

High stakes slow responding, but do not help overcome Pavlovian biases in humans

2

3 Johannes Algermissen¹ and Hanneke E. M. den Ouden¹

⁴ ¹Radboud University, Donders Institute for Brain, Cognition, and Behaviour, Nijmegen, The
⁵ Netherlands

6

7 Draft version 1.0, 27/12/2023. This paper has not been peer reviewed.

8

9

10

Author Note

12 Johannes Algermissen  <https://orcid.org/0000-0002-1694-7069>

13 Hanneke E. M. den Ouden <https://orcid.org/0000-0001-7039-5130>

14 Johannes Algermissen is now at the Department of Experimental Psychology, University of Oxford,
15 Oxford, United Kingdom.

Conflict of interest: We have no known conflict of interest to disclose.

17 **Funding:** J. Algermissen was funded by a PhD position from the Donders Centre of Cognition,
18 Faculty of Social Sciences, Radboud University, the Netherlands. Hanneke E.M. den Ouden was
19 supported by a Netherlands Organization for Scientific Research (NWO) VIDI grant 452-17-016.

20 **Acknowledgements:** We thank Theresa Altmeyer, Max Molhuizen, Eva Saragosa, and
21 Madeleine Stahlhacke for assistance with data collection.

22 Correspondence concerning this article should be addressed to Johannes Algermissen or
23 Hanneke E.M. den Ouden, Thomas van Aquinostraat 4, 6526 GD Nijmegen. E-mail:
24 johannes.algermissen@donders.ru.nl and hanneke.denouden@donders.ru.nl.

25

Abstract

26 “Pavlovian” or “motivational” biases are the phenomenon that the valence of prospective outcomes
27 modulates action invigoration: Reward prospect invigorates action, while punishment prospect
28 suppresses it. While effects of the valence of prospective outcomes are well established, it is unclear
29 how the magnitude of outcomes modulates these biases. In this pre-registered study ($N = 55$), we
30 manipulated stake magnitude (high vs. low) in an orthogonalized Motivational Go/ NoGo Task. We
31 tested whether higher stakes (a) strengthen biases or (b) elicit cognitive control recruitment, enhancing
32 the suppression of biases in motivationally incongruent conditions. Confirmatory tests yielded that high
33 stakes slowed down responses independently of the Pavlovian biases, especially in motivationally
34 incongruent conditions, without affecting response selection. Reinforcement-learning drift-diffusion
35 models (RL-DDMs) fit to the data suggested that this effect was best captured by stakes prolonging the
36 non-decision time, but not affecting the response threshold as in typical speed-accuracy tradeoffs. In
37 sum, these results suggest that high stakes result in a slowing-down of the decision process without
38 affecting the expression of Pavlovian biases in behavior. We speculate that this slowing under high
39 stakes might reflect heightened cognitive control, which is however ineffectively used, or reflect
40 positive conditioned suppression, i.e., the suppression of locomotion by high-value immanent rewards,
41 as phenomenon previously observed in rodents that might also exist in humans. Pavlovian biases and
42 slowing under high stakes seem to arise in parallel to each other.

43

44 Key words: Pavlovian biases, motivation, cognitive control, choking, incentive

45

46

47

48

49

50

Introduction

51 The behavior of humans and other animals reflects the interplay of multiple, partly independent
52 decision-making systems (Collins & Cockburn, 2020; Daw, Niv, & Dayan, 2005; Dickinson & Balleine,
53 1994; Metcalfe & Mischel, 1999; Shiffrin & Schneider, 1977; Strack & Deutsch, 2004). One such
54 system is the Pavlovian system which rigidly triggers response invigoration to the prospect of reward
55 and response inhibition to the threat of punishment (Boureau & Dayan, 2011; Dayan, Niv, Seymour, &
56 Daw, 2006; O'Doherty, Cockburn, & Pauli, 2017). Its actions are visible in the form of “Pavlovian” or
57 “motivational” biases, which have been proposed to underlie many seemingly maladaptive behaviors
58 in humans and other animals (Dayan et al., 2006).

59

59 Pavlovian mechanisms might explain seemingly “irrational” behaviors in animals, including
60 the facilitation of instrumental approach behavior by unrelated, but reward-predictive cues (Estes, 1943,
61 1948; LoLordo, McMillan, & Riley, 1974; Lovibond, 1983; Rescorla & Solomon, 1967; Schwartz,
62 1976), or the development of “sign-tracking” behavior, i.e., reward-predictive cues distracting an
63 animal from a focal task (Hearst & Jenkins, 1974; Jenkins & Moore, 1973). Recently, sign-tracking has
64 been suggested to constitute a phenomenon shared across species, including humans (Colaizzi et al.,
65 2020; Garofalo & di Pellegrino, 2015), which might contribute to the etiology and maintenance of drug
66 abuse (Flagel & Robinson, 2017; Flagel, Watson, Robinson, & Akil, 2007). A better understanding of
67 when Pavlovian biases occur and how they interact with other systems regulating behavior promises
68 insights into the development and maintenance of psychiatry conditions such as alcohol or drug abuse
69 (Chen, Garbusow, Sebold, Zech, et al., 2022; Schad et al., 2020).

70

70 There are several proposed accounts for the ecological rationality of Pavlovian biases, i.e.,
71 under which circumstances strong biases are adaptive. Pavlovian control is generally contrasted against
72 instrumental control, i.e. the ability to flexibly adapt behavior to different response-outcome
73 contingencies. There is agreement that Pavlovian control is both faster and cheaper, but at the same
74 time more rigid than instrumental control (Boureau & Dayan, 2011; Dayan et al., 2006). It might thus
75 be particularly adaptive in situations in which instrumental control yields no benefits beyond Pavlovian
76 control, e.g. in novel, unfamiliar, or uncontrollable environments (Dorfman & Gershman, 2019).

77 Another idea is Pavlovian control acts as an “emergency action control system” in high-stakes situations
78 that are critical for survival (O’Doherty et al., 2017), such as when facing a lethal predator, in which it
79 overrides instrumental systems. Yet another idea is that Pavlovian and instrumental control do not
80 compete, but can act in synergy, with instrumental control recruiting Pavlovian control to achieve
81 responses that are faster and more robust to inference (Algermissen & den Ouden, 2023). Hence, besides
82 selecting an appropriate action, strong Pavlovian biases could provide advantages in speed or caution.

83 However, none of these accounts specifies how behavior is guided in the presence of rewards
84 and/or threats of different magnitudes. Several arguments suggest that these biases should be sensitive
85 to the magnitude of these prospective outcomes (or “stakes”). Agents frequently face situations in which
86 they have to select amongst multiple rewards of varying magnitude. It could be beneficial if Pavlovian
87 biases would automatically direct the agent towards the largest reward. Particularly, on its way to
88 attaining the largest reward, an agent might have to ignore smaller, more proximal rewards. Hence,
89 Pavlovian biases should not be triggered by any reward, but distinguish between smaller rewards on the
90 one hand, which might be arbitrated against other goals an agent pursues using deliberational processes,
91 and sufficiently large rewards on the other hand, which escape such an arbitration and instead elicit
92 unconditional approach behavior. Similar, the danger level of potential threats (or “threat magnitude”)
93 needs to be considered: A human hunter who freezes upon the sight of a lion might have a competitive
94 advantage over someone who continues to forage. However, a hunter who freezes upon the sight of a
95 small spider might have a disadvantage compared to other foragers, demonstrating again that Pavlovian
96 biases can only be adaptive if they take the magnitude of rewards and threats into account and ignore
97 smaller outcomes in service of pursuing larger outcomes.

98 Evidence that the strength of Pavlovian biases varies with stake magnitude has been mixed so
99 far. A few studies using Pavlovian-to-Instrumental Transfer (PIT) tasks, in which task-irrelevant cues
100 associated with rewards/ punishments are presented in the background, have observed slight increases
101 in response rates and somewhat faster reaction times for higher rewards (Algermissen & den Ouden,
102 2023; Schad et al., 2020) as well as decreased response rates and slower reaction times for larger
103 punishments (Geurts, Huys, den Ouden, & Cools, 2013b, 2013a). However, many other studies have

104 not observed such modulations (Chen, Garbusow, Sebold, Kuitunen-Paul, et al., 2022; Chen,
105 Schlagenhauf, et al., 2022; Garbusow et al., 2019, 2016; Sommer et al., 2020, 2017). Other tasks varying
106 the reward on offer, specifically versions of the monetary incentive delay task (Knutson, Adams, Fong,
107 & Hommer, 2001; Luo, Ainslie, Giragosian, & Monterosso, 2009) have observed faster reaction times
108 to larger rewards. A study using a virtual predation game found slower reaction times under larger
109 threats (Bach, 2015). However, in the latter studies, it remained unclear whether reward-induced
110 invigoration/ punishment-induced slowing followed from automatic, Pavlovian effects or rather
111 participants' deliberate strategies, reflecting their beliefs about which behavior was conducive to reward
112 attainment/ punishment avoidance (Mahlberg et al., 2021; Westbrook, Frank, & Cools, 2021). To
113 disentangle automatic from strategic effects, there must be task conditions that incentivize the
114 suppression of Pavlovian biases—a unique feature of the Motivational Go/NoGo Task.

115 Pavlovian biases can most unequivocally be measured with the (orthogonalized) Motivational
116 Go/NoGo Task. In this task, individuals learn through trial-and-error to perform either a Go or NoGo
117 response to a number of different cues. For some cues (“Win cues”), they can gain points for correct
118 performance (with no change in score for incorrect performance), while for other cues (“Avoid” cues),
119 they can lose points for incorrect performance (with no change in score for correct performance; Fig.
120 1A-C). In this task, humans typically show higher accuracy in performing active “Go” actions to Win
121 cues than passive “NoGo” actions to Win cues, while the reverse is true for Avoid cues, reflecting the
122 influence of Pavlovian biases (Guitart-Masip, Duzel, Dolan, & Dayan, 2014; Guitart-Masip et al., 2012;
123 Swart et al., 2017). Beyond differences in accuracy, humans also show faster responses to Win than
124 Avoid cues. In order to perform well on this task, participants need to detect when Pavlovian biases are
125 incongruent with the required response and inhibit their biases on these trials (Cavanagh, Eisenberg,
126 Guitart-Masip, Huys, & Frank, 2013; Swart et al., 2018). Unlike PIT tasks, every cue signaling whether
127 to perform a Go or NoGo response has a fixed valence, either providing the chance to win or to lose
128 points, typically eliciting stronger biases than tasks in which task-irrelevant cues are presented in the
129 background.

130 While Pavlovian biases might lead to adaptive behavior in a number of situations, their
131 influence becomes most apparent in situations in which they conflict with optimal behavior: Sometimes,
132 agents have to wait to secure a reward, e.g., in situations akin to the Marshmallow Test (Mischel &
133 Ebbesen, 1970), or they have to take active steps to prevent or fight a threat, e.g., in exposure therapy
134 to treat arachnophobia. In such circumstances, agents have to suppress Pavlovian biases, a requirement
135 animals usually struggle with (Breland & Breland, 1961; Hershberger, 1986) and even humans only
136 imperfectly master (Cavanagh et al., 2013; Swart et al., 2018). The ability to suppress automatic,
137 unwanted action tendencies is usually regarded to require cognitive control (Cohen, 2017). For several
138 decades, cognitive control has been seen as a limited resource or ability that can fail, leading to action
139 slips and undesired behavior (Hofmann, Friese, & Strack, 2009). In contrast, more recent perspectives,
140 most notably the *expected value of control theory* (EVC) (Lieder, Shenhav, Musslick, & Griffiths, 2018;
141 Shenhav, Botvinick, & Cohen, 2013) have suggested that cognitive control is not inherently limited,
142 but follows from a cost-benefit trade-off that weighs the potential benefits of exerting additional control
143 against the costs of doing so. In line with this idea, a number of studies using conflict tasks, such as the
144 Stroop, Simon, or Flanker task, have shown that compatibility effects—taken to reflect cognitive control
145 limitations—become smaller when participants are offered financial incentives for recruiting control
146 (Boehler, Hopf, Stoppel, & Krebs, 2012; Chiew & Braver, 2014; Dixon & Christoff, 2012; Fröber &
147 Dreisbach, 2016; Krebs, Boehler, & Woldorff, 2010). From this perspective, higher stakes should
148 motivate an agent to exert additional cognitive control in order to suppress biases in situations in which
149 those are maladaptive. In these situations, notably, the EVC theory makes predictions directly opposite
150 to the above-described case of high stakes strengthening biases: while ecological considerations suggest
151 that higher stakes should lead to stronger biases, EVC predicts more control and thus weaker biases. To
152 suppress biases, additional time might be required to recruit control processes, leading to higher
153 accuracy on behalf of longer RTs, i.e., a speed-accuracy tradeoff. In contrast, in situations in which
154 biases lead to adaptive behavior, EVC predicts no effect of stakes on behavior.

155 In this study, we directly tested these two opposing predictions against each other. We collected
156 data from 55 participants performing the motivational Go/NoGo Task in which the magnitude of stakes

157 (high or low) was manipulated on a trial-by-trial basis. Following the first hypothesis that higher stakes
 158 drive stronger Pavlovian biases, we predicted an interaction between congruency and the stakes
 159 magnitude, with a stronger congruency effect (indicative of the Pavlovian bias) and higher performance
 160 on congruent, but lower performance on incongruent trials under high compared to low stakes (Fig.
 161 1D). In contrast, following the EVC hypothesis, we predicted an interaction effect in the opposite
 162 direction, with a weaker congruency effect (reflecting cognitive control recruitment) and selectively
 163 higher performance on incongruent trials (but slower RTs) under high compared to low stakes (Fig.
 164 1E).

165

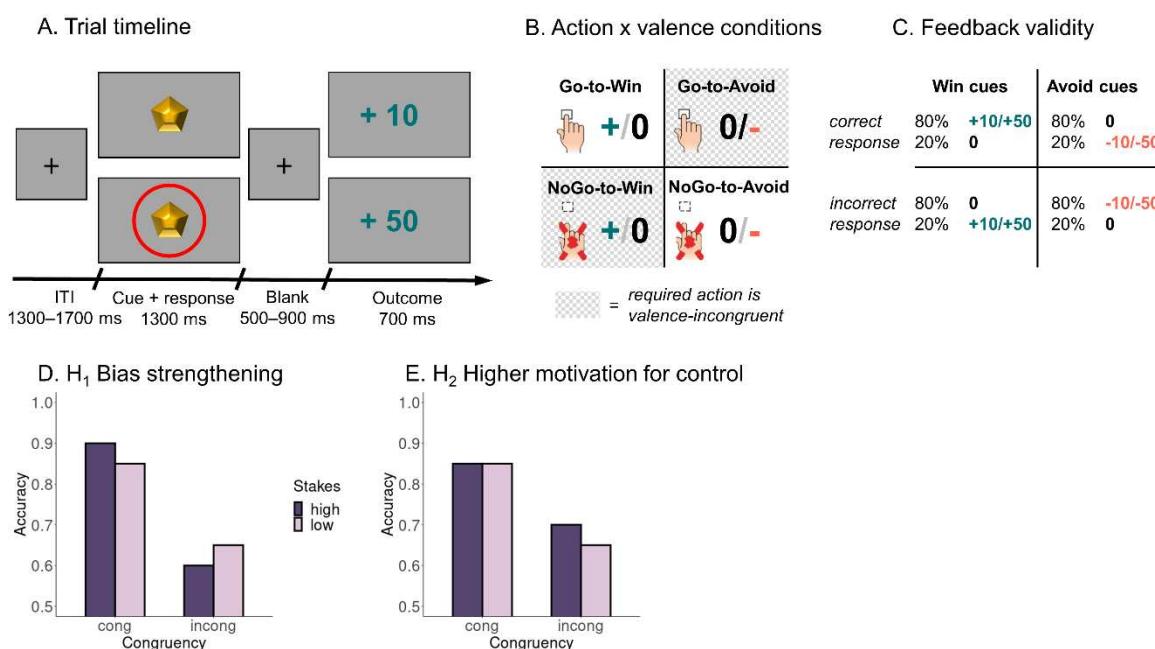


Figure 1. Task and behavioral predictions. **A.** Time course of each trial. Participant see one of four cues (“gems”) and have to decide whether to respond to it with a button press (“Go”) or not (“NoGo”). On half of the trials, the cue is surrounded by a red circle, indicating that stakes are five times as high and points gained/ lost in this trial will be multiplied with 5. After a variable interval, participants receive an outcome (increase in points, no change, or decrease in points). **B.** Task conditions. Half of the cues are “Win” cues for which points can be gained (or no change in the point score occurs), while the other half are “Avoid” cues for which points can be lost (or no change in the point score occurs). Orthogonal to cue valence is the correct action required for each cue, which is either Go or NoGo. **C.** Feedback given cue valence and response accuracy. For Win cues, correct responses mostly lead to an increase in points (+10 or +50, depending on whether the trial was high or low stakes), but occasionally lead to no change in score (0). For Avoid cues, correct responses mostly lead to no change in score (0), while occasionally lead to a loss of points (-10 or -50, depending on whether the trial was high or low stakes). For incorrect responses, probabilities are reversed. **D.** Prediction from a “bias strengthening” hypothesis. High stakes strengthen biases, leading to higher accuracy for bias-congruent cues (for which required action and valence match), but lower accuracy for bias-incongruent cues. **E.** Prediction from the “motivation for control” hypothesis. High stakes motivate cognitive control, which inhibits biases when they are incongruent with the required action, leading to higher accuracy selectively for bias-incongruent cues (for which the bias-triggered response has to be inhibited).

