 14.9%). When breaking down the interaction by Response, participants showed higher accuracy for no-go-to-avoid than no-go-to-escape (χ^2 = 24.3, p = .000002), whereas there was no difference in go-to-avoid and go-to-escape accuracy (χ^2 = 0.3, p = .61; Figure 3A).

RT

Like accuracy, we analyzed RT with four GLMERs, starting with a null model and incrementally adding fixed factors and an interaction term. The model with the interaction term provided the best fit (Table 1; χ^2 = 14.03, p < .016).

The winning model revealed a significant Condition × Response interaction (b = −0.13, 95% CI [−0.25, −0.01], p = .035) where, although participants had faster RTs for all required responses to escape, rather than avoid, aversive feedback, this difference was larger for no-go cues (no-go-to-escape: M = 775.94, SD = 223.30; no-go-to-avoid: M = 965.35, SD = 310.38) than for go cues (go-to-escape: M = 588.49, SD = 150.21; go-to-avoid: M = 656.62, SD = 190.87). In addition, go-to-escape trials had significantly faster RTs than go-to-avoid trials (χ^2 = 17.17, p < .001), suggesting that go-to-escape induced a more vigorous response (Figure 3B).

Computational Modeling Results

As discussed above, we sought to identify a model in which the expected choice probabilities captured the observed behavioral effects for accuracy and RT. Model M0 that did not include a Pavlovian bias parameter failed to capture most qualitative features of the accuracy and RT results. Therefore, we discarded M0 and compared only M1 and M2, with different parameterizations of the Pavlovian bias.