166

Methods

167 Participants and Exclusion Criteria

168 Fifty-five human participants ($M_{\text{age}} = 22.31$, $SD_{\text{age}} = 2.21$, range 18–29; 42 women, 13 men; 47
169 right-handed, 8 left-handed) participated in an experiment of about 45 minutes. The study design,
170 hypotheses, and analysis plan were pre-registered on OSF under <https://osf.io/ue397>. Individuals who
171 were 18–30 years old, spoke and understood English, and did not suffer from colorblindness were
172 recruited via the SONA Radboud Research Participation System of Radboud University. Their data
173 were excluded from all analyses for two (pre-registered) reasons: (a) guessing the hypotheses of the
174 experiment on the first question of the debriefing, which was not the case for any participant; (b)
175 performance not significantly above chance (tested by using required action to predict performed action
176 with a logistic regression; only participants with $p < .05$ were included), which was the case for one
177 participant. All the results presented in the main text are thus based on a final sample of $N = 54$. See the
178 Supplementary Material S03 for results based on all 55 participants, which led to identical conclusions.
179 This research was approved by the local ethics committee of the Faculty of Social Sciences at Radboud
180 University (proposal no. ECSW-2018-171) in accordance with the Declaration of Helsinki.

181 The sample size was not based on a power analysis, but on lab availability for this project (three
182 weeks). This study was conducted as part of final year thesis projects, which received special lab access
183 in this period. The final sample size of $N = 54$ was larger than previous studies investigating Pavlovian
184 biases with the same task (Algermissen, Swart, Scheeringa, Cools, & den Ouden, 2022; Swart et al.,
185 2018) and more than twice as large as comparable studies investigating the effect of incentives on
186 cognitive control recruitment (Chiew & Braver, 2016; Krebs et al., 2010). A post-hoc sensitivity power
187 analysis yielded that, given 54 participants providing 320 trials, thus 17,280 trials in total, assuming an
188 intra-cluster coefficient of 0.043 for responses and 0.094 for RTs (estimated from the data), the effective
189 sample size was $n = 5,281$ for responses and $n = 2,877$ for RTs, which allowed us to detect effects of β
190 $> .039$ (standardized regression coefficient) for responses and $\beta > .052$ for RTs with 80% power (Aarts,
191 Verhage, Veenvliet, Dolan, & van der Sluis, 2014).

192 **Procedure**

193 Participants completed a single experimental session that lasted about 45 minutes. After
194 providing informed consent, participants received computerized instructions and performed four
195 practice trials for each of the four task conditions. Afterwards, they completed 320 trials of the
196 Motivational Go/NoGo Task. After the task, participants performed the V5-D MESA Digit Span Test
197 measuring forward and backward digit span (Fitzpatrick et al., 2015) and filled in the non-planning
198 subscale of the Barratt Impulsiveness Scale (Patton, Stanford, & Barratt, 1995) and the neuroticism sub-
199 scale of the Big Five Aspects Scales (DeYoung, Quilty, & Peterson, 2007). These measures were part
200 of final year thesis projects and not of focal interest (see the pre-registration); results are reported in
201 Supplementary Material S05. Finally, participants went through a funnel debriefing asking them about
202 their guesses of the hypothesis of the study, whether they used specific strategies to perform the task,
203 whether they found the task more or less difficult to perform on high stakes trials, and if so, whether
204 they had an explanation of why this was the case. At the end, they received course credit for participation
205 as well as a small extra candy reward when they scored more than 960 points (equivalent to 67%
206 accuracy across trials, equivalence unknown to participants), which was announced in the instructions.

207 **Task**

208 Participants completed 320 trials (80 per condition; 40 each with high and low stakes
209 respectively) of the Motivational Go/ NoGo learning task. Each trial started with one of four abstract
210 geometric cues presented for 1,300 ms (Fig. 1A). The assignment of cues to task conditions was
211 counterbalanced across participants. Participants needed to learn from trial-and-error about the cue
212 valence, i.e., whether the cue was a Win cue (point gain for correct responses; no change in point score
213 for incorrect responses) or an Avoid cue (no change in point score for correct responses; point loss for
214 incorrect responses), and the required action, i.e., whether the correct response was Go (a key press of
215 the space bar) or NoGo (no action; Fig. 1B). Participants could perform Go responses while the cue was
216 on the screen. In 50% of trials, the cue was surrounded by a dark red circle (RGB [255, 0, 0]), signaling
217 the chance to win or avoid losing 50 points (high stakes condition). On all other trials, 10 points could
218 be won or lost (low stakes condition). After a variable inter-stimulus interval of 500–900 ms (uniform

219 distribution in steps of 100 ms), numerical feedback was presented for 700 ms (+10/+50 in green font
220 for point wins, -10/-50 in red font for point losses; 000 in grey font for no change in point score).
221 Feedback was probabilistic such that correct responses were followed by favorable outcomes (point win
222 for Win cues, no change for Avoid cues) on only 80% of trials, while on the other 20% of trials,
223 participants received unfavorable outcomes (no change for Win cues, point loss for Avoid cues; Fig.
224 1C). These probabilities were reversed for incorrect responses. Probabilistic feedback was used to make
225 learning more difficult and induce a slower learning curve. Trials ended with a variable inter-trial
226 interval of 1,300–1,700 ms (uniform distribution in steps of 100 ms).

227 The task was administered in four blocks of 80 trials each. Each block featured a distinct set of
228 four cues for which participants had to learn the correct response. Probabilistic feedback and renewal
229 of the cue set were used to slow down learning, given previous findings that biases disappear when
230 accuracy approaches 100% (Swart et al., 2017).

231 **Data Analysis**

232 ***Data Preprocessing***

233 (Trials with) RTs faster than 300 ms were excluded from all analyses as those were assumed to
234 be too fast to reflect processing of the cue. This was the case for 103 out of 17,600 trials (per participant:
235 $M = 1.91$, $SD = 5.89$, range 0–41). See Supplementary Material S02 for results using all reaction times
236 from all trials.

237 ***Mixed-effects Regression Models***

238 We tested hypotheses using mixed-effects linear regression (function `lmer`) and logistic
239 regression (function `glmer`) as implemented in the package `lme4` in R (Bates, Mächler, Bolker, &
240 Walker, 2015). We used generalized linear models with a binomial link function (i.e., logistic
241 regression) for binary dependent variables such as accuracy (correct vs. incorrect) and response (Go vs.
242 NoGo), and linear models for continuous variables such as RTs. We used zero-sum coding for
243 categorical independent variables. All continuous dependent and independent variables were
244 standardized such that regression weights can be interpreted as standardized regression coefficients. All
245 regression models contained a fixed intercept. We added all possible random intercepts, slopes, and

246 correlations to achieve a maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013). *P*-
247 values were computed using likelihood ratio tests with the package afex (Singmann, Bolker, Westfall,
248 & Aust, 2018). We considered *p*-values smaller than $\alpha = 0.05$ as statistically significant.

249 ***Evidence for absence of an effect***

250 We plot the condition means for each participant and provide confidence intervals for every
251 effect. Every possible point estimate of an effect that would fall outside the estimated confidence
252 interval can be rejected at a level of $\alpha = 0.05$.

253 ***Computational modeling of responses and reaction times***

254 **Combining reinforcement learning with a drift-diffusion choice rule.** A class of
255 computational models that allows to jointly model both responses and reaction times are so called
256 “evidence accumulation” or “sequential sampling” models such as the drift-diffusion model (DDM)
257 (Ratcliff, 1978). These models formalize a decision process in which evidence for two (or more)
258 response options is accumulated until a fixed threshold, and a response is elicited upon reaching this
259 threshold. The process is captured through four parameters (Wabersich & Vandekerckhove, 2014): the
260 drift rate δ , reflecting the speed with which evidence is accumulated; the decision threshold α ,
261 describing the distance of the threshold from the starting point; the starting point bias β , reflecting if the
262 accumulation process starts in the middle between both bounds ($\beta = 0.5$) or closer to one of the
263 boundaries, reflecting an overall response bias; and the non-decision time τ ; capturing the duration of
264 all perceptual or motor processes that contribute to RT, but are not part of the decision process itself.

265 Typically, DDMs aim to explain choices when response requirements given a certain visual
266 input are clear to the participant. However, in the current study, participants learn the correct response
267 for each cue over time, leading to progressively faster and more accurate responses. Recent advances
268 in computational modeling propose that it is possible to combine drift-diffusion models with a
269 reinforcement learning (RL) process, yielding a reinforcement-learning drift-diffusion model (RL-
270 DDM) (Fontanesi, Gluth, Spektor, & Rieskamp, 2019; Miletić, Boag, & Forstmann, 2020; Pedersen,
271 Frank, & Biele, 2017). We employed a simple Rescorla-Wagner model which uses outcomes r (+1 for

272 rewards, 0 for neutral outcomes, -1 for punishments) to compute prediction errors $r - Q$, which we then
273 used to update the action value Q for the chosen action a towards cue s :

274
$$Q_t(a_t, s_t) = Q_{t-1}(a_t, s_t) + \varepsilon * (r - Q_{t-1}(a_t, s_t)) \quad (1)$$

275 Here, the difference in Q -values between choice options ($Q_{Go} - Q_{NoGo}$) serves as the input to
276 the drift rate. This difference is initially zero, but grows with learning (positive difference if “Go” leads
277 to more rewards, and negative difference if “NoGo” leads to more rewards). This Q -value difference is
278 then multiplied with a constant drift rate parameter. At the beginning of the learning process, the
279 resulting low drift rates lead to more stochastic choices and slow RTs, but, as the Q -value difference
280 grows, higher drift rates result in more deterministic choices and faster RTs. The learning process
281 requires an additional free parameter, i.e., the learning rate parameter ε , which determines the impact
282 of the prediction error on belief updating. The drift rate parameter acts akin to the inverse temperature
283 parameter used in the softmax choice rule, with higher drift rates leading to more deterministic choices.

284 One peculiarity of the Motivational Go-NoGo Task is the NoGo response option, which by
285 definition does not yield RTs. Variants of the DDM allow for such responses by integrating over the
286 latent RT distribution of the implicit NoGo decision boundary (Gomez, Ratcliff, & Perea, 2007;
287 Ratcliff, Huang-Pollock, & McKoon, 2018), for which an approximation exists (Blurton, Kesselmeier,
288 & Gondan, 2012). This implementation has previously been used to model another variant of
289 motivational Go/ NoGo task (Millner, Gershman, Nock, & den Ouden, 2017) and is implemented in the
290 HDDM toolbox (Wiecki, Sofer, & Frank, 2013).

291 Note that RL-DDMs were not mentioned in the pre-registration, which only mentioned
292 reinforcement learning models to-be fitted to participants’ choices. In light of the results from the
293 regression analyses, incorporating RTs into the model and testing alternative mechanisms by which
294 stakes could influence the choice process seemed warranted.

295 **Model space.** We fit a series of increasingly complex models. We first tested whether an RL-
296 DDM fit the data better than a standard DDM; then tested the computational implementation of the
297 Pavlovian bias, and lastly tested the effect of stakes on model parameters. Model **M1** (parameters α , τ ,

298 $\beta, \delta_{\text{INT}}$) just featured the DDM model with a constant drift rate parameter, but no learning, assuming
299 that participants have a constant propensity to make a Go response for any trial, irrespective of the
300 presented cue. **M2** (parameters $\alpha, \tau, \beta, \delta_{\text{INT}}, \delta_{\text{SLOPE}}, \varepsilon$) added a reinforcement learning process, updating
301 Q-values for Go and NoGo for each cue with the observed feedback, multiplying the Q-value difference
302 ($Q_{\text{Go}} - Q_{\text{NoGo}}$) with the drift rate parameter δ_{SLOPE} and finally adding it to the drift-rate intercept δ_{INT} to
303 obtain the net drift rate. Including a drift-rate intercept δ_{INT} , i.e., an overall tendency towards making a
304 Go/NoGo response even when the Q-value difference was zero, which is similar to an overall Go bias
305 parameter, yielded considerably better fit than models without such an intercept. If people learned the
306 task, model M2 should fit their data better than M1. Next, M3 and M4 comprised different
307 implementations of the Pavlovian bias, either assuming separate starting point biases (**M3**; parameters
308 $\alpha, \tau, \beta_{\text{WIN}}, \beta_{\text{AVOID}}, \delta_{\text{INT}}, \delta_{\text{SLOPE}}, \varepsilon$) or alternatively separate drift rate intercepts (**M4**; parameters $\alpha, \tau, \beta,$
309 $\delta_{\text{WIN}}, \delta_{\text{AVOID}}, \delta_{\text{SLOPE}}, \varepsilon$) for Win and Avoid cues, two plausible implementations considered in previous
310 literature (Millner et al., 2017). Next, models M5-M8 (parameters $\alpha, \tau, \beta, \delta_{\text{WIN}}, \delta_{\text{AVOID}}, \delta_{\text{SLOPE}}, \varepsilon$, one
311 additional parameter π for high stakes) extended M4 and tested possible effects of the stakes on a single
312 parameter, implementing effect of the stakes on the threshold (**M5**), the non-decision time (**M6**), the
313 bias (**M7**) and the drift rate intercept (**M8**). As a control, models M9-M11 (parameters $\alpha, \tau, \beta, \delta_{\text{WIN}},$
314 $\delta_{\text{AVOID}}, \delta_{\text{SLOPE}}, \varepsilon$, two additional parameters π and θ for high stakes) tested effects of stakes on two
315 parameters (only combinations that could potentially give rise to response slowing), namely on both the
316 threshold and the non-decision time (**M9**), the threshold and the drift rate (**M10**; i.e. the two parameters
317 typically modulated by speed-accuracy trade-offs), and the non-decision time and drift rate (**M11**).
318 Finally, given the results from model comparison of these earlier models, **M12** tested whether the effect
319 of stakes of non-decision time was different for congruent and incongruent cues.

320 **Priors, transformations, parameterization, and starting values.** We fitted models in a
321 hierarchical Bayesian fashion, modeling group-level parameters (means and standard deviations) that
322 served as priors for the subject-level parameters using the probabilistic programming language Stan
323 (Carpenter et al., 2017) in R (rstan). Stan implements a Hamiltonian Monte-Carlo (HMC) Markov-
324 chain algorithm with a No-U-Turn sampler (NUTS). We used the following group-level hyperpriors:

325 $M_\delta \sim N(5, 2)$, $M_\alpha \sim N(0, 1)$, $M_\beta \sim N(0, 1)$, $M_\tau \sim N(0, 1)$, $M_\varepsilon \sim N(0, 1)$, $M_\pi \sim N(0, 1)$, $M_\theta \sim N(0, 1)$, and
326 for all SDs: $SD \sim N(0, 1)$. The parameters δ , α , τ were constrained to be positive by using the $y = \log(1$
327 $+ \exp(x))$ transformation, which is $y = 0$ for negative numbers, smoothly asymptotes 0 for small positive
328 numbers, and is roughly $y = x$ for large positive numbers. The parameters β and ε were constrained to
329 be in the range $[0, 1]$ by using a softmax transformation $y = \exp(x) / (1 + \exp(x))$. In line with previous
330 DDM implementations in Stan (Fontanesi et al., 2019; Kraemer, Fontanesi, Spektor, & Gluth, 2021),
331 we used a non-centered parameterization in which individual-subject parameters are modeled with a
332 standard normal prior $N(0, 1)$ that is first multiplied with the group-level standard deviation and then
333 added to the group-level mean parameter. Furthermore, again in line with previous DDM
334 implementations in Stan (Fontanesi et al., 2019; Kraemer et al., 2021), we set the following starting
335 values: $M_\alpha = -0.18$, $M_\tau = -10$, $M_\beta \sim N(0.5, 0.1)$, $M_{\delta\text{INT}} \sim N(0, 1)$, $M_{\delta\text{SLOPE}} \sim N(0, 1)$, $M_\pi \sim N(0, 0.1)$, $M_\theta \sim$
336 $N(0, 0.1)$, all group-level $SDs = 0.001$, all subject level parameters as $\sim N(0, 1)$. For models with an
337 effect of stakes on the non-decision-time (M_6 , M_9 , M_{11}), τ (low stakes) had to be initialized to be
338 considerably smaller than π (high stakes), which was accomplished by $M_\tau \sim N(0, 1e-6)$ and $SD_\tau = 1e-6$.

339 **Model fitting and convergence checks.** For each model, we used four chains with 10,000
340 iterations each (5,000 as warm-up), yielding a total of 20,000 samples contributing to the posteriors.
341 We checked that Rhats for all parameters were below 1.01, effective sample sizes for all parameters
342 were at least 400, that chains were stationary and well-mixing (using trace plots), that the Bayesian
343 fraction of missing information (BFMI) for each chain was above 0.2, and that (if possible) no divergent
344 transitions occurred (Baribault & Collins, 2023). To minimize the occurrence of divergent transitions,
345 we increased the target average proposal acceptance probability (adapt_delta) to 0.99. We visually
346 inspected that posterior densities were unimodal and no strong trade-offs between parameters across
347 samples occurred.

348 **Model comparison.** For model comparison, we used the LOO-IC (efficient approximate leave-
349 one-out cross-validation information criterion) based on Pareto-smoothed importance sampling (PSIS)
350 (Vehtari, Gelman, & Gabry, 2017). For completeness, we also report the WAIC (widely applicable
351 information criterion) in Supplementary Material S07, but give priority to the LOO-IC, which is more

352 robust to weak priors or influential observations (Vehtari et al., 2017). Both WAIC and LOO-IC behave
353 like the negative log-likelihood, with lower numbers indicating better model fit.