Qualitatively, the expected choice probabilities for both M1 and M2 captured the average accuracy differences between Condition (i.e., higher accuracy for go-to-escape and higher accuracy for no-go-to-avoid [although this effect was marginally significant in the observed data and significant in the model]; Figure 3A [M2 is not displayed but qualitatively appears very similar]) and most of the average RT differences (i.e., faster RT for go-to-escape than go-to-avoid; Figure 3B) as well as captured accuracy and RT changes over the task (Figure 3C and D). The one exception was for M2, with separate escape and avoid drift rates. The behavioral data
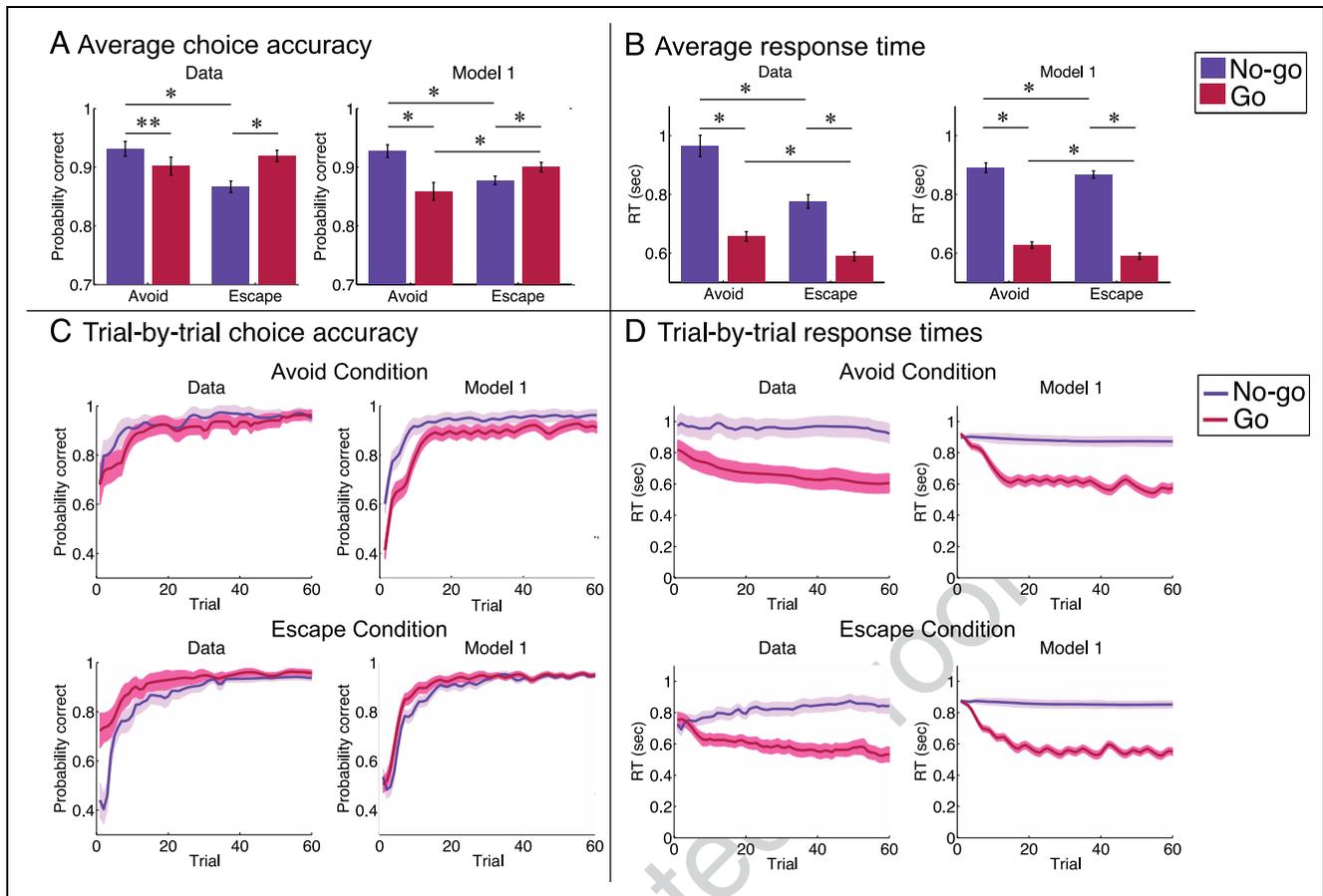
Figure 3. Accuracy and RT results for the empirical data and winning model. (A) Average accuracy and (B) RT for empirical data and model fits (derived from Model 1). Error bars denote *SEM*. Performance is more accurate on no-go trials than on go responses in the avoid condition (indicating Pavlovian inhibition when aversive stimuli are impending); this pattern reverses in the escape condition (indicating Pavlovian activation when aversive stimuli are ongoing). Consistent with this interpretation, responses are also overall slower in the avoid condition. Participants also show overall slower RTs when they erroneously "go" on no-go cues across both escape and avoid conditions. These patterns are captured qualitatively by Model 1, which posits that escape and avoid conditions induce different starting points for the drift-diffusion process. (C) Proportion correct and (D) average RTs for each trial with smoothing based on robust spline smoothing (Garcia, 2010). We chose to not display the results from Model 2 because, qualitatively, the plots for Models 1 and 2 appear very similar. Error bars represent *SEM*. $*p < .05$; $**p = .058$.

showed that RT during no-go-to-escape trials was significantly faster than that during no-go-to-avoid trials; however, M2 failed to capture this difference. M1, with separate starting points for escape and avoid conditions, captured this significant RT difference but with a smaller effect than in the observed result (Figure 3B). Both M1 and M2 also resulted in significantly higher accuracy for go-to-escape compared with go-to-avoid, which was not observed in the empirical data (Figure 3A). Overall, both models captured the qualitative effects well, including Pavlovian influence on both response choice and vigor, but M1 captured all the effects, whereas M2 failed to capture the no-go RT effect.

Consistent with the qualitative comparison, random-effects Bayesian model selection (Rigoux et al., 2014) showed that M1 was favored over M2 (Table 2). However, like the similar qualitative effects, the model comparison did not strongly favor M1 over M2, because both models captured a large amount of variance in the data (Table 2).

Finally, for the favored model, M1, we assessed the fitted parameters. A paired *t* test revealed that the starting point for the escape condition was significantly higher (i.e., biased toward a go response) compared with the starting point for the avoid condition, $t(50) = 4.74$, $p < .0001$. This suggests that, in the escape trials, the presence of the aversive noise pushes the starting point of the decision-making process closer to the go decision boundary, requiring less evidence that the go response has a higher value than the no-go response, which makes a go response more likely and faster. Conversely, the presence of a potential punishment during avoid trials pushes the starting point closer to the no-go decision boundary, making a no-go response more likely.

DISCUSSION

Aversive Pavlovian biases have been implicated in several psychiatric disorders including depression (Huys et al.,

Table 2. Bootstrapped 95% Confidence Intervals for All Model Parameters

		Model 1: Separate Starting Points	Model 2: Separate Drift Rates
Nondecision time	T	[0.043, 0.119]	[0.039, 0.119]
Constant go bias	β_0	[0.233, 0.588]	–
Escape go bias	β_{escape}	–	[0.396, 0.792]
Avoid go bias	β_{avoid}	–	[0.059, 0.431]
Shared go bias	β_1	[3.181, 4.906]	[1.401, 3.493]
Starting point	w	–	[0.303, 0.390]
Escape starting point	w_{escape}	[0.346, 0.435]	–
Avoid starting point	w_{avoid}	[0.285, 0.372]	–
Learning rate	α	[0.155, 0.203]	[0.170, 0.259]
Boundary separation	ω	[1.915, 2.184]	[1.907, 2.150]
Model comparison	BIC	141.8 (238.8)	169.5 (321.3)
	PXP	0.53	0.47

Separate starting points capture variation in response bias, whereas separate drift rates capture variation in action discrimination. BIC = Bayesian information criterion (mean and standard deviation across participants); PXP = protected exceedance probability.

2016), but we have limited understanding of how they operate. Prior work in humans has generally associated aversive Pavlovian biases with behavioral inhibition (Guitart-Masip et al., 2012; Crockett et al., 2009). However, in this study, we report that aversive Pavlovian biases can motivate behavioral activation and inhibition, depending on the specific aversive context. Specifically, Pavlovian response biases promote active responses to escape an ongoing aversive stimulus while promoting passive responses to avoid an impending aversive stimulus. Furthermore, we found that active responses to escape were more vigorous, as measured by faster RTs, compared with active responses to avoid an impending negative outcome, and this was the case regardless of whether the context required an active response or not.

We developed a computational model that embedded a classic RL model within a DDM to capture how Pavlovian biases may drive these changes in both choice and response speed. The model that best captured participants' choice and RT patterns suggested that the Pavlovian influence is best understood as a response bias in which the presence of an ongoing aversive state (i.e., escape) pushed the starting point of the decision process toward an active response (go) and the presence of an impending punishment (i.e., avoid) pushed the starting point toward withholding responding (no-go).

These results replicate and extend prior work examining Pavlovian response biases to valenced stimuli. Specifically, we replicated the finding that punishment-predictive cues were associated with Pavlovian inhibitory no-go responses (Swart et al., 2017; Cavanagh et al., 2013; Guitart-Masip et al., 2012) and extended this by showing a Pavlovian bias that promotes a vigorous go response in

the context of an ongoing aversive stimulus. The current results also add to the significant established evidence that Pavlovian responses interfere with instrumental performance. This was particularly evident for no-go-to-escape trials (i.e., where no-go was the best choice to stop aversive stimulus), as this condition exhibited the poorest accuracy, similar to participants' difficulty learning to withhold responding to obtain a reward in tasks with an appetitive condition (Cavanagh et al., 2013; Guitart-Masip et al., 2012). In line with this choice bias, escape trials also demonstrated more vigorous responses for both go and no-go cues, compared with the corresponding avoid cues. Together, these results suggest that the Pavlovian bias initiated by the aversive sound promoted strong invigoration of action.

There was an important difference between the current study and past studies using similar paradigms (Cavanagh et al., 2013; Guitart-Masip et al., 2012). During the escape condition, the valenced stimulus (i.e., the sound) was present during the cue, whereas prior studies have only had valenced stimuli during feedback. This meant that the escape cues were unconditioned in that they did not require learning to acquire an aversive valence. On the other hand, similar to all conditions in prior studies, the avoid cue was conditioned because it acquired a negative valence by predicting an aversive outcome. The strong Pavlovian activation bias on choice and RT during escape trials may have been due, in part, to the presence of the sound and its unconditioned nature.