354 **Posterior predictive checks.** For the winning model M12, we randomly drew 1,000 samples
355 from the posteriors of each participants' subject-level parameters, simulated a data set for each
356 participant for each of these 1,000 parameter settings, and computed the mean simulated $p(\text{Go})$,
357 $p(\text{Correct})$, and RT for each participant for each trial across parameter settings. We then plotted the
358 mean simulated $p(\text{Go})$, $p(\text{Correct})$, and RT as a function of relevant task conditions to verify that the
359 model could reproduce key qualitative patterns from the empirical data (Palminteri, Wyart, & Koechlin,
360 2017).

361 **Parameter recovery.** For the winning model M12, we fitted a multivariate normal distribution
362 to the mean subject-level parameters across participants and sampled 1,000 new parameter settings from
363 this distribution. We simulated a data set for each parameter setting and fitted model M12 to the
364 simulated data. We then correlated the “ground-truth” generative parameters used to simulate each data
365 set to the fitted parameters obtained when fitting M12 to it. To evaluate whether correlations were
366 significantly higher than expectable by chance, we computed a permutation null distribution of the on-
367 diagonal correlations. For this purpose, over 1,000 iterations, we randomly permuted the assignment of
368 fitted parameter values to data sets, correlated generative and fitted parameter values, and saved the on-
369 diagonal correlations. We tested empirical correlations against the 95th percentile of this permutation
370 null distribution.

371 **Model recovery.** For each of the 12 models, we fitted a multivariate normal distribution to the
372 mean subject-level parameters across participants and sampled 1,000 new parameter settings from it
373 (with the constraints that learning rates were required to be > 0.05 and parameter differences sampled
374 from the upper 50% of the parameter distribution to keep models distinguishable). We simulated a new
375 data set for each parameter setting, resulting in total in 12,000 data sets. We fitted each of the 12 models
376 to each data set, resulting in 144,000 model fits. For each data set, we identified the model with the
377 lowest LOO-IC. We counted how often each fitted model Y emerged as the winning model for the data
378 sets of each generative model X, computing the forward confusion matrix containing conditional

379 probabilities $p(\text{best fitting model} = Y | \text{generative model} = X)$ for each combination of generative model
380 X and fitted model Y (Wilson & Collins, 2019). We also computed the inverse confusion matrix
381 containing $p(\text{generative model} = X | \text{best-fitting model} = Y)$; see Supplementary Material S07). To
382 evaluate whether these probabilities were significantly higher than expectable by chance, we computed
383 a permutation null distribution of the on-diagonal probabilities. For this purpose, over 1,000 iterations,
384 we randomly permuted the LOO-IC values of all fitted models for a given data set, counted how often
385 each fitted model emerged as the winning model for the data sets of each generative model, and
386 extracted the on-diagonal probabilities. We tested empirical probabilities against the 95th percentile of
387 this null distribution.

388 **Transparency and openness**

389 We report how we determined our sample size, all data exclusions, all manipulations, and all measures
390 in the study. All data, analysis code, and research materials will be shared upon publication. The study
391 design, hypotheses, and analysis plan were pre-registered on OSF under <https://osf.io/ue397>. Data were
392 analyzed using R, version 4.1.3 (R Core Team, 2022). Models were fitted with the package lme4,
393 version 1.1.31 (Bates et al., 2015). Plots were generated with ggplot, version 3.4.2 (Wickham, 2016).

394 **Results**

395 **Manipulation checks: Learning and Pavlovian biases**

396 As a manipulation check and in order to compare the results from this study to previous studies
397 (Algermissen et al., 2022; Swart et al., 2018, 2017), we fitted a mixed-effects logistic regression with
398 responses (Go/ NoGo) as dependent variable as well as required action (Go/ NoGo) and valence (Win/
399 Avoid) and independent variables (see Supplementary Material S01 for an overview of all regression
400 results; see Supplementary Material S04 for means and standard deviations per condition). Participants
401 made significantly more Go responses to Go cues than NoGo cues (required action), $b = 1.441$, 95%-
402 CI [1.252, 1.630], $\chi^2(1) = 87.873$, $p < .001$, indicating that they learned the task. They also showed
403 significantly more Go responses to Win than Avoid cues (cue valence), $b = 0.750$, 95%-CI [0.609,
404 0.889], $\chi^2(1) = 59.587$, $p < .001$, reflecting a Pavlovian bias (Fig. 2A–C). There was no evidence for the

405 Pavlovian bias being stronger for either Go or NoGo cues (required action x valence), $b = 0.019$, 95%-
406 CI [-0.100, 0.137], $\chi^2(1) = 0.093$, $p = .760$.

407 Next, we performed a similar mixed-effects linear regression with reaction times (RTs) as
408 dependent variable. Note that RTs were naturally only available for (correct and incorrect) Go
409 responses. Participants showed significantly faster (correct) responses to Go cues than (incorrect)
410 responses to NoGo cues (required action), $b = -0.109$, 95%-CI [-0.145, -0.073], $\chi^2(1) = 27.494$, $p < .001$,
411 and significantly faster responses to Win than Avoid cues (cue valence), $b = -0.191$, 95%-CI [-0.227, -
412 0.155], $\chi^2(1) = 59.204$, $p < .001$, again reflecting the Pavlovian bias (Fig. 3A–C). The cue valence effect
413 (Pavlovian bias) on RTs was slightly stronger for (correct) response to Go cues than (incorrect)
414 responses to NoGo cues (required action x cue valence) , $b = -0.032$, 95%-CI [-0.061, -0.003], $\chi^2(1) =$
415 4.384, $p = .036$. The strength of the Pavlovian bias (both in responses and RTs) was neither correlated
416 with working memory span, nor impulsivity, nor neuroticism (Supplementary Material S05). In sum,
417 participants learned the task and exhibited a Pavlovian bias in both responses and RTs.

418

419

420

421

422

423

424

425

426

427

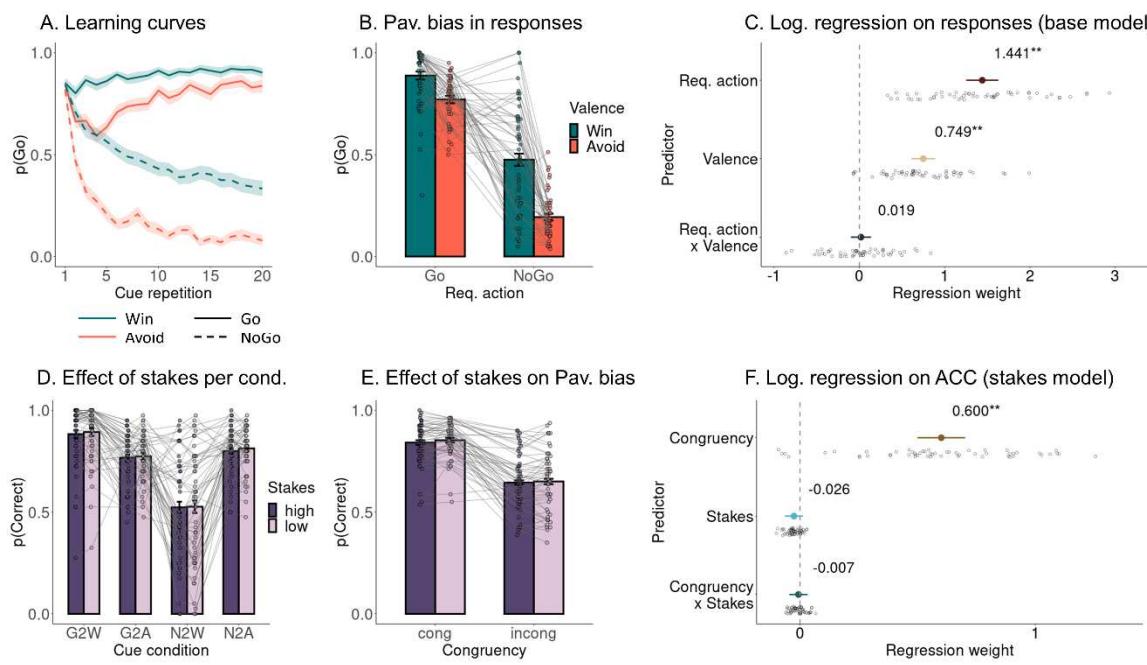


Figure 2. Effect on propensity of Go responses. **A.** Learning curves per cue condition. **B.** Proportion of Go responses per cue condition (individual dots are individual participant means). Participants show more Go responses to Go than NoGo cues (indicative of learning the task) and more Go responses to Win cues than Avoid cues (indicative of Pavlovian biases). **C.** Group-level (colored dot, 95%-CI) and individual-participant (grey dots) regression coefficients from a mixed-effects logistic regression of responses on required action, cue valence, and their interaction. **D.** Accuracy per cue condition and stakes condition. There is no effect of stakes on responses for any cue condition. **E.** Accuracy per valence-action congruency and stakes condition. Accuracy is higher for congruent than incongruent conditions, but this congruency effect is not modulated by stakes. **F.** Group-level and individual-participant regression coefficients from a mixed-effects logistic regression of responses on congruency, stakes, and their interaction.

428

429 **Confirmatory analyses: Modulation by stakes**

430 As the first set of confirmatory, pre-registered analyses, we fitted a mixed-effects logistic
 431 regression with accuracy (correct/ incorrect) as dependent variable and congruency (congruent/
 432 incongruent) and stakes (high/ low) as independent variables. There was a significant main effect of
 433 congruency, $b = 0.600$, 95%-CI [0.499, 0.702], $\chi^2(1) = 67.867$, $p < .001$, with higher accuracy to
 434 congruent than incongruent cues, again reflecting the Pavlovian bias. However, neither the main effect
 435 of stakes, $b = -0.026$, 95%-CI [-0.065, 0.013], $\chi^2(1) = 1.430$, $p = .232$, nor the interaction between
 436 congruency and stakes, $b = -0.007$, 95%-CI [0.046, 0.032], $\chi^2(1) = 0.094$, $p = .759$, was significant (Fig.
 437 2E, F).

438 Exploratory post-hoc tests for each cue condition separately did not show any effect of stakes
 439 on responses for any cue condition (Go-to-Win: $z = -0.590$, $p = .555$; Go-to-Avoid: $z = -0.184$, $p = .854$;
 440 NoGo-to-Win: $z = -0.145$, $p = .885$; NoGo-to-Avoid: $z = -0.963$, $p = .3357$; Fig. 2D). In further
 441 exploratory analyses, we tested whether an effect of stakes on responses emerged (or disappeared) over

442 time, either within the learning trajectory of a cue (cue repetition; 1 – 20) or across the entire task (trial
443 number: 1–320). Neither the interaction between cue repetition and stakes, $b = -0.002$, 95%-CI [-0.039,
444 0.035], $\chi^2(1) = 0.020$, $p = .898$, nor the interaction between trial number and stakes, $b = -0.012$, 95%-
445 CI [-0.048, 0.023], $\chi^2(1) = 0.401$, $p = .527$, was significant, providing no evidence for stakes influencing
446 responses selectively at certain time points during learning or during the task. In sum, there was no
447 evidence for stakes modulating the Pavlovian bias in participants' responses.

448 As the second set of confirmatory, pre-registered analyses, we fitted a mixed-effects linear
449 regression with RTs as dependent variable and congruency (congruent/ incongruent) and stakes (high/
450 low) as independent variables. Participants responded significantly faster to congruent than incongruent
451 cues (congruency), $b = -0.131$, 95%-CI [-0.160, -0.102], $\chi^2(1) = 49.546$, $p < .001$, reflecting the
452 Pavlovian bias. Furthermore, they responded significantly more slowly under high compared to low
453 stakes (stakes), $b = 0.072$, 95%-CI [0.051, 0.092], $\chi^2(1) = 33.702$, $p < .001$ (Fig 3E, F). Finally, the
454 interaction between congruency and stakes was significant, $b = -0.019$, 95%-CI [-0.037, -0.001], $\chi^2(1)$
455 = 3.856, $p = .049$, with a stronger congruency effect under high compared to low stakes. This effect was
456 also significant ($p = .046$) when including RTs < 300 ms (see Supplementary Material S02), but only
457 marginally significant ($p = .060$) when adding the data of remaining participant with not-above-chance
458 performance (see Supplementary Material S03). The effect of stakes on RTs was correlated neither with
459 working memory span, impulsivity, or neuroticism (Supplementary Material S05).

460 Exploratory post-hoc tests for each cue condition separately yielded a significant effect of
461 stakes on RTs for three out of four cue conditions, including in particular the two incongruent conditions
462 Go-to-Avoid and NoGo-to-Win (Go-to-Win: $z = 2.973$, $p = .003$; Go-to-Avoid: $z = 4.528$, $p < .001$;
463 NoGo-to-Win: $z = 4.975$, $p < .001$; NoGo-to-Avoid: $z = 1.414$, $p = .158$; Fig. 3D). In further exploratory
464 analyses, we tested whether the effect of stakes on responses got stronger or weaker with time, either
465 within the learning trajectory of a cue (cue repetition) or across the entire task (trial number). Neither
466 the interaction between stakes and cue repetition, $b = -0.012$, 95%-CI [-0.030, 0.006], $\chi^2(1) = 1.599$, p
467 = .206, nor the interaction between stakes and trial number, $b = 0.025$, 95%-CI [-0.021, 0.018], $\chi^2(1) =$
468 0.480, $p = .489$, was significant, providing no evidence for a change in the effect of stakes on RTs over

469 time. See Supplementary Material S06 for tests for non-linear changes with time, again finding no
 470 evidence for changes in the effect of stakes over time. In sum, these results suggest that high stakes
 471 affected participant responses in that they overall slowed down responses. This slowing was slightly
 472 stronger for incongruent than congruent cues and appeared to be constant over time. However, stakes
 473 did not affect response accuracy nor the degree of Pavlovian bias as indexed by the decisions to make
 474 a Go or NoGo response.

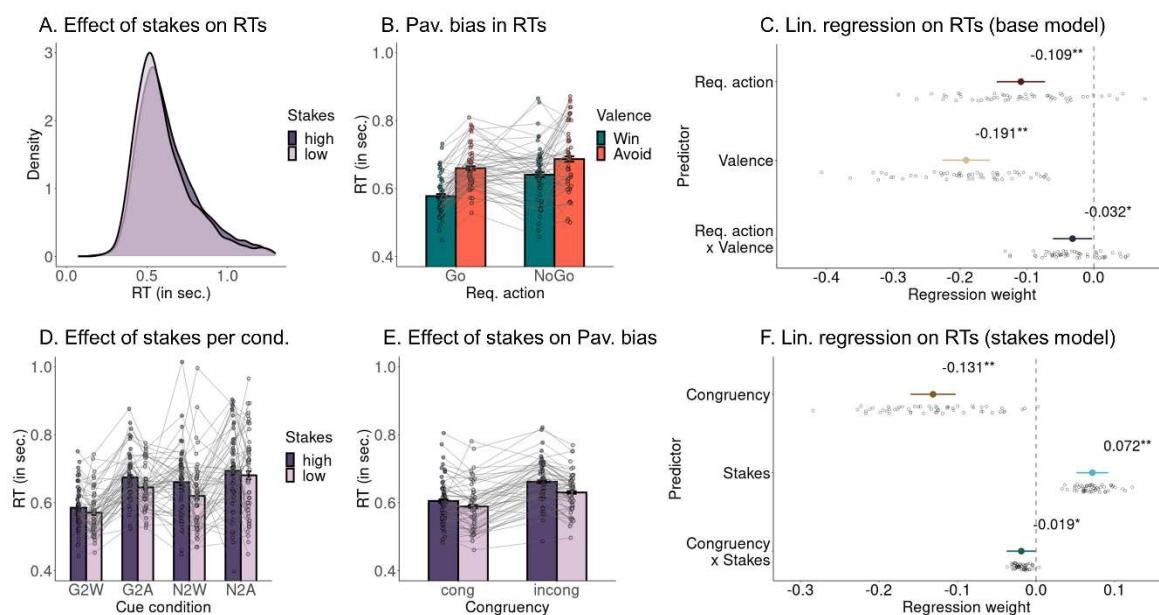


Figure 3. Effect on propensity of reaction times (RTs). **A.** Distribution of RTs for high and low stakes. RTs are slower under high stakes. **B.** RTs per cue condition. Participants show faster RTs to (correct) Go responses than (incorrect) Go responses to NoGo cues and faster RTs for Go to Win cues than Avoid cues (indicative of Pavlovian biases). **C.** Group-level (colored dot, 95%-CI) and individual-participant (grey dots) regression coefficients from a mixed-effects linear regression of RTs on required action, cue valence, and their interaction. **D.** RTs per cue condition and stakes condition. RTs are significantly slower under high stakes in the Go-to-Win (G2W), Go-to-Avoid (G2A), and NoGo-to-Win (NG2W) conditions. **E.** RTs per valence-action congruency and stakes condition. RTs are significantly slower under high compared to low stakes. This effect is significantly stronger for incongruent than congruent cue conditions. **F.** Group-level and individual-participant regression coefficients from a mixed-effects linear regression of RTs on congruency, stakes, and their interaction.