Although results showing increased invigoration for aversive trials in an escape context are not necessarily predicted by theories positing a Valence × Action interaction coupling activation to reward and inhibition to

punishment (Boureau & Dayan, 2011; Cools et al., 2011), they are consistent with studies on fear conditioning and on defensive responses to natural predators. For example, similar to escape trials in the current study, rats show active responses to escape an unconditioned stimulus, and similar to avoid trials, they exhibit passive responses to cues that predict an electric shock (Myers, 1971, as cited in Gray & MacNaughton, 2003, p. 51). Similarly, influential work on defensive responses suggest that animals have (species-specific [Bolles, 1970]) responses depending on the threat level (Gray & MacNaughton, 2003; Deakin & Graeff, 1991): More severe, proximal threats (e.g., predator attack) are associated with active responses to escape, whereas less severe, distal threat (e.g., smell of a predator) is associated with action inhibition (see Gray & MacNaughton, 2003).

In the current paradigm, escape cues contained both an unconditioned aversive stimulus (i.e., the sound) and indicated the response required to gain relief from the aversive stimulus. Thus, we could not tease apart whether the observed Pavlovian go bias and increased vigor were the result of a (putatively serotonergic) unconditioned response to the aversive state or were elicited (putatively through dopaminergic pathways) because the escape cue is associated with safety, in line with the two-stage theory of active avoidance (Lloyd & Dayan, 2016; Mowrer, 1956), or a combination of both. In other words, the presence of aversive stimuli could directly drive the observed motor responses and increased vigor in the form of a defensive response. Consistent with this, studies with computer-game predators that emit real electric shocks show that, as a predator attack became likely, people show increased ratings of dread, locomotor errors and activated defense-related brain circuits (Mobbs et al., 2007, 2009).

An alternative possibility, related to the two-stage theory of active avoidance and more recently worked out in more detail by Lloyd and Dayan (2016), is that appetitive and aversive active responses rely on neural systems that encode the relative change from baseline rather than the absolute negative or positive valence of the context (Lloyd & Dayan, 2016; Mowrer, 1956). In this model, shortly after its onset, the aversive state becomes the baseline state from which neutral feedback represents an appetitive outcome (i.e., safety). Achieving safety then results in a positive reward prediction error. Consistent with this, active avoidance (where an animal continually avoids punishment by actively moving away from a punishment-predicting cue; Mowrer, 1956) has been successfully modeled using safety states as akin to obtaining a reward (Maia, 2010). A similar process may lead people informed of the degree of a potential punishment ahead of a choice to show more vigorous responses to more aversive potential punishments (Griffiths & Beierholm, 2017). Along these same lines, the observed vigorous responses to escape in the current study could have also arisen from an instrumental mechanism that selects both

a discrete action and its vigor (Niv, Daw, Joel, & Dayan, 2007), to the extent that vigorous responses that terminated the aversive state were instrumentally reinforced. Given that participants showed a go bias from the first trial, before learning, we argue that the results suggest at least the presence of a Pavlovian mechanism driven by the aversive stimulus, but it is possible that both mechanisms are involved.

Determining the mechanisms underlying the results of the current study is important because it may provide insight into similar questions regarding aversive emotions associated with mental disorders. Although theoretical models have attributed different mental disorders to functional subsystems of the overall defense systems, these models have not been tested extensively in clinical populations (Gray & MacNaughton, 2003). Part of our reasoning for using an aversive sound, rather than using physically painful stimuli, was to study aversive states that approximated aversive psychological states. Physical pain may involve different responses and neurocircuitry, and using it to approximate aversive psychological states is problematic because some psychiatric conditions, such as nonsuicidal self-injury, are characterized by higher physical pain tolerance (Hooley, Ho, Slater, & Lockshin, 2010) and the use of physical pain to regulate negative affective states (Franklin, Aaron, Arthur, Shorkey, & Prinstein, 2012).