475

476 Computational Modeling of Responses and RTs (RL-DDMs)

477 To better understand the mechanisms by which cue valence and stakes influenced responses
 478 and RTs, we fit a series of increasingly complex reinforcement-learning drift-diffusion models (RL-
 479 DDMs). A past study using a similar paradigm found evidence for cue valence modulating the starting
 480 point bias in an evidence-accumulation framework rather than the drift rate (Millner et al., 2017),
 481 although evidence in that study remained mixed. Furthermore, past studies suggested that response-

482 slowing might reflect a speed-accuracy trade-off, with stakes leading to response caution and higher
483 decision thresholds, leading to higher accuracy at the cost of slower responses (Bogacz, Brown,
484 Moehlis, Holmes, & Cohen, 2006; Shevlin, Smith, Hausfeld, & Krajbich, 2022; Wiecki & Frank, 2013).
485 We implemented different mechanisms of how cue valence and stakes might influence the various
486 parameters (decision threshold, non-decision time, starting point bias, drift rate intercept) in an evidence
487 accumulation framework and compared the fit of different, increasingly complex models.

488 Behavior was better described by an RL-DDM (M2) in which participants learned cue-specific
489 Q-values rather than an standard DDM (M1) with a fixed propensity to emit Go/ NoGo responses (Fig.
490 4A), reflecting that participants learned the task and that learned affected responses and RTs. Model fit
491 was further improved when incorporating a Pavlovian bias (M3–M4), specifically when fitting separate
492 drift rate intercepts for Win and Avoid cues (M4; with high drift rate intercepts for Win than Avoid
493 cues, see Fig. 4B). Next, we assessed different mechanisms through which stake magnitude could affect
494 responding, which further improved model fit (M5–M8). Here, the best model was one in which stakes
495 modulate the non-decision time (M6). Note that, although M6 showed a superior fit to M4, group-level
496 non-decision times for high and low stakes were not significantly different from each other ($M_{diff} =$
497 0.012, 95%-CI [-0.017, 0.041]), suggestive of the presence of individual differences with an overall
498 mean close to zero. Allowing stakes to modulate two instead of one parameter did not yield any
499 substantial improvement in fit (M9–M11). Specifically, a model implementing a “classical” speed-
500 accuracy tradeoff by allowing stakes to influence both the threshold and the drift rate (M10) performed
501 worse than a model allowing stakes to influence the non-decision time (M6). Lastly, model fit was
502 further improved by when splitting the effect of stakes into separate parameters for congruent and
503 incongruent cues (M12), which was overall the best fitting model in the model comparison. Note that
504 M12 has the same number of parameters as models M9–M11, suggesting that the increase in fit is not
505 due to a mere increase in the number of parameters, but due to the specific mechanism implemented.
506 Also note that, although M12 with separate non-decision times under high stakes for congruent and
507 incongruent cues outperformed M6 with a single non-decision time under high stakes, there was no

508 group-level difference between the parameters for congruent vs. incongruent cues ($M_{diff} = -0.003$, 95%
509 CI [-0.033, 0.027], Fig. 4B), suggestive of individual differences with a group-level mean close to 0.

510 We performed several model validation checks to verify that the winning model M12 was able
511 to capture key qualitative features of the empirical data (posterior predictive checks), could identify
512 data-generating parameters reliably (parameter recovery), and could be distinguished from other models
513 (model recovery). Data simulated from M12 reproduced a Pavlovian bias in responses and RTs,
514 reproduced an overall slowing under high stakes, but somewhat underestimated the difference in RT
515 slowing between congruent and incongruent cues (Fig. 4C; see also Supplementary Material S07 for
516 further plots). Furthermore, generative and fitted parameters were overall highly correlated, indicative
517 of a successful parameter recovery ($M_r = 0.83$, $SD_r = 0.14$, range 0.62–0.98; 95th percentile of
518 permutation null distribution: $r = 0.08$; Fig. 4D; see Supplementary Material S07 for scatter plots of on-
519 diagonal correlations). Besides correlations between generative parameters with their corresponding
520 fitted parameters, there were two notable cases of off-diagonal correlations: first, the different non-
521 decision times (under low stakes, under high stakes for congruent cues, and under high stakes for
522 incongruent cues) were correlated ($r = 0.71$ and $r = 0.77$; Fig. 4D), reflecting an overall tendency
523 towards faster/ slower responses that is naturally shared across all three parameters. Second, learning
524 rates and drift rate slopes were negatively correlated across parameter settings ($r = -0.56$; Fig. 4D),
525 which mimics the frequently observed trade-off between learning rate and inverse temperature
526 parameters in more classic reinforcement learning models of choices (Ballard & McClure, 2019). In
527 RL-DDMs, the drift rate slope is multiplied with the Q-value difference, so that steeper slopes lead to
528 more deterministic choices and shallower slopes lead to more stochastic choices, similar to an inverse
529 temperature parameter. Finally, model recovery was successful, particularly for the winning model
530 M12, which was the best fitting model for 98% of data sets for which it was the generative model
531 (forward confusion matrix; Fig. 4E). Recovery for the other models was not quite as high, though still
532 significantly above chance for all models ($M_p = 0.31$, $SD_p = 0.32$, range 0.13–0.98; 95th percentile of
533 permutation null distribution: $p = 0.10$). See Supplementary Material S07 for matrices involving only
534 the five nested sub-versions of M12 (i.e., M1, M2, M4, M6, M12). In this restricted subset, recovery

535 was much higher ($M_p = 0.74$, $SD_p = 0.24$, range 0.44–0.99; 95th percentile of permutation null
536 distribution: $p = 0.22$). Also, see Supplementary Material S07 also for the inverse confusion matrix.

537 In sum, model comparison results were in line with the regression results, yielding a selective
538 effect of stakes in prolonging the non-decision time, and separately so for incongruent and congruent
539 cues. Stakes did not affect the threshold and/or the drift rate as typically observed in a speed-accuracy
540 trade-off. Hence, we conclude that stakes do not shift the speed-accuracy trade-off, but rather lead to a
541 response slowing independent of response selection.

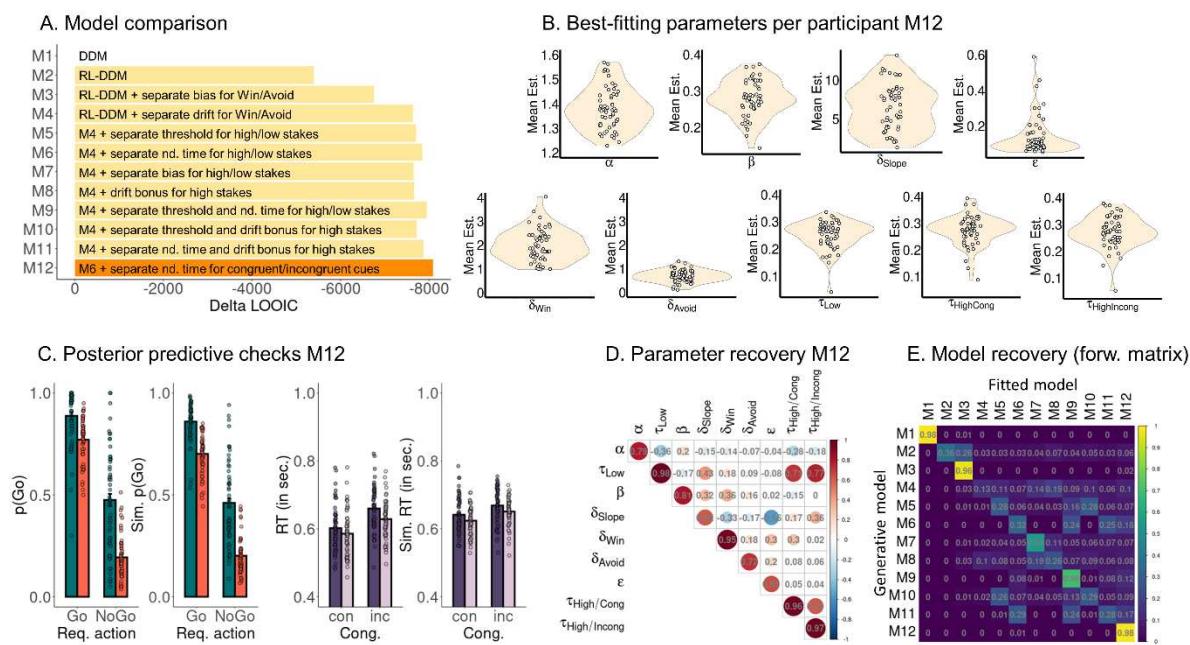


Figure 4. Reinforcement-learning drift-diffusion models. **A.** Model comparison. LOO-IC favors model M12, implementing separate drift rate intercepts for Win and Avoid cues and separate non-decision times for low stakes, congruent cues under high stakes, and incongruent cues under high stakes. **B.** Densities of best fitting parameters for model M12 per participant. Drift rate intercepts for Win cues are consistently higher than drift rate intercepts for Avoid cues. Note that, although the winning model implements separate non-decision times for high/ low stakes and congruent/ incongruent cues, the parameter values for these different conditions are not significantly different from each other. **C.** Posterior predictive checks for the winning model M12. Left panel: Simulated proportion of Go responses per required action and cue valence averaged over simulations and participants. The winning model M12 reproduces Pavlovian biases in responses and RTs (see Supplementary Material S07). Right panel: Simulated RTs per cue congruency per stakes level averaged over simulations and participants. The winning model M12 reproduces the overall slowing under high stakes as well as differences in slowing between congruent and incongruent cues, but underestimates this difference compared to the empirical data. For further plots, see Supplementary Material S07. **D.** Parameter recovery for the winning model M12. Correlations between generative parameters used for simulating 1,000 data sets based on M12 and parameters obtained when fitting M12 to simulated data. All correlations between generative and fitted parameters (on-diagonal correlations) are significantly above chance. **E.** Model recovery for model M1-M12. The forward confusion matrix displays the conditional probabilities that model Y is the best fitting model (columns) if model X (rows) is the underlying generative model used to simulate a given data set. On-diagonal probabilities indicate the probability of reidentifying the generative model. All on-diagonal probabilities are significantly above chance. Especially recovery for M12 is exceptionally high. For the inverse confusion matrix and matrix on subsets of models, see Supplementary Material S07.

542

Discussion

543 In this pre-registered experiment, we found evidence that increasing stake magnitude slowed
544 down responses in a Motivational Go/NoGo Learning Task, especially for incongruent cue conditions,

545 without affecting whether participants responded or not. In line with previous literature, participants
546 exhibited a Pavlovian bias in both responses and RTs (Algermissen et al., 2022; Swart et al., 2017),
547 with more and faster Go responses to Win than Avoid cues. On trials with high stakes (i.e., larger
548 rewards or punishments at stake), they slowed down, particularly for the two incongruent conditions
549 Go-to-Avoid and NoGo-to-Win. This response slowing was best described by high stakes prolonging
550 the non-decision time in a drift-diffusion model framework, particularly so for incongruent trials. This
551 finding is inconsistent with both hypotheses put forward in the introduction, i.e., high stakes
552 strengthening Pavlovian biases or high stakes motivating cognitive control to suppress them on
553 incongruent trials. In sum, higher stakes slow down response selection, but neither strengthen nor
554 weaken Pavlovian biases in responses. We propose two possible explanations for this (somewhat
555 surprising) result: response slowing under high stakes might reflect (flexibly recruited) cognitive
556 control, which is however ineffectively used, or it might reflect (automatic/ reflexive) positive condition
557 suppression, i.e., the suppression of locomotion by large immanent rewards as previously observed in
558 animal studies.

559 **No evidence for bias strengthening or bias suppression**

560 On trials with high stakes, participants took longer to make a Go response, but did not exhibit
561 any altered tendency for Go/ NoGo responses, i.e. no reduction or enhancement of Pavlovian biases.
562 Apart from the null effect on responses, RTs slowed down under high stakes, an effect that was highly
563 consistent across participants (Fig. 3E, F). These two findings are incompatible with the first hypothesis
564 posited, i.e., high stakes strengthening Pavlovian biases. Slowing (instead of speeding) of responses
565 under high rewards might appear quite surprising given a large body of literature showing higher
566 incentives to speed up responses (Fontanesi et al., 2019; Knutson et al., 2001; Luo et al., 2009; Pirrone,
567 Azab, Hayden, Stafford, & Marshall, 2018; Smith & Krajbich, 2018) and some evidence for larger PIT
568 effects for high compared to low value cues (Algermissen & den Ouden, 2023; Schad et al., 2020).
569 Notably, response slowing occurred for both appetitive and aversive cues, suggesting that the effect is
570 independent of cue valence and orthogonal to the Pavlovian biases. Note that 50% of trials were high
571 stake trials, arguing against the possibility of surprise (i.e., oddball effects) driving the response

572 slowing. High and low stake trials were visually very distinct, arguing against differences in processing
573 demands between both trial types. In sum, the size of Pavlovian biases in the Motivational Go/NoGo
574 Task appears to be unaffected by stake magnitude, which instead induced a response slowing orthogonal
575 to the biases.

576 Response slowing under high stakes might be partly compatible with the second hypothesis
577 (EVC), i.e., high stakes increasing cognitive control in order to suppress biases, given that heightened
578 cognitive control recruitment is often inferred from/ accompanied by prolonged reaction times (Frank,
579 2006; Shenhav et al., 2013; Wessel & Aron, 2017). Specifically, in line with our preregistered
580 hypothesis that high stakes increase cognitive control recruitment, response slowing was stronger on
581 motivationally incongruent trials on which Pavlovian biases had to be suppressed in order to execute
582 the correct response. This effect suggests that participants did distinguish the different cue conditions
583 with respect to whether they could benefit from increased cognitive control recruitment and prolonged
584 deliberation times (i.e., situations in which control could in theory change the emitted response) or not.
585 However, the increased deliberation time putatively afforded by cognitive control recruitment was
586 inconsequential for response selection, and the size of Pavlovian biases (in terms of the proportion of
587 Go responses for Win vs. Avoid cues) was unaltered under high stakes. One might thus conclude that
588 participants recruited additional cognitive control, but did not effectively use it to suppress their
589 Pavlovian biases when they were unhelpful.

590 An alternative explanation for response slowing under high stakes might be the phenomenon
591 of “choking under pressure”, i.e., the fear of failure in high-stakes situations inducing rumination and
592 thus decreasing performance (Beilock & Carr, 2001, 2005), an option we had considered in our pre-
593 registration. Choking under pressure predicts a pattern opposite to the second hypothesis (EVC), with
594 high stakes undermining cognitive control recruitment and leading to lower performance in incongruent
595 conditions. While the observed slowing of RTs could be interpreted as a kind of “choking under
596 pressure”, we did not observe corresponding performance decrements. Hence, this finding does not fall
597 under the phenomenon of “choking under pressure” as investigated in previous literature. In sum, these

598 results are most compatible with the idea of high stakes leading to increased cognitive control
599 recruitment, though without any consequences for response selection and accuracy.

600 **No evidence for a speed-accuracy tradeoff**

601 Past computational models have proposed mechanisms of how decision accuracy—which is
602 particularly warranted in high stakes situations—can be prioritized over speed by increasing decision
603 thresholds in an evidence accumulation framework (Bogacz et al., 2006). Such increased decision
604 bounds have been typically investigated in situations in which choice options are close in value and
605 thus eliciting cognitive conflict. Neuro-computational models suggest that such conflict is detected by
606 the anterior cingulate cortex and presupplementary motor area, which—via the hyperdirect pathway
607 involving the subthalamic nucleus—project to the globus pallidus and increase decision thresholds in
608 the basal ganglia action selection circuits, leading to a higher requirement for positive evidence to elicit
609 a response (Cavanagh et al., 2011; Forstmann et al., 2008; Frank, 2006; Frank et al., 2015; Wiecki &
610 Frank, 2013). This decision threshold adjustment will lead to a higher proportion of correct, but overall
611 slower responses. It is plausible that the same mechanism could lead to response caution in the context
612 of high-value cues. In fact, a series of recent studies found that cues indicating an upcoming choice
613 between high-value options (but not the presence of high-value options per se) slowed down of RTs,
614 which was best captured by a heightened decision threshold (Shevlin et al., 2022). However, in contrast,
615 the data of the present study were best explained by a model embodying prolonged non-decision times
616 rather than heightened response thresholds. It is thus unclear whether the same computational and neural
617 mechanisms proposed for implementing speed-accuracy tradeoffs are also responsible for the response
618 slowing observed in this data. Future studies using neuroimaging of cortical and subcortical activity
619 (Algermissen et al., 2022) and instructions to prioritize speed or accuracy during the task (Forstmann
620 et al., 2008) while simultaneously manipulating stakes could shed light on shared vs. separate neural
621 mechanisms.

622 **Response slowing as positive conditioned suppression**

623 Another possible interpretation of our findings is that the response slowing under large stake
624 magnitudes is an instance of positive conditioned suppression as previously reported in rodents (Azrin

625 & Hake, 1969; Marshall, Halbout, Munson, Hutson, & Ostlund, 2023; Van Dyne, 1971). In positive
626 conditioned suppression, cues signaling the immanent receipt of a reward suppress responding.
627 Specifically, a cue announcing an immanent reward suppresses exploratory behavior that would move
628 the animal away from a food site, and instead invigorates and prolongs engagement with the site of
629 reward delivery until the reward is obtained (Marshall, Munson, Maidment, & Ostlund, 2020).
630 However, this suppression can extend backwards in time such that it even affects the instrumental
631 response required to obtain the reward (i.e., a lever press). A recent study found small rewards to
632 invigorate responding in line with classical PIT findings (Marshall et al., 2023). However, large rewards
633 suppressed instrumental lever pressing and diminished PIT effects, suggestive of positive conditioned
634 suppression interfering with PIT in a way similar of our findings.

635 One speculation on the adaptive nature of this phenomenon is that it may prevent agents to
636 become distracted by other reward opportunities and forget to collect the reward they previously worked
637 for (Timberlake, Wahl, & King, 1982). Notably, the prolongation of RTs in the present data was
638 particularly strong for motivationally incongruent cues, which perhaps argues against a purely
639 automatic, “reflexive” nature of the observed effect of stake magnitude on RTs (such as positive
640 conditioned suppression), and instead in favor of an adaptive effect that is (at least partially) sensitive
641 to task requirements. It is thus possible that both (automatic) positive conditioned suppression and
642 (voluntary) heightened cognitive control recruitment triggered by motivational conflict are present, or
643 that positive conditioned suppression is (partially) a consequence of cognitive control recruitment.
644 Future studies could test whether the slowing induced by high stakes is sensitive to the temporal delay
645 between response execution and outcome delivery, which would argue for interference between reward
646 collection and response selection as the cause of slowing (Delamater & Holland, 2008; Marshall et al.,
647 2023; Marshall & Ostlund, 2018; Meltzer & Hamm, 1978; Miczek & Grossman, 1971).

648 Furthermore, conditioned suppression has yet not been studied in the context of avoiding
649 aversive outcomes. Slowing induced by conditioned suppression will look highly similar to slowing
650 induced by the Pavlovian bias itself. In our data, the finding that effects of action-valence congruency
651 (i.e. Pavlovian bias) and stake magnitude on RTs were additive suggests independent mechanisms.

652 Future research might try to disentangle these two effects further by using an “escape” context in which
653 participants must select actions to terminate an ongoing punishment (e.g. loud noise), which typically
654 inverts the Pavlovian bias and leads to an increased tendency towards action (Millner et al., 2017).
655 Varying the punishment magnitude in such a context could potentially elucidate joint or independent
656 contributions of Pavlovian biases and conditioned suppression on RTs.

657 **Normative aspects**

658 The presented results suggest that high stakes do not strengthen or weaken Pavlovian biases per
659 se; rather, they globally slow or pause behavior. This slowing down can be adaptive in high threat
660 situations in which response postponement mimics nonresponding, similar to freezing itself (Bach,
661 2015), although in our data, the slowing did not affect participants’ eventual propensity to execute a Go
662 response. This slowing might also be adaptive from the perspective of positive conditioned suppression
663 in focusing an agent on reward collection and consumption rather than exploring other options in the
664 meantime (Marshall et al., 2023). The ability to inhibit behavior and wait for rewards has been proposed
665 to be serotonergic in nature, as serotonin is likely implicated in mediating aversive inhibition (Crockett,
666 Clark, Apergis-Schoute, Morein-Zamir, & Robbins, 2012; Crockett, Clark, & Robbins, 2009; Geurts et
667 al., 2013b). Indeed, serotonin depletion has been shown to abolish the slowing observed under high
668 reward stakes (Bari & Robbins, 2013; den Ouden et al., 2015; Soubrié, 1986), while the activation of
669 serotonergic neurons facilitates waiting for rewards (K. Miyazaki, Miyazaki, & Doya, 2011; K.
670 Miyazaki et al., 2020; K. W. Miyazaki et al., 2014) and persistence in foraging (Lottem et al., 2018).
671 Future research should explicitly test the putatively serotonergic nature of high stakes-induced response
672 slowing in the Motivational Go/NoGo Task in particular and of positive conditioned suppression, more
673 generally.

674 **Limitations and relations to other stakes manipulations**

675 A limitation of the current study is that high stakes were explicitly signaled via a red circle
676 around the task cue. In this way, the task mimicked situations in which high stakes can be inferred
677 directly from simple visual features, e.g. when telling apart a lion from a spider. However, it does not
678 mimic situations in which high value must be inferred indirectly from past experiences or by combining

679 set of features, e.g., in detecting a good bargain house or car. In the context of the Motivational
680 Go/NoGo Task, stakes were irrelevant for selection the optimal action, and evidence from a similar task
681 (Algermissen & den Ouden, 2023) suggests that participants ignore differences in outcome value when
682 learning about the optimal action. Hence, stakes might only play a role when explicitly signaled or
683 easily perceivable from the environment, but not when they have to be inferred from past experiences.
684 This is an important consideration for task designs that might explain the mixed literature on stakes
685 effects in PIT tasks. Finally, the presented finding mimics cases where “high stakes” describes the entire
686 situation rather than a single option (Shevlin et al., 2022), but is unlike cases where only a single option
687 is more valuable and dominates all other options.

688 Another limitation might be that stakes were not varied in a continuous fashion, but
689 categorically as two discrete levels. Again, it might be plausible that agents represent situations (e.g.
690 trials) as overall “high stakes” or not, irrespective of the particular value of single options (Shevlin et
691 al., 2022). Varying the stakes magnitude in a continuous fashion would increase processing demands
692 and thus already slow down responses due to perceptual (irrespective of additional decision) difficulty.
693 Furthermore, participants might subjectively recode stakes levels relative to the mean stake level,
694 representing low rewards as disappointing and thus akin to punishments, while perceiving low
695 punishments as a relief and thus akin to rewards (Klein, Ullsperger, & Jocham, 2017; Palminteri,
696 Khamassi, Joffily, & Coricelli, 2015). These considerations support the ecological validity of
697 dichotomizing stakes into high and low levels. However, it remains to be empirically tested whether
698 continuous stakes levels lead to similar or different effects.

699 Conclusion

700 In sum, while possibilities to gain rewards/ avoid punishments induce Pavlovian biases,
701 increasing the stakes of these prospects does not alter the strength of the bias. However, high stakes
702 motivate humans to slow down their responses. One interpretation is that this slowing is adaptive in
703 allowing time for conflict detection and cognitive control recruitment in case motivational biases have
704 to be suppressed. However, the slowing is not associated with changes in response selection, i.e., also
705 not with the degree to which participants suppress their Pavlovian biases when these are unhelpful,

706 suggesting that humans do not use this additional time effectively. An alternative interpretation is that
707 prolonged reaction times reflect positive conditioned suppression, i.e. attraction by the reward value
708 that interferes with action selection itself as previously observed in rodents. Taken together, this study
709 suggests that high stakes might have a similar effect in both humans and rodents in the context of
710 Pavlovian/ instrumental interactions on action selection.

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

References

729 Aarts, E., Verhage, M., Veenvliet, J. V., Dolan, C. V., & van der Sluis, S. (2014). A solution to
730 dependency: Using multilevel analysis to accommodate nested data. *Nature Neuroscience*,
731 17(4), 491–496. doi: 10.1038/nn.3648

732 Algermissen, J., & den Ouden, H. E. M. (2023). Goal-directed recruitment of Pavlovian biases
733 through selective visual attention. *Journal of Experimental Psychology: General*, 152(10),
734 2941–2956. doi: 10.1037/xge0001425

735 Algermissen, J., Swart, J. C., Scheeringa, R., Cools, R., & den Ouden, H. E. M. (2022). Striatal
736 BOLD and midfrontal theta power express motivation for action. *Cerebral Cortex*, 32(14),
737 2924–2942. doi: 10.1093/cercor/bhab391

738 Azrin, N. H., & Hake, D. F. (1969). Positive conditioned suppression: Conditioned suppression using
739 positive reinforcers as the unconditioned stimuli. *Journal of the Experimental Analysis of
740 Behavior*, 12(1), 167–173. doi: 10.1901/jeab.1969.12-167

741 Bach, D. R. (2015). Anxiety-like behavioural inhibition is normative under environmental threat-
742 reward correlations. *PLOS Computational Biology*, 11(12), e1004646. doi:
743 10.1371/journal.pcbi.1004646

744 Ballard, I. C., & McClure, S. M. (2019). Joint modeling of reaction times and choice improves
745 parameter identifiability in reinforcement learning models. *Journal of Neuroscience Methods*,
746 317(June 2018), 37–44. doi: 10.1016/j.jneumeth.2019.01.006

747 Bari, A., & Robbins, T. W. (2013). Inhibition and impulsivity: Behavioral and neural basis of
748 response control. *Progress in Neurobiology*, 108, 44–79. doi:
749 10.1016/j.pneurobio.2013.06.005

750 Baribault, B., & Collins, A. G. E. (2023). Troubleshooting Bayesian cognitive models. *Psychological
751 Methods*. doi: 10.1037/met0000554

752 Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory
753 hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. doi:
754 10.1016/j.jml.2012.11.001

755 Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using
756 lme4. *Journal of Statistical Software*, 67, 1–48. doi: 10.18637/jss.v067.i01

757 Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking
758 under pressure? *Journal of Experimental Psychology: General*, 130(4), 701–725. doi:
759 10.1037/0096-3445.130.4.701

760 Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and “choking
761 under pressure” in math. *Psychological Science*, 16(2), 101–105. doi: 10.1111/j.0956-
762 7976.2005.00789.x

763 Blurton, S. P., Kesselmeier, M., & Gondan, M. (2012). Fast and accurate calculations for cumulative
764 first-passage time distributions in Wiener diffusion models. *Journal of Mathematical
765 Psychology*, 56(6), 470–475. doi: 10.1016/j.jmp.2012.09.002

766 Boehler, C. N., Hopf, J.-M., Stoppel, C. M., & Krebs, R. M. (2012). Motivating inhibition – reward
767 prospect speeds up response cancellation. *Cognition*, 125(3), 498–503. doi:
768 10.1016/j.cognition.2012.07.018

769 Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal
770 decision making: A formal analysis of models of performance in two-alternative forced-
771 choice tasks. *Psychological Review*, 113(4), 700–765. doi: 10.1037/0033-295X.113.4.700

772 Boureau, Y.-L., & Dayan, P. (2011). Opponency revisited: Competition and cooperation between
773 dopamine and serotonin. *Neuropsychopharmacology*, 36(1), 74–97. doi:
774 10.1038/npp.2010.151

775 Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist*, 16(11),
776 681–684. doi: 10.1037/h0040090

777 Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A.
778 (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1).
779 doi: 10.18637/jss.v076.i01

780 Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q. J. M., & Frank, M. J. (2013). Frontal theta
781 overrides Pavlovian learning biases. *Journal of Neuroscience*, 33(19), 8541–8548. doi:
782 10.1523/JNEUROSCI.5754-12.2013

783 Cavanagh, J. F., Wiecki, T. V., Cohen, M. X., Figueroa, C. M., Samanta, J., Sherman, S. J., & Frank,
784 M. J. (2011). Subthalamic nucleus stimulation reverses mediofrontal influence over decision
785 threshold. *Nature Neuroscience*, 14(11), 1462–1467. doi: 10.1038/nn.2925

786 Chen, K., Garbusow, M., Sebold, M., Kuitunen-Paul, S., Smolka, M. N., Huys, Q. J. M., ... Heinz, A.
787 (2022). Alcohol approach bias is associated with both behavioral and neural Pavlovian-to-
788 instrumental transfer effects in alcohol-dependent patients. *Biological Psychiatry Global
789 Open Science*. doi: 10.1016/j.bpsgos.2022.03.014

790 Chen, K., Garbusow, M., Sebold, M., Zech, H. G., Zimmermann, U., & Heinz, A. (2022). Automatic
791 approach behaviors in alcohol dependence: Does a cognitive bias modification training affect
792 Pavlovian-to-instrumental transfer effects? *Neuropsychobiology*, 81(5), 387–402. doi:
793 10.1159/000526805

794 Chen, K., Schlagenhauf, F., Sebold, M., Kuitunen-Paul, S., Chen, H., Huys, Q. J. M., ... Garbusow,
795 M. (2022). The association of non-drug-related Pavlovian-to-instrumental transfer effect in
796 nucleus accumbens with relapse in alcohol dependence: A replication. *Biological Psychiatry*.
797 doi: 10.1016/j.biopsych.2022.09.017

798 Chiew, K. S., & Braver, T. S. (2014). Dissociable influences of reward motivation and positive
799 emotion on cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, 14(2), 509–
800 529. doi: 10.3758/s13415-014-0280-0

801 Chiew, K. S., & Braver, T. S. (2016). Reward favors the prepared: Incentive and task-informative
802 cues interact to enhance attentional control. *Journal of Experimental Psychology: Human
803 Perception and Performance*, 42(1), 52–66. doi: 10.1037/xhp0000129

804 Cohen, J. D. (2017). Cognitive control. Core constructs and current considerations. In T. Egner (Ed.),
805 *The Wiley Handbook of Cognitive Control* (pp. 1–28). Chichester, UK: John Wiley & Sons,
806 Ltd. doi: 10.1002/9781118920497.ch1

807 Colaizzi, J. M., Flagel, S. B., Joyner, M. A., Gearhardt, A. N., Stewart, J. L., & Paulus, M. P. (2020).
808 Mapping sign-tracking and goal-tracking onto human behaviors. *Neuroscience &
809 Biobehavioral Reviews*, 111(July 2019), 84–94. doi: 10.1016/j.neubiorev.2020.01.018

810 Collins, A. G. E., & Cockburn, J. (2020). Beyond dichotomies in reinforcement learning. *Nature Reviews Neuroscience*. doi: 10.1038/s41583-020-0355-6

811 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(10), 2244–2252. doi: 10.1038/npp.2012.75

812 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(38), 11993–11999. doi: 10.1523/JNEUROSCI.2513-09.2009

813 Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8(12), 1704–1711. doi: 10.1038/nn1560

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

815 Delamater, A. R., & Holland, P. C. (2008). The influence of CS-US interval on several different indices of learning in appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, 34(2), 202–222. doi: 10.1037/0097-7403.34.2.202

816 den Ouden, H. E. M., Swart, J. C., Schmidt, K., Fekkes, D., Geurts, D. E. M., & Cools, R. (2015). Acute serotonin depletion releases motivated inhibition of response vigour. *Psychopharmacology*, 232(7), 1303–1312. doi: 10.1007/s00213-014-3762-4

817 DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, 93(5), 880–896. doi: 10.1037/0022-3514.93.5.880

818 Dickinson, A., & Balleine, B. (1994). Motivational control of goal-directed action. *Animal Learning & Behavior*, 22(1), 1–18. doi: 10.3758/BF03199951

819 Dixon, M. L., & Christoff, K. (2012). The decision to engage cognitive control is driven by expected reward-value: Neural and behavioral evidence. *PLoS ONE*, 7(12), e51637. doi: 10.1371/journal.pone.0051637

838 Dorfman, H. M., & Gershman, S. J. (2019). Controllability governs the balance between Pavlovian
839 and instrumental action selection. *Nature Communications*, 10(1), 5826. doi: 10.1038/s41467-
840 019-13737-7

841 Estes, W. K. (1943). Discriminative conditioning. I. A discriminative property of conditioned
842 anticipation. *Journal of Experimental Psychology*, 32(2), 150–155. doi: 10.1037/h0058316

843 Estes, W. K. (1948). Discriminative conditioning. II. Effects of a Pavlovian conditioned stimulus
844 upon a subsequently established operant response. *Journal of Experimental Psychology*,
845 38(2), 173–177. doi: 10.1037/h0057525

846 Fitzpatrick, A. L., Rapp, S. R., Luchsinger, J., Hill-Briggs, F., Alonso, A., Gottesman, R., ... Seeman,
847 T. (2015). Sociodemographic correlates of cognition in the Multi-Ethnic Study of
848 Atherosclerosis (MESA). *The American Journal of Geriatric Psychiatry*, 23(7), 684–697. doi:
849 10.1016/j.jagp.2015.01.003

850 Flagel, S. B., & Robinson, T. E. (2017). Neurobiological basis of individual variation in stimulus-
851 reward learning. *Current Opinion in Behavioral Sciences*, 13, 178–185. doi:
852 10.1016/j.cobeha.2016.12.004

853 Flagel, S. B., Watson, S. J., Robinson, T. E., & Akil, H. (2007). Individual differences in the
854 propensity to approach signals vs goals promote different adaptations in the dopamine system
855 of rats. *Psychopharmacology*, 191(3), 599–607. doi: 10.1007/s00213-006-0535-8

856 Fontanesi, L., Gluth, S., Spektor, M. S., & Rieskamp, J. (2019). A reinforcement learning diffusion
857 decision model for value-based decisions. *Psychonomic Bulletin & Review*. doi:
858 10.3758/s13423-018-1554-2

859 Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., &
860 Wagenmakers, E.-J. (2008). Striatum and pre-SMA facilitate decision-making under time
861 pressure. *Proceedings of the National Academy of Sciences*, 105(45), 17538–17542. doi:
862 10.1073/pnas.0805903105

863 Frank, M. J. (2006). Hold your horses: A dynamic computational role for the subthalamic nucleus in
864 decision making. *Neural Networks*, 19(8), 1120–1136. doi: 10.1016/j.neunet.2006.03.006

865 Frank, M. J., Gagne, C., Nyhus, E., Masters, S., Wiecki, T. V., Cavanagh, J. F., & Badre, D. (2015).
866 fMRI and EEG predictors of dynamic decision parameters during human reinforcement
867 learning. *Journal of Neuroscience*, 35(2), 485–494. doi: 10.1523/JNEUROSCI.2036-14.2015

868 Fröber, K., & Dreisbach, G. (2016). How performance (non-)contingent reward modulates cognitive
869 control. *Acta Psychologica*, 168, 65–77. doi: 10.1016/j.actpsy.2016.04.008

870 Garbusow, M., Nebe, S., Sommer, C., Kuitunen-Paul, S., Sebold, M., Schad, D. J., ... Heinz, A.
871 (2019). Pavlovian-to-instrumental transfer and alcohol consumption in young male social
872 drinkers: Behavioral, neural and polygenic correlates. *Journal of Clinical Medicine*, 8(8),
873 1188. doi: 10.3390/jcm8081188