Our computational model was similar to prior models used in conjunction with this paradigm (Cavanagh et al., 2013; Guitart-Masip et al., 2012) in that it used a simple RL model and included an overall go bias to predict choice behavior, but it also differed in several important ways. First, the Pavlovian bias parameter was static, unlike prior studies that included a dynamic (i.e., learned) Pavlovian bias, because there was no ambiguity as to the nature of the cue, similar to a previous study (Swart et al., 2017). The fact that a paradigm model with a static parameter was able to recapitulate the behavioral results supports the idea that Pavlovian biases guide behavior by influencing action selection rather than by affecting how action values are learned (Cavanagh et al., 2013; Guitart-Masip et al., 2012). Second, our model integrated RL models and DDMs. Other researchers have used similar models (Pedersen et al., 2017; Frank et al., 2015; Milosavljevic et al., 2010), but this was the first to do so to examine Pavlovian biases.

To identify how Pavlovian biases exert influence, we compared models in which Pavlovian biases started the decision process closer to the promoted action (M1) or amplified the signal of the promoted action (M2). We found that both models captured most of the empirical choice and RT effects. One slight exception was that both M1 and M2 showed a significantly higher choice accuracy for no-go-to-avoid compared with go-to-avoid, whereas this effect was only marginally significant in the empirical data. Although, overall, results for M1 and M2 were very similar, M2 could not capture the relative speeding on

no-go escape trials to no-go avoid trials, whereas M1 was able to capture all of the qualitative behavioral effects and was the quantitatively favored model. This winning model suggests that the observed Pavlovian biases in choice and response speed arise through biased promoting of specific actions given certain contexts. Once the context/state (in this case, escape or avoid) is recognized, the bias is carried out by moving the starting point of the decision-making process closer to the promoted action and therefore requiring less "evidence" of a higher value signal to select that response.

Our results should be considered in the context of the study's limitations. First, in line with previous studies using similar tasks, there was a 1-sec gap between the onset of the cue and the opportunity to respond (Cavanagh et al., 2013; Guitart-Masip et al., 2012). Although the aversive context affected RT, this gap meant that the decision process was not fully indexed by RT. This could have affected which model provided the best fit to the data and our interpretations. Second, during the task instructions, participants received explicit practice with all four trial types and could have inferred the design of the task, which may explain the high accuracy across the conditions. This inference would have increased the precision of beliefs in instrumental values, thereby aiding instrumental control and reducing the impact of conflicting Pavlovian biases.

In conclusion, this study suggests that Pavlovian processes interfere with instrumental learning by promoting action to escape an ongoing aversive stimulus and promoting behavioral inhibition to avoid an impending aversive stimulus. These results add nuance to prior models that argued for valence-dependent action by showing that negatively valenced context can spur action depending on whether the aversive stimulus is present or impending.

Acknowledgments

REFERENCES

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology, 4,* 328.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1–48.

Beierholm, U., Guitart-Masip, M., Economides, M., Chowdhury, R., Düzel, E., Dolan, R., et al. (2013). Dopamine modulates reward-related vigor. *Neuropsychopharmacology, 38,* 1495–1503.

Blanchard, D. C., & Blanchard, R. J. (1988). Ethoexperimental approaches to the biology of emotion. *Annual Review of Psychology, 39,* 43–68.

Bolles, R. C. (1970). Species-specific defense reactions and avoidance learning. *Psychological Review, 77,* 32–48.

Boureau, Y.-L., & Dayan, P. (2011). Opponency revisited: Competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology, 36,* 74–97.

Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist, 16,* 681–684.

Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q. J. M., & Frank, M. J. (2013). Frontal theta overrides Pavlovian learning biases. *Journal of Neuroscience, 33,* 8541–8548.

Cools, R., Nakamura, K., & Daw, N. D. (2011). Serotonin and dopamine: Unifying affective, activational, and decision functions. *Neuropsychopharmacology, 36,* 98–113.

Crockett, M. J., Clark, L., Apergis-Schoute, A. M., Morein-Zamir, S., & Robbins, T. W. (2012). Serotonin modulates the effects of Pavlovian aversive predictions on response vigor. *Neuropsychopharmacology, 37,* 2244–2252.