874 Garbusow, M., Schad, D. J., Sebold, M., Friedel, E., Bernhardt, N., Koch, S. P., ... Heinz, A. (2016).
875 Pavlovian-to-instrumental transfer effects in the nucleus accumbens relate to relapse in
876 alcohol dependence. *Addiction Biology*, 21(3), 719–731. doi: 10.1111/adb.12243

877 Garofalo, S., & di Pellegrino, G. (2015). Individual differences in the influence of task-irrelevant
878 Pavlovian cues on human behavior. *Frontiers in Behavioral Neuroscience*, 9. Retrieved from
879 <https://www.frontiersin.org/articles/10.3389/fnbeh.2015.00163>

880 Geurts, D. E. M., Huys, Q. J. M., den Ouden, H. E. M., & Cools, R. (2013a). Aversive Pavlovian
881 control of instrumental behavior in humans. *Journal of Cognitive Neuroscience*, 25(9), 1428–
882 1441. doi: 10.1162/jocn_a_00425

883 Geurts, D. E. M., Huys, Q. J. M., den Ouden, H. E. M., & Cools, R. (2013b). Serotonin and aversive
884 Pavlovian control of instrumental behavior in humans. *Journal of Neuroscience*, 33(48),
885 18932–18939. doi: 10.1523/JNEUROSCI.2749-13.2013

886 Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental
887 Psychology: General*, 136(3), 389–413. doi: 10.1037/0096-3445.136.3.389

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

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

893 Hearst, E., & Jenkins, H. (1974). *The stimulus-reinforcer relation and directed action*. Austin, TX:
894 Monograph of the Psychonomic Society. doi: 10.2815/64877

895 Hershberger, W. A. (1986). An approach through the looking-glass. *Animal Learning & Behavior*,
896 14(4), 443–451. doi: 10.3758/BF03200092

897 Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems
898 perspective. *Perspectives on Psychological Science*, 4(2), 162–176. doi: 10.1111/j.1745-
899 6924.2009.01116.x

900 Jenkins, H. M., & Moore, B. R. (1973). The form of the auto-shaped response with food or water
901 reinforcers. *Journal of the Experimental Analysis of Behavior*, 20(2), 163–181. doi:
902 10.1901/jeab.1973.20-163

903 Klein, T. A., Ullsperger, M., & Jocham, G. (2017). Learning relative values in the striatum induces
904 violations of normative decision making. *Nature Communications*, 8(May 2016), 16033. doi:
905 10.1038/ncomms16033

906 Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary
907 reward selectively recruits Nucleus Accumbens. *Journal of Neuroscience*, 21(16), RC159–
908 RC159. doi: 10.1523/JNEUROSCI.21-16-j0002.2001

909 Kraemer, P. M., Fontanesi, L., Spektor, M. S., & Gluth, S. (2021). Response time models separate
910 single- and dual-process accounts of memory-based decisions. *Psychonomic Bulletin &
911 Review*, 28(1), 304–323. doi: 10.3758/s13423-020-01794-9

912 Krebs, R. M., Boehler, C. N., & Woldorff, M. G. (2010). The influence of reward associations on
913 conflict processing in the Stroop task. *Cognition*, 117(3), 341–347. doi:
914 10.1016/j.cognition.2010.08.018

915 Lieder, F., Shenhav, A., Musslick, S., & Griffiths, T. L. (2018). Rational metareasoning and the
916 plasticity of cognitive control. *PLOS Computational Biology*, 14(4), e1006043. doi:
917 10.1371/journal.pcbi.1006043

918 LoLordo, V. M., McMillan, J. C., & Riley, A. L. (1974). The effects upon food-reinforced pecking
919 and treadle-pressing of auditory and visual signals for response-independent food. *Learning
920 and Motivation*, 5(1), 24–41. doi: 10.1016/0023-9690(74)90035-6

921 Lottem, E., Banerjee, D., Vertechi, P., Sarra, D., Lohuis, M. O., & Mainen, Z. F. (2018). Activation of
922 serotonin neurons promotes active persistence in a probabilistic foraging task. *Nature Communications*, 9(1), 1000. doi: 10.1038/s41467-018-03438-y

923 Lovibond, P. F. (1983). Facilitation of instrumental behavior by a Pavlovian appetitive conditioned
924 stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, 9(3), 225–247.
925 doi: 10.1037/0097-7403.9.3.225

926 Luo, S., Ainslie, G., Giragosian, L., & Monterosso, J. R. (2009). Behavioral and neural evidence of
927 incentive bias for immediate rewards relative to preference-matched delayed rewards. *Journal of Neuroscience*, 29(47), 14820–14827. doi: 10.1523/JNEUROSCI.4261-09.2009

928 Mahlberg, J., Seabrooke, T., Weidemann, G., Hogarth, L., Mitchell, C. J., & Moustafa, A. A. (2021).
929 Human appetitive Pavlovian-to-instrumental transfer: A goal-directed account. *Psychological Research*, 85(2), 449–463. doi: 10.1007/s00426-019-01266-3

930 Marshall, A. T., Halbout, B., Munson, C. N., Hutson, C., & Ostlund, S. B. (2023). Flexible control of
931 Pavlovian-instrumental transfer based on expected reward value. *Journal of Experimental Psychology: Animal Learning and Cognition*, 49, 14–30. doi: 10.1037/xan0000348

932 Marshall, A. T., Munson, C. N., Maidment, N. T., & Ostlund, S. B. (2020). Reward-predictive cues
933 elicit excessive reward seeking in adolescent rats. *Developmental Cognitive Neuroscience*, 45, 100838. doi: 10.1016/j.dcn.2020.100838

934 Marshall, A. T., & Ostlund, S. B. (2018). Repeated cocaine exposure dysregulates cognitive control
935 over cue-evoked reward-seeking behavior during Pavlovian-to-instrumental transfer. *Learning & Memory*, 25(9), 399–409. doi: 10.1101/lm.047621.118

936 Meltzer, D., & Hamm, R. J. (1978). Differential conditioning of conditioned enhancement and
937 positive conditioned suppression. *Bulletin of the Psychonomic Society*, 11(1), 29–32. doi: 10.3758/BF03336757

938 Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of
939 willpower. *Psychological Review*, 106(1), 3–19. doi: 10.1037/0033-295X.106.1.3

947 Miczek, K. A., & Grossman, S. P. (1971). Positive conditioned suppression: Effects of CS duration.

948 *Journal of the Experimental Analysis of Behavior*, 15(2), 243–247. doi:
949 10.1901/jeab.1971.15-243

950 Miletic, S., Boag, R. J., & Forstmann, B. U. (2020). Mutual benefits: Combining reinforcement
951 learning with sequential sampling models. *Neuropsychologia*, 136, 107261. doi:
952 10.1016/j.neuropsychologia.2019.107261

953 Millner, A. J., Gershman, S. J., Nock, M. K., & den Ouden, H. E. M. (2017). Pavlovian control of
954 escape and avoidance. *Journal of Cognitive Neuroscience*, 26(3), 1–12. doi:
955 10.1162/jocn_a_01224

956 Mischel, W., & Ebbesen, E. B. (1970). Attention in delay of gratification. *Journal of Personality and
957 Social Psychology*, 16(2), 329–337. doi: 10.1037/h0029815

958 Miyazaki, K., Miyazaki, K. W., & Doya, K. (2011). Activation of dorsal raphe serotonin neurons
959 underlies waiting for delayed rewards. *Journal of Neuroscience*, 31(2), 469–479. doi:
960 10.1523/JNEUROSCI.3714-10.2011

961 Miyazaki, K., Miyazaki, K. W., Sivori, G., Yamanaka, A., Tanaka, K. F., & Doya, K. (2020).
962 Serotonergic projections to the orbitofrontal and medial prefrontal cortices differentially
963 modulate waiting for future rewards. *Science Advances*, 6(48), eabc7246. doi:
964 10.1126/sciadv.abc7246

965 Miyazaki, K. W., Miyazaki, K., Tanaka, K. F., Yamanaka, A., Takahashi, A., Tabuchi, S., & Doya, K.
966 (2014). Optogenetic activation of dorsal raphe serotonin neurons enhances patience for future
967 rewards. *Current Biology*, 24(17), 2033–2040. doi: 10.1016/j.cub.2014.07.041

968 O'Doherty, J. P., Cockburn, J., & Pauli, W. M. (2017). Learning, reward, and decision making.
969 *Annual Review of Psychology*, 68(1), 73–100. doi: 10.1146/annurev-psych-010416-044216

970 Palminteri, S., Khamassi, M., Joffily, M., & Coricelli, G. (2015). Contextual modulation of value
971 signals in reward and punishment learning. *Nature Communications*, 6, 8096. doi:
972 10.1038/ncomms9096

973 Palminteri, S., Wyart, V., & Koechlin, E. (2017). The importance of falsification in computational
974 cognitive modeling. *Trends in Cognitive Sciences*, 21(6), 425–433. doi:
975 10.1016/j.tics.2017.03.011

976 Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness
977 Scale. *Journal of Clinical Psychology*, 51(6), 768–774. doi: 10.1002/1097-
978 4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1

979 Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in
980 reinforcement learning. *Psychonomic Bulletin & Review*, 24(4), 1234–1251. doi:
981 10.3758/s13423-016-1199-y

982 Pirrone, A., Azab, H., Hayden, B. Y., Stafford, T., & Marshall, J. A. R. (2018). Evidence for the
983 speed–value trade-off: Human and monkey decision making is magnitude sensitive. *Decision*,
984 5(2), 129–142. doi: 10.1037/dec0000075

985 R Core Team. (2022). R: A language and environment for statistical computing. In *R Foundation for
986 Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved
987 from www.R-project.org

988 Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108. doi:
989 10.1037/0033-295X.85.2.59

990 Ratcliff, R., Huang-Pollock, C., & McKoon, G. (2018). Modeling individual differences in the go/no-
991 go task with a diffusion model. *Decision*, 5(1), 42–62. doi: 10.1037/dec0000065

992 Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory: Relationships between
993 Pavlovian conditioning and instrumental learning. *Psychological Review*, 74(3), 151–182.
994 doi: 10.1037/h0024475

995 Schad, D. J., Rapp, M. A., Garbusow, M., Nebe, S., Sebold, M., Obst, E., ... Huys, Q. J. M. (2020).
996 Dissociating neural learning signals in human sign- and goal-trackers. *Nature Human
997 Behaviour*, 4(2), 201–214. doi: 10.1038/s41562-019-0765-5

998 Schwartz, B. (1976). Positive and negative conditioned suppression in the pigeon: Effects of the locus
999 and modality of the CS. *Learning and Motivation*, 7(1), 86–100. doi: 10.1016/0023-
1000 9690(76)90019-9

1001 Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative
1002 theory of anterior cingulate cortex function. *Neuron*, 79(2), 217–240. doi:
1003 10.1016/j.neuron.2013.07.007

1004 Shevlin, B. R. K., Smith, S. M., Hausfeld, J., & Krajbich, I. (2022). High-value decisions are fast and
1005 accurate, inconsistent with diminishing value sensitivity. *Proceedings of the National
1006 Academy of Sciences*, 119(6), e2101508119. doi: 10.1073/pnas.2101508119

1007 Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II.
1008 Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84(2),
1009 127–190. doi: 10.1037/0033-295X.84.2.127

1010 Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2018). *afex: Analysis of factorial experiments*.
1011 Retrieved from <https://cran.r-project.org/package=afex>

1012 Smith, S. M., & Krajbich, I. (2018). Gaze amplifies value in decision making. *Psychological Science*,
1013 095679761881052. doi: 10.1177/0956797618810521

1014 Sommer, C., Birkenstock, J., Garbusow, M., Obst, E., Schad, D. J., Bernhardt, N., ... Zimmermann,
1015 U. S. (2020). Dysfunctional approach behavior triggered by alcohol-unrelated Pavlovian cues
1016 predicts long-term relapse in alcohol dependence. *Addiction Biology*, 25(1), 1–10. doi:
1017 10.1111/adb.12703

1018 Sommer, C., Garbusow, M., Jünger, E., Pooseh, S., Bernhardt, N., Birkenstock, J., ... Zimmermann,
1019 U. S. (2017). Strong seduction: Impulsivity and the impact of contextual cues on instrumental
1020 behavior in alcohol dependence. *Translational Psychiatry*, 7(8), e1183–e1183. doi:
1021 10.1038/tp.2017.158

1022 Soubrié, P. (1986). Reconciling the role of central serotonin neurons in human and animal behavior.
1023 *Behavioral and Brain Sciences*, 9(2), 319–335. doi: 10.1017/S0140525X00022871

1024 Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior.
1025 *Personality and Social Psychology Review*, 8(3), 220–247. doi:
1026 10.1207/s15327957pspr0803_1

1027 Swart, J. C., Frank, M. J., Määttä, J. I., Jensen, O., Cools, R., & den Ouden, H. E. M. (2018). Frontal
1028 network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in
1029 motivated action. *PLOS Biology*, 16(10), e2005979. doi: 10.1371/journal.pbio.2005979

1030 Swart, J. C., Froböse, M. I., Cook, J. L., Geurts, D. E., Frank, M. J., Cools, R., & den Ouden, H. E.
1031 (2017). Catecholaminergic challenge uncovers distinct Pavlovian and instrumental
1032 mechanisms of motivated (in)action. *eLife*, 6, e22169. doi: 10.7554/eLife.22169

1033 Timberlake, W., Wahl, G., & King, D. (1982). Stimulus and response contingencies in the
1034 misbehavior of rats. *Journal of Experimental Psychology. Animal Behavior Processes*, 8(1),
1035 62–85.

1036 Van Dyne, G. C. (1971). Conditioned suppression with a positive us in the rat. *Journal of
1037 Comparative and Physiological Psychology*, 77(1), 131–135. doi: 10.1037/h0031576

1038 Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out
1039 cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. doi:
1040 10.1007/s11222-016-9696-4

1041 Wabersich, D., & Vandekerckhove, J. (2014). The RWiener Package: An R Package Providing
1042 Distribution Functions for the Wiener Diffusion Model. *The R Journal*, 6(June), 49–56.

1043 Wessel, J. R., & Aron, A. R. (2017). On the globality of motor suppression: Unexpected events and
1044 their influence on behavior and cognition. *Neuron*, 93(2), 259–280. doi:
1045 10.1016/j.neuron.2016.12.013

1046 Westbrook, A., Frank, M. J., & Cools, R. (2021). A mosaic of cost–benefit control over cortico-
1047 striatal circuitry. *Trends in Cognitive Sciences*, 25(8), 710–721. doi:
1048 10.1016/j.tics.2021.04.007

1049 Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York, NY: Springer-Verlag.
1050 Retrieved from <https://ggplot2.tidyverse.org>

1051 Wiecki, T. V., & Frank, M. J. (2013). A computational model of inhibitory control in frontal cortex
1052 and basal ganglia. *Psychological Review*, 120(2), 329–355. doi: 10.1037/a0031542

1053 Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-
1054 Diffusion Model in Python. *Frontiers in Neuroinformatics*, 7(August), 1–10. doi:
1055 10.3389/fninf.2013.00014

1056 Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of
1057 behavioral data. *eLife*, 8, 1–35. doi: 10.7554/eLife.49547

1058

Supplemental Material S01: Overview results mixed-effects regression models

3 Here, we report an overview over all major statistical results reported in the main text and the
4 supplementary material. For details on how mixed-effects regression were performed, see the Methods
5 section of the main text.

Model ID	DV	IV	b	SE	$\chi^2(1)$	p
1	Response	Required action	1.441	0.096	87.873	< .001
		Valence	0.749	0.072	59.587	< .001
		Required action x cue valence	0.019	0.060	0.093	.760
2	RT	Required action	-0.109	0.019	27.494	< .001
		Valence	-0.191	0.019	59.204	< .001
		Required action x cue valence	0.031	0.015	4.384	.036
3	Accuracy	Congruency	0.600	0.052	67.867	< .001
		Stakes	-0.026	0.020	1.430	.232
		Congruency x Stakes	-0.007	0.020	0.094	.759
4	RT	Congruency	-0.131	0.015	49.546	< .001
		Stakes	0.072	0.010	33.702	< .001
		Congruency x Stakes	-0.019	0.010	3.856	.049

Table S01. Overview of the results from all mixed-effects regression models reported the main text of the manuscript. Featuring data from $N = 54$ participants, trial with RTs < 0.300 sec. are excluded from RT analyses.

6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

Supplemental Material S02: Overview results mixed-effects regression models on reaction times on all trials

Model ID	DV	IV	b	SE	$\chi^2(1)$	p
1	RT	Required action	-0.103	0.019	23.936	< .001
		Valence	-0.183	0.019	53.550	< .001
		Required action x cue valence	-0.039	0.015	6.138	.013
2	RT	Congruency	-0.133	0.014	51.704	< .001
		Stakes	0.070	0.011	31.210	< .001
		Congruency x Stakes	-0.019	0.010	3.982	.046

Table S02. Overview of RT regression models from $N = 54$ participants when including all trials (also those with RTs < 0.3 sec.).