Crockett, M. J., Clark, L., & Robbins, T. W. (2009). Reconciling the role of serotonin in behavioral inhibition and aversion: Acute tryptophan depletion abolishes punishment-induced inhibition in humans. *Journal of Neuroscience, 29,* 11993–11999.

Daw, N. D., Kakade, S., & Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks, 15,* 603–616.

Dayan, P., & Huys, Q. J. M. (2009). Serotonin in affective control. *Annual Review of Neuroscience, 32,* 95–126.

Dayan, P., & Huys, Q. J. M. (2015). Serotonin's many meanings elude simple theories. *eLife, 4,* e07390.

Dayan, P., Niv, Y., Seymour, B., & Daw, N. D. (2006). The misbehavior of value and the discipline of the will. *Neural Networks, 19,* 1153–1160.

De Boeck, P., Bakker, M., Zwitser, R., Nivard, M., Hofman, A., Tuerlinckx, F., et al. (2011). The estimation of item response models with the lmer function from the lme4 package in R. *Journal of Statistical Software, 39,* 1–28.

Deakin, J. F. W. (2013). The origins of "5-HT and mechanisms of defence" by Deakin and Graeff: A personal perspective. *Journal of Psychopharmacology, 27,* 1084–1089.

Deakin, J. F. W., & Graeff, F. G. (1991). 5-HT and mechanisms of defence. *Journal of Psychopharmacology, 5,* 305–315.

Dickinson, A., & Balleine, B. (1994). Motivational control of goal-directed action. *Animal Learning & Behavior, 22,* 1–18.

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience, 17,* 51–72.

Frank, M. J., Gagne, C., Nyhus, E., Masters, S., Wiecki, T. V., Cavanagh, J. F., et al. (2015). fMRI and EEG predictors of dynamic decision parameters during human reinforcement learning. *Journal of Neuroscience, 35,* 485–494.

Franklin, J. C., Aaron, R. V., Arthur, M. S., Shorkey, S. P., & Prinstein, M. J. (2012). Nonsuicidal self-injury and diminished pain perception: The role of emotion dysregulation. *Comprehensive Psychiatry, 53,* 691–700.

Garcia, D. (2010). Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational Statistics & Data Analysis, 54,* 1167–1178.

Geurts, D. E. M., Huys, Q. J. M., den Ouden, H. E. M., & Cools, R. (2013). Aversive Pavlovian control of instrumental behavior in humans. *Journal of Cognitive Neuroscience, 25,* 1428–1441.

Glickman, S. E., & Schiff, B. B. (1967). A biological theory of reinforcement. *Psychological Review, 74,* 81–109.

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General, 136,* 389–413.

Graeff, F. G., Netto, C. F., & Zangrossi, H., Jr. (1998). The elevated T-maze as an experimental model of anxiety. *Neuroscience and Biobehavioral Reviews, 23,* 237–246.

Graeff, F. G., & Silveira Filho, N. G. (1978). Behavioral inhibition induced by electrical stimulation of the median raphe nucleus of the rat. *Physiology & Behavior, 21,* 477–484.

Gray, J. A., & MacNaughton, N. (2003). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system* (2nd ed.). Oxford: Oxford University Press.

Griffiths, B., & Beierholm, U. R. (2017). Opposing effects of reward and punishment on human vigor. *Scientific Reports, 7,* srep42287.

Guitart-Masip, M., Economides, M., Huys, Q. J. M., Frank, M. J., Chowdhury, R., Duzel, E., et al. (2013). Differential, but not opponent, effects of L-DOPA and citalopram on action learning with reward and punishment. *Psychopharmacology, 231,* 955–966.

Guitart-Masip, M., Huys, Q. J. M., 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,* 154–166.

Hershberger, W. A. (1986). An approach through the looking-glass. *Animal Learning & Behavior, 14,* 443–451.

Hooley, J. M., Ho, D. T., Slater, J., & Lockshin, A. (2010). Pain perception and nonsuicidal self-injury: A laboratory investigation. *Personality Disorders: Theory, Research, and Treatment, 1,* 170–179.

Huys, Q. J. M., Cools, R., Gölzer, M., Friedel, E., Heinz, A., Dolan, R. J., et al. (2011). Disentangling the roles of approach, activation and valence in instrumental and Pavlovian responding. *PLoS Computational Biology, 7,* e1002028.

Huys, Q. J. M., Gölzer, M., Friedel, E., Heinz, A., Cools, R., Dayan, P., et al. (2016). The specificity of Pavlovian regulation is associated with recovery from depression. *Psychological Medicine, 46,* 1027–1035.

Kravitz, A. V., Tye, L. D., & Kreitzer, A. C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature Neuroscience, 15,* 816–818.

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software, 69,* 1–33.

Lloyd, K., & Dayan, P. (2016). Safety out of control: Dopamine and defence. *Behavioral and Brain Functions, 12,* 15.

Maia, T. V. (2010). Two-factor theory, the actor-critic model, and conditioned avoidance. *Learning & Behavior, 38,* 50–67.

Martinez, H. D. R. (2015). Analysing interactions of fitted models. Retrieved from https://cran.r-project.org/web/packages/phia/vignettes/phia.pdf.

McNaughton, N., & Corr, P. J. (2004). A two-dimensional neuropsychology of defense: Fear/anxiety and defensive distance. *Neuroscience and Biobehavioral Reviews, 28,* 285–305.

Mellgren, R. L., Nation, J. R., & Wrather, D. M. (1975). Magnitude of negative reinforcement and resistance to extinction. *Learning and Motivation, 6,* 253–263.

Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., & Rangel, A. (2010). The Drift Diffusion Model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgment and Decision Making, 5,* 437–449.

Mobbs, D., Marchant, J. L., Hassabis, D., Seymour, B., Tan, G., Gray, M., et al. (2009). From threat to fear: The neural organization of defensive fear systems in humans. *Journal of Neuroscience, 29,* 12236–12243.

Mobbs, D., Petrovic, P., Marchant, J. L., Hassabis, D., Weiskopf, N., Seymour, B., et al. (2007). When fear is near: Threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science, 317,* 1079–1083.

Mowrer, O. H. (1956). Two-factor learning theory reconsidered, with special reference to secondary reinforcement and the concept of habit. *Psychological Review, 63,* 114–128.

Navarro, D. J., & Fuss, I. G. (2009). Fast and accurate calculations for first-passage times in Wiener diffusion models. *Journal of Mathematical Psychology, 53,* 222–230.

Navratilova, E., & Porreca, F. (2014). Reward and motivation in pain and pain relief. *Nature Neuroscience, 17,* 1304–1312.

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology, 191,* 507–520.

Niv, Y., & Montague, P. R. (2008). Theoretical and empirical studies of learning. *Neuroeconomics: Decision Making and the Brain,* 329–350.

Nock, M. K. (2009). Why do people hurt themselves? New insights into the nature and functions of self-injury. *Current Directions in Psychological Science, 18,* 78–83.

Paul, E. D., & Lowry, C. A. (2013). Functional topography of serotonergic systems supports the Deakin/Graeff hypothesis of anxiety and affective disorders. *Journal of Psychopharmacology, 27,* 1090–1106.

Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic Bulletin & Review, 24,* 1234–1251.

Ratcliff, R., Huang-Pollock, C., & McKoon, G. (2016). Modeling individual differences in the go/no-go task with a diffusion model. *Decision*.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences, 20,* 260–281.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory, 2,* 64–99.

Rigoux, L., Stephan, K. E., Friston, K. J., & Daunizeau, J. (2014). Bayesian model selection for group studies—Revisited. *Neuroimage, 84,* 971–985.

Shin, L. M., & Liberzon, I. (2010). The neurocircuitry of fear, stress, and anxiety disorders. *Neuropsychopharmacology, 35,* 169–191.

Swart, J. C., Froböse, M. I., Cook, J. L., Geurts, D. E., Frank, M. J., Cools, R., et al. (2017). Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. *eLife, 6,* e22169.

White, C. N., & Poldrack, R. A. (2014). Decomposing bias in different types of simple decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 385–398.

Yager, J. (2015). Addressing patients' psychic pain. *American Journal of Psychiatry, 172,* 939–943.