26

27

28

29

30

31

32

33

54

55

50

87

18

41

42

43

44

45

46

47

48

49

50

Supplemental Material S03: Overview results mixed-effects regression models including additional participant

Model ID	DV	IV	b	SE	$\chi^2(1)$	p
1	Response	Required action	1.417	0.098	86.250	< .001
		Valence	0.736	0.071	59.174	< .001
		Required action x cue valence	0.019	0.059	0.097	.756
2	RT	Required action	-0.105	0.018	26.984	< .001
		Valence	-0.189	0.018	60.877	< .001
		Required action x cue valence	-0.033	0.015	4.824	.028
3	Accuracy	Congruency	0.591	0.052	67.189	< .001
		Stakes	-0.025	0.019	1.365	0.243
		Congruency x Stakes	-0.009	0.020	0.169	0.681
4	RT	Congruency	-0.130	0.014	50.997	< .001
		Stakes	0.071	0.010	34.566	< .001
		Congruency x Stakes	-0.018	0.009	3.527	.060

Table S03. Overview of all regression models when including data from all $N = 55$ participants (also the one participant excluded from analyses reported in the main text for not performing above chance level).

54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75

76 **Supplemental Material S04: Overview response, accuracy, and RT**
77 **means and standard deviations per condition**

Responses

Req. Act.	Go	Go	NoGo	NoGo
Valence	Win	Avoid	Win	Avoid
Mean	0.888	0.771	0.475	0.194
SD	0.140	0.111	0.261	0.107

Table S04. Means and standard deviations of Go/NoGo responses across participants per required action x valence condition.

78

Responses

Req. Act.	Go	Go	Go	Go	NoGo	NoGo	NoGo	NoGo
Valence	Win	Win	Avoid	Avoid	Win	Win	Avoid	Avoid
Stakes	High	Low	High	Low	High	Low	High	Low
Mean	0.883	0.893	0.767	0.774	0.477	0.474	0.200	0.187
SD	0.151	0.138	0.125	0.114	0.253	0.276	0.116	0.109

Table S05. Means and standard deviations of Go/NoGo responses across participants per required action x valence x stakes condition.

79

Accuracy

Req. Act.	Go	Go	NoGo	NoGo
Valence	Win	Avoid	Win	Avoid
Mean	0.888	0.771	0.525	0.806
SD	0.140	0.111	0.261	0.107

Table S06. Means and standard deviations of accuracy across participants per required action x valence condition.

80

Accuracy

Req. Act.	Go	Go	Go	Go	NoGo	NoGo	NoGo	NoGo
Valence	Win	Win	Avoid	Avoid	Win	Win	Avoid	Avoid
Stakes	High	Low	High	Low	High	Low	High	Low
Mean	0.883	0.893	0.767	0.774	0.523	0.526	0.800	0.813
SD	0.151	0.138	0.125	0.114	0.253	0.276	0.116	0.109

Table S07. Means and standard deviations of accuracy across participants per required action x valence x stakes condition.

81

RTs

Req. Act.	Go	Go	NoGo	NoGo
Valence	Win	Avoid	Win	Avoid
Mean	0.578	0.660	0.641	0.687
SD	0.059	0.062	0.085	0.098

Table S08. Means and standard deviations of reaction times across participants per required action x valence condition.

82

RTs

Req. Act.	Go	Go	Go	Go	NoGo	NoGo	NoGo	NoGo
Valence	Win	Win	Avoid	Avoid	Win	Win	Avoid	Avoid
Stakes	High	Low	High	Low	High	Low	High	Low
Mean	0.585	0.570	0.675	0.645	0.660	0.620	0.695	0.681
SD	0.064	0.063	0.072	0.063	0.098	0.099	0.117	0.122

Table S09. Means and standard deviations of reaction times across participants per required action x valence x stakes condition.

83 **Supplemental Material S05: Correlations with questionnaires**

84 In line with the exploratory analysis plans in mentioned in our pre-registration, we extracted the
85 per-participant coefficients (fixed plus random effects) for (a) the effect of cue valence on responses
86 (Pavlovian bias), (b) the effect of stakes on accuracy, (c) the effect of valence on RTs (Pavlovian bias),
87 and (d) the effect of stakes on RTs. We then computed correlations of these coefficients with forward
88 memory span (Fitzpatrick et al., 2015), backwards memory span, the non-planning subscale of the
89 Barratt Impulsiveness Scale (Patton, Stanford, & Barratt, 1995), and the neuroticism subscale of the
90 neuroticism sub-scale of the Big Five Aspects Scales (DeYoung, Quilty, & Peterson, 2007). One might
91 plausibly hypothesize that impulsivity is related to the Pavlovian bias since many impulsive behaviors
92 can be conceptualized as automatic, cue-triggered behaviors. Hence, individuals high on impulsivity
93 might show stronger Pavlovian biases in responses and reaction times. Furthermore, one might
94 hypothesize that the phenomenon of choking under pressure arises from rumination and worrying, which
95 is typically increased in individuals scoring high on neuroticism (DeCaro, Thomas, Albert, & Beilock,
96 2011). Also, the effects of rumination on performance might be stronger in individuals with a low
97 working memory score (Beilock & Carr, 2005; Bijleveld & Veling, 2014; DeCaro et al., 2011). Hence,
98 individuals high on neuroticism and/or low on working memory span might show stronger effects of
99 stakes on behavior.

100 See Figures S01 and S02 for scatterplots of all bivariate associations. None of the correlations
101 were significant, providing no evidence for the strength of the Pavlovian bias or the effect of stakes on
102 responses and RTs being related to either working memory span, impulsivity, or neuroticism.

103

STAKE MAGNITUDE IN PAVLOVIAN BIASES

6

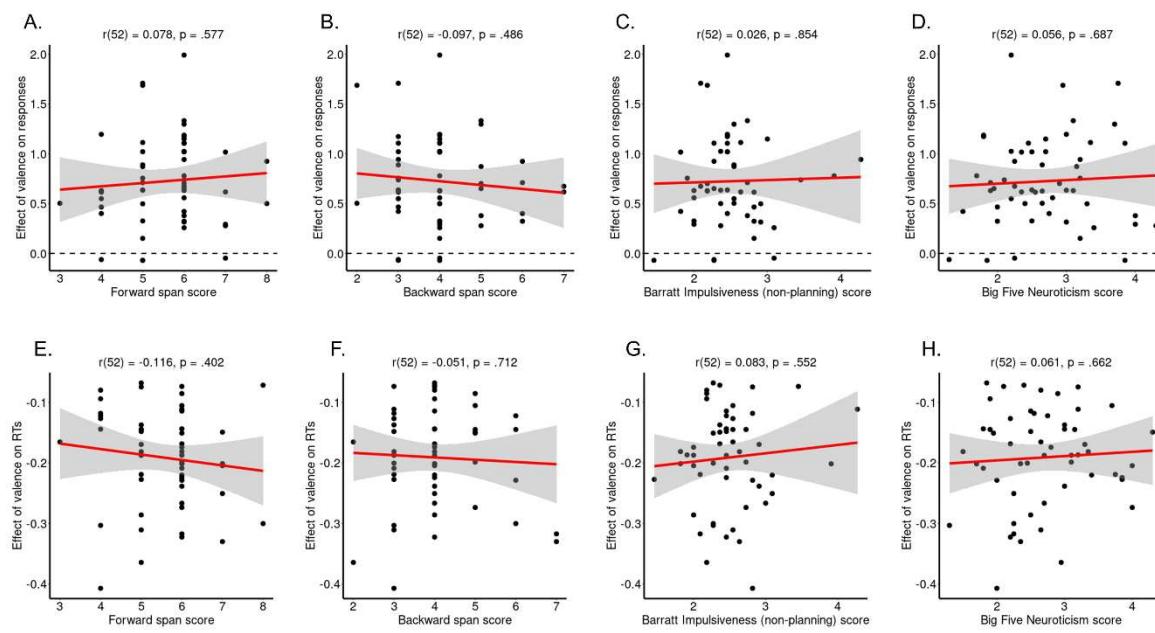


Figure S01. Association of memory performance, impulsivity, and neuroticism with the valence and stakes effects on responses. Correlations between the effect of valence on responses (A–D), reflecting Pavlovian biases, and the effect of stakes on accuracy (E–H) with (A/F) forward working memory span, (B/F) backwards working memory span, (C/G) impulsivity (Barratt Impulsiveness Scale, non-planning subscale) and (D/H) neuroticism. Black dots represent per-participant scores, the red line the best-fitting regression line, they grey shade the 95%-confidence interval. None of the displayed correlations is significant at $\alpha = .05$.

104

105

106

107

108

109

110

111

STAKE MAGNITUDE IN PAVLOVIAN BIASES

7

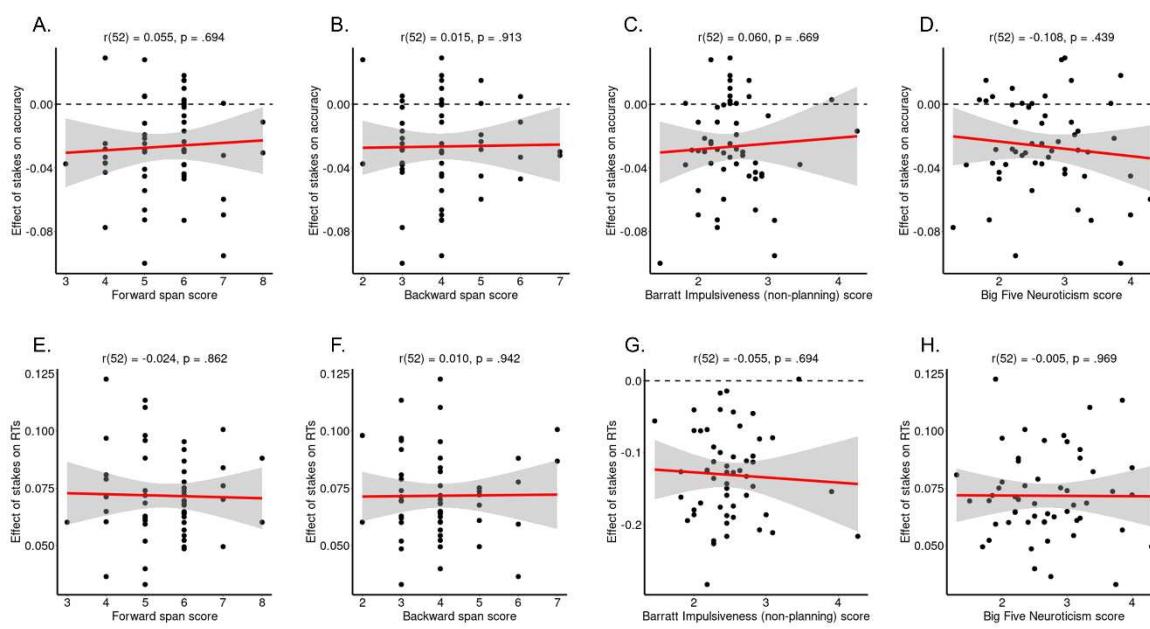


Figure S03. Association of memory performance, impulsivity, and neuroticism with the valence and stakes effects on RTs. Correlations between the effect of valence on RTs (A–D), reflecting Pavlovian biases, and the effect of stakes on RTs (E–H) with (A/F) forward working memory span, (B/F) backwards working memory span, (C/G) impulsivity (Barratt Impulsiveness Scale, non-planning subscale) and (D/H) neuroticism. Black dots represent per-participant scores, the red line the best-fitting regression line, they grey shade the 95%-confidence interval. None of the displayed correlations is significant at $\alpha = .05$.

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126 **Supplemental Material S06: Effect of stakes on RTs over time**

127 In the results in the main text, we report linear associations between time on task (cue repetition,
128 trial number with blocks, trial number across blocks) and reaction time. All associations were non-significant.
129 A more sensitive approach to detect possible non-linear changes over time are so called additive models,
130 which model a time series as a mixture of smooth functions (i.e., thin plate regression splines) for each
131 condition, and allow to test whether (a) a given time series is significant different from a flat line, and (b)
132 whether the time series of different conditions are significantly different from each other (Baayen et al., 2017;
133 Wood, 2017). A smooth function regularizes a raw times series and suppresses high-frequency (i.e., trial-by-
134 trial) noise. Furthermore, it allows for non-zero auto-correlation between residuals, which are assumed to be
135 zero in linear models.

136 In order to test whether the effect of task conditions of stakes on RTs changed over time, we fit three
137 generalized additive mixed-effects models with the z-standardized trial-by-trial RT as dependent variable,
138 modelled as an effect of cue repetition (1–20) with separate time series for (a) each cue condition (Go-to-
139 Win, Go-to-Avoid, NoGo-to-Win, NoGo-to-Avoid), (b) for each stakes condition (high, low), or (c) the
140 interaction between congruency (congruent, incongruent) and stakes (high, low). We modeled the time course
141 of cue repetition as a factor smooth (which has a similar, but potentially non-linear effect as adding a random
142 intercept and a random slope) for each participant for each block, allowing for the possibility that condition
143 differences were different in different participants in different blocks (equivalent to a full random-effects
144 structure). We used a scaled *t*-distribution instead of a Gaussian distribution for the RT variable as it led to
145 lower AIC values. We also investigated whether fit further improved by adding an AR(1) auto-regressive
146 model, which was not the case. For all fitted models, We visually checked that residuals were approximately
147 normally distributed using quantile-quantile plots and whether auto-correlation was near zero using auto-
148 correlation plots (van Rij et al., 2019).

149 The model testing for differences between cue conditions suggested that RTs overall significantly
150 decreased over time in all conditions (see Table S10; Fig. S03A). Further, RTs started to differ between cue
151 conditions from repetition 1 or 2 onwards (see Table S11). Overall, RTs were faster for responses to Win
152 than Avoid cues and faster for (correct) responses to Go cues than (incorrect) responses to NoGo cues.
153 Overall, RT differences between conditions persisted throughout the block.

154 The model testing for differences between stakes levels suggested again that RTs overall
155 significantly decreased over time in both conditions (Table S10; Fig. S03B). Furthermore, throughout the

156 block, RTs were slower for responses on high-stakes trials than for responses on low-stakes trials (Table
157 S11). This difference persisted throughout the block.

158 Finally, the model testing for differences between congruency conditions and stakes levels found
159 again a significant decrease in RTs over time (Table S10; Fig. S03C). RTs were slower for responses to
160 incongruent than to congruent cues, and slower on high-stakes trials than on low-stakes trials. Importantly,
161 RTs were slower on high-stakes trials compared to low-stakes trials both for congruent and for incongruent
162 cues, similarly, although this differences tended to be bigger for incongruent trials. These differences
163 persisted throughout the task.

164 In sum, these results show that condition differences and differences between stakes in RTs emerge
165 on the very first trials (cue repetitions) of a task and persist until the end of a block, with little change in these
166 condition differences.

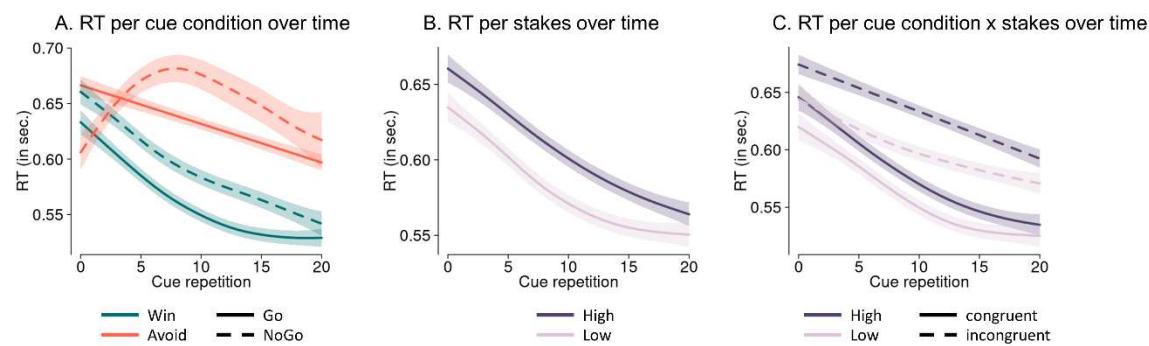


Figure S03. Time course of RTs over cue repetitions within a block as predicted by a generalized additive mixed-effects model, separated by conditions. Overall, RTs speed up over time. **A.** Differences between cue conditions as predicted by the fit of a . RTs are significantly faster for responses to Win than responses to Avoid cues, and faster for (correct) responses to Go cues than (incorrect) responses to NoGo cues throughout a block. **B.** Differences between stakes levels. RTs are significantly slower on high-stakes trials compared to low-stakes trials throughout a block. **C.** Differences between stakes levels separately per condition. Both for congruent and incongruent cues, RTs on high-stakes trials are significantly slower than RTs on low-stakes trials. This difference tends to be larger for incongruent cues.

167

Model	Parametric coefficient (Linear change within condition)	Smooth (non-linear change within each condition)
Cue conditions:		
Go-to-Win	$t(3, 0.103) = 115.249, p < .001$	$F(3.094, 3.788) = 27.370, p < .001$
Go-to-Avoid	$t(3, 0.103) = 22.778, p < .001$	$F(1.000, 1.000) = 35.715, p < .001$
NoGo-to-Win	$t(3, 0.103) = 8.036, p < .001$	$F(2.497, 3.061) = 26.894, p < .001$
NoGo-toAvoid	$t(3, 0.103) = 12.887, p < .001$	$F(2.963, 3.629) = 1.530, p = .107$
Stakes:		
High	$t(3, 0.107) = 122.148, p < .001$	$F(2.266, 2.746) = 35.926, p < .001$
Low	$t(3, 0.107) = -9.346, p < .001$	$F(2.857, 3.478) = 24.505, p < .001$
Congruency x Stakes:		
Congruent/high	$t(3, 0.104) = 108.940, p < .001$	$F(2.426, 2.973) = 30.546, p < .001$
Congruent/low	$t(3, 0.104) = -4.957, p < .001$	$F(2.679, 3.284) = 22.242, p < .001$
Incongruent/high	$t(3, 0.104) = 15.148, p < .001$	$F(1.000, 1.000) = 44.597, p < .001$
Incongruent/low	$t(3, 0.104) = 6.496, p < .001$	$F(1.947, 2.381) = 15.505, p < .001$

Table S10. Results from generalized additive mixed models (GAMMs) with separate smooth per condition. The parametric term reflects a linear change in time, while the smooth terms reflects any non-linear changes. Both add up to the total term.

168

Model	Parametric coefficient (Intercept difference)	Smooth (non-linear differences)	Windows of significant differences
Cue conditions:			
<i>G2W – G2A</i>	$t(3, 0.096) = 23.320, p < .001$	$F(3.779, 4.618) = 5.052, p < .001$	1 – 20
<i>G2W – NG2W</i>	$t(3, 0.081) = 8.383, p < .001$	$F(1.000, 1.000) = 0.606, p = .436$	0 – 20
<i>G2W – NG2A</i>	$t(3, 0.081) = 12.400, p < .001$	$F(3.255, 3.933) = 14.710, p < .001$	2 – 20
<i>G2A – NG2W</i>	$t(3, 0.108) = -9.870, p < .001$	$F(2.587, 3.168) = 5.080, p = .001$	2 – 20
<i>G2A – NG2A</i>	$t(3, 0.112) = 3.234, p = .001$	$F(2.878, 3.476) = 7.412, p < .001$	0 – 2, 5 – 16
<i>NG2W – NG2A</i>	$t(3, 0.098) = 6.939, p < .001$	$F(3.376, 4.042) = 11.760, p < .001$	0 – 1, 3 – 20
Stakes:			
<i>High – Low</i>	$t(3, 0.107) = -9.317, p < .001$	$F(1.424, 1.706) = 1.715, p = .278$	0 – 20
Congruency x Stakes:			
<i>Cong/High – Cong/Low</i>	$t(3, 0.081) = -4.997, p < .001$	$F(1.000, 1.000) = 0.039, p = .844$	0 – 20
<i>Inccong/High – Inccong/Low</i>	$t(3, 0.108) = -8.337, p < .001$	$F(1.000, 1.000) = 0.369, p = .543$	0 – 20
<i>Cong/High – Inccong/High</i>	$t(3, 0.999) = 15.430, p < .001$	$F(1.711, 2.085) = 2.757, p = .061$	0 – 20
<i>Cong/Low – Inccong/Low</i>	$t(3, 0.102) = 11.470, p < .001$	$F(1.000, 1.000) = 5.196, p = .023$	0 – 20

Table S11. Results from generalized additive mixed models (GAMMs) with difference smooths between two conditions. The parametric term reflects a linear difference between conditions, while the smooth terms reflects any non-linear difference. Both add up to the total term. The time window of significant condition differences is automatically returned by the model.

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189
190
191

Supplemental Material S07: Reinforcement-learning drift diffusion models

	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12
WAIC	14501	9284	8011	7029	7023	6848	7082	6996	6821	7025	6843	6682
LOO-IC	14365	8970	7611	6734	6656	6512	6722	6706	6420	6646	6494	6278
α	2.011 [1.97, 2.02]	1.377 [1.372, 1.404]	1.466 [1.449, 1.482]	1.442 [1.428, 1.456]		1.409 [1.397, 1.421]	1.430 [1.417, 1.444]	1.444 [1.430, 1.457]			1.408 [1.396, 1.421]	1.379 [1.367, 1.390]
α_{Low}					1.406 [1.393, 1.420]				1.375 [1.361, 1.389]	1.410 [1.397, 1.424]		
α_{High}					1.479 [1.466, 1.492]				1.429 [1.414, 1.444]	1.475 [1.462, 1.488]		
τ	0.128 [0.119, 0.136]	0.234 [0.226, 0.241]	0.228 [0.220, 0.236]	0.232 [0.224, 0.239]	0.233 [0.226, 0.240]		0.234 [0.227, 0.241]	0.231 [0.224, 0.239]		0.233 [0.225, 0.240]		
τ_{Low}						0.237 [0.230, 0.245]			0.243 [0.236, 0.250]		0.238 [0.231, 0.245]	0.244 [0.237, 0.251]
τ_{High}						0.249 [0.242, 0.256]			0.247 [0.240, 0.254]		0.249 [0.241, 0.256]	
$\tau_{High/Cong}$												0.266 [0.260, 0.273]
$\tau_{High/Incong}$												0.264 [0.256, 0.271]
β	0.061 [0.058, 0.063]	0.259 [0.251, 0.267]		0.251 [0.244, 0.258]	0.250 [0.243, 0.257]	0.264 [0.257, 0.270]		0.249 [0.243, 0.256]	0.266 [0.259, 0.273]	0.250 [0.243, 0.256]	0.264 [0.257, 0.271]	0.277 [0.270, 0.284]
β_{Win}			0.318 [0.308, 0.328]									
β_{Avoid}			0.167 [0.161, 0.172]									
β_{Low}						0.268 [0.262, 0.274]						
β_{High}						0.247 [0.241, 0.253]						
δ_{Int}	3.617 [3.558, 3.675]	1.358 [1.310, 1.407]	1.542 [1.483, 1.602]									
δ_{Win}				2.086 [2.018, 2.152]	2.100 [2.032, 2.167]	2.037 [1.890, 2.105]	2.041 [1.971, 2.112]	2.159 [2.091, 2.228]	2.026 [1.957, 2.094]	2.130 [2.061, 2.199]	2.074 [2.005, 2.142]	1.981 [1.910, 2.053]
δ_{Avoid}					0.796 [0.757, 0.834]	0.803 [0.765, 0.841]	0.736 [0.698, 0.774]	0.763 [0.726, 0.799]	0.867 [0.827, 0.908]	0.727 [0.691, 0.764]	0.831 [0.791, 0.871]	0.774 [0.736, 0.813]
δ_{Slope}		6.823 [6.354, 7.267]	6.283 [5.872, 6.681]	6.093 [5.718, 6.446]	6.149 [5.773, 6.508]	6.191 [5.810, 6.550]	6.102 [5.734, 6.458]	6.109 [5.741, 6.464]	6.218 [5.834, 6.590]	6.151 [5.777, 6.510]	6.219 [5.834, 6.586]	6.273 [5.896, 6.633]
δ_{High}								-0.128 [-0.148, -0.108]		-0.054 [-0.076, -0.033]	-0.075 [-0.096, -0.054]	
ε		0.100 [0.088, 0.110]	0.102 [0.092, 0.112]	0.121 [0.110, 0.131]	0.121 [0.109, 0.131]	0.120 [0.108, 0.130]	0.120 [0.109, 0.130]	0.122 [0.110, 0.132]	0.120 [0.109, 0.130]	0.121 [0.109, 0.131]	0.120 [0.109, 0.131]	0.119 [0.108, 0.129]

Table S12. Mean [25th percentile, 75th percentile] of the posterior densities of group-level parameters. α = decision threshold, τ = non-decision time, β = starting point bias, δ = drift rate, ε = learning rate. WAIC and LOO-IC are reported as measures of model fit, with smaller values indicating a better fit.

192
193
194
195

Posterior densities of group-level parameters M12

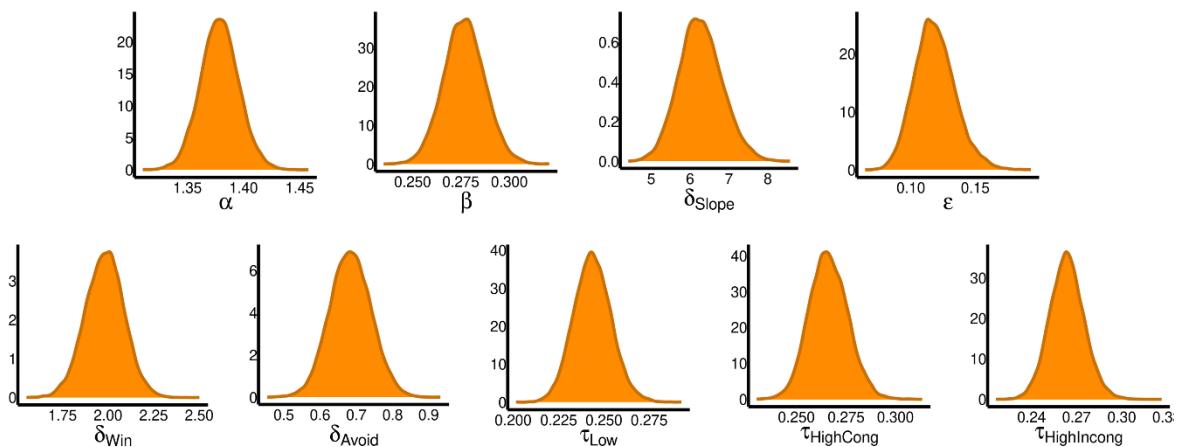


Figure S04. Posterior densities of the group-level parameters of the winning model M12. α = decision threshold, τ = non-decision time, β = starting point bias, δ = drift rate, ϵ = learning rate.

196

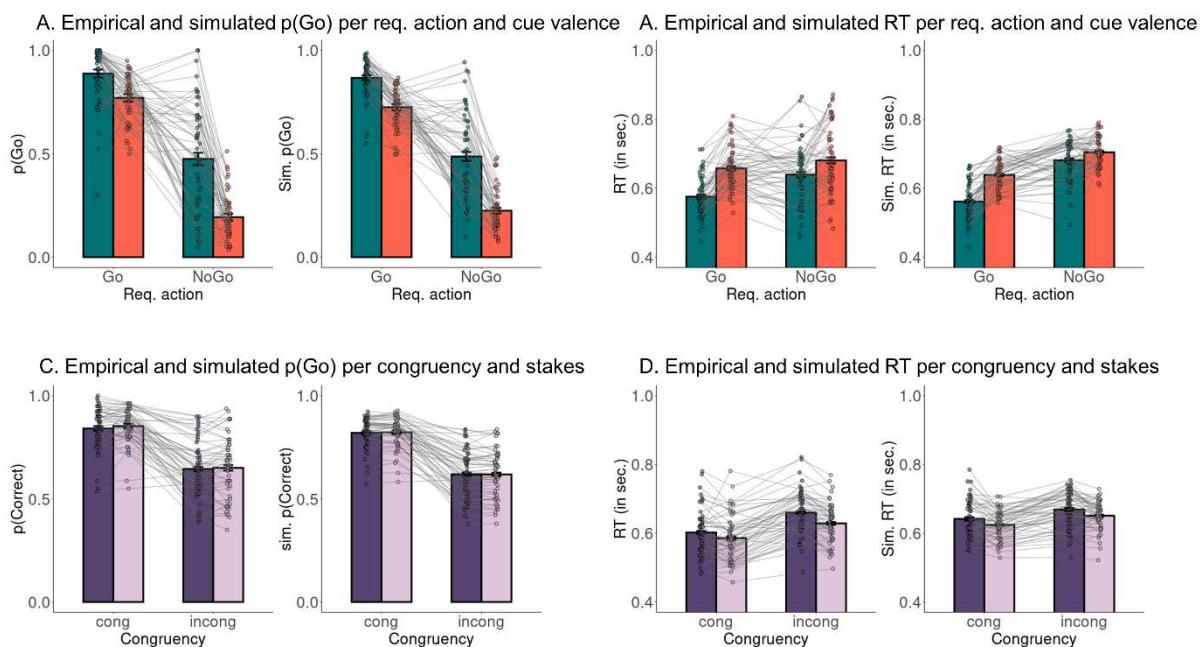


Figure S05. Posterior predictive checks for data simulated from the winning model M12. **A.** Both in empirical data (left panel) and data simulated from the winning model M12 (right panel), (simulated) participants performed more Go responses to Go than NoGo cues (learning) and more Go responses to Win than Avoid cues (Pavlovian bias). Simulated data matched the empirical data pattern. **B.** Both in empirical and simulated data, (simulated) participants showed faster responses to Go than NoGo cues and to Win than Avoid cues. Simulated data matched the empirical data pattern. **C.** Both in empirical and simulated data, (simulated) participants performed more accurately for congruent than incongruent cues, with no difference between high and low stakes. **D.** Both in empirical and simulated data, (simulated) participants performed faster for congruent than incongruent cues and under low compared to high stakes. In empirical participants, the stakes effect was stronger for incongruent than congruent cues, but this difference was somewhat underestimated by the winning model M12.

197

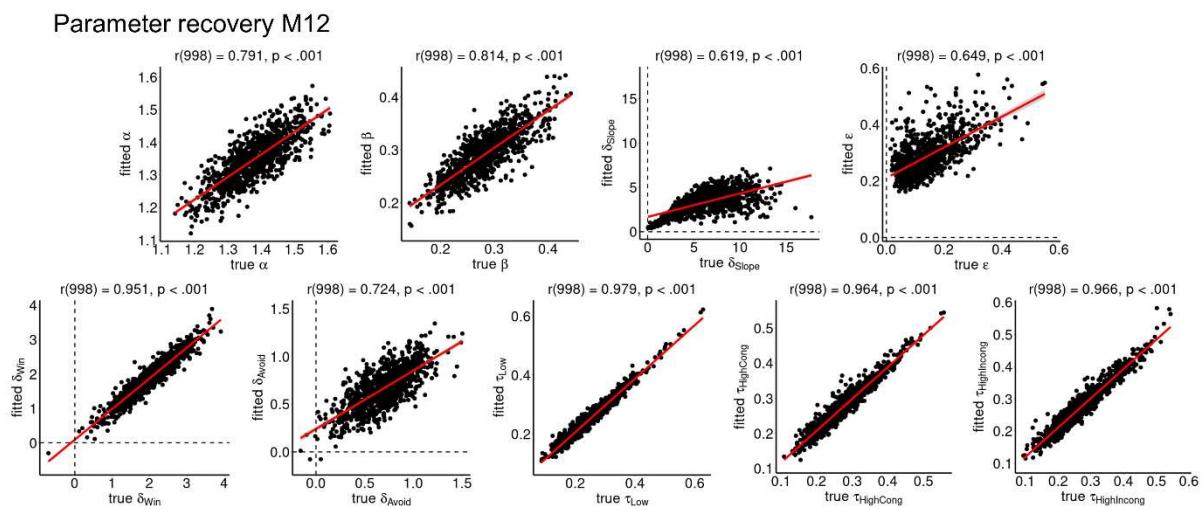


Figure S06. Parameter recovery results for the winning model M12. The correlation between generative and fitted parameters is overall very high. Recovery is overall very high. It is least optimal (but still strongly significant) for δ_{Slope} and ϵ , which trade off against each other (see Fig. 4D main text). α = decision threshold, τ = non-decision time, β = starting point bias, δ = drift rate, ϵ = learning rate.

198

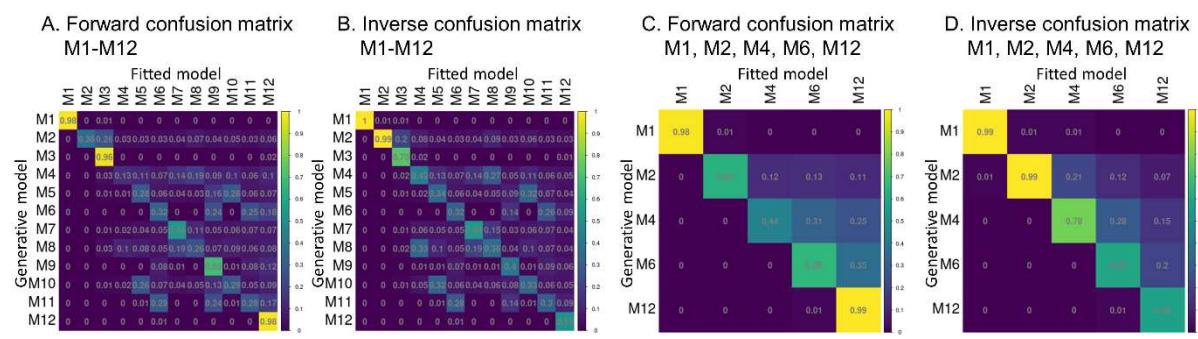


Figure S07. Forward and inverse confusion matrices from model recovery of all models and of nested sub-versions of the winning model M12. **A.** The forward confusion matrix displays the conditional probabilities that model Y is the best fitting model (columns) if model X (rows) is the underlying generative model used to simulate a given data set (identical to Fig. 4E main text). Rows sum to 100%. On-diagonal probabilities indicate the probability of reidentifying the generative model. All on-diagonal probabilities are significantly above chance (range 0.13–0.98; 95th percentile of permutation null distribution: $p = 0.10$). Especially recovery for M12 is exceptionally high (98%). **B.** The inverse confusion matrix displays the conditional probabilities that model X is the generative model (rows) if model Y (rows) is the best fitting model for a given data set. Columns sum to 100%. On-diagonal probabilities indicate the probability of reidentifying the generative model. All on-diagonal probabilities are significantly above chance (range 0.30–1.00; 95th percentile of permutation null distribution: $p = 0.10$). **C.** Forward confusion matrix only for the five models that are nested sub-versions of M12 (i.e., M1, M2, M4, M6, M12). Recovery is overall much higher (range 0.44–0.99; 95th percentile of permutation null distribution: $p = 0.22$). **D.** Inverse confusion matrix only for the five models that are nested sub-versions of M12. Recovery is overall much higher (range 0.58–0.99; 95th percentile of permutation null distribution: $p = 0.22$).

